Research Article

Modeling Illicit Drug Use Dynamics and Its Optimal Control Analysis

Steady Mushayabasa and Gift Tapedzesa

University of Zimbabwe, Department of Mathematics, P.O. Box MP 167, Harare, Zimbabwe

Correspondence should be addressed to Steady Mushayabasa; steadymushaya@gmail.com

Received 20 April 2015; Accepted 13 October 2015

Academic Editor: Seiya Imoto

Copyright © 2015 S. Mushayabasa and G. Tapedzesa. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The global burden of death and disability attributable to illicit drug use, remains a significant threat to public health for both developed and developing nations. This paper presents a new mathematical modeling framework to investigate the effects of illicit drug use in the community. In our model the transmission process is captured as a social "contact" process between the susceptible individuals and illicit drug users. We conduct both epidemic and endemic analysis, with a focus on the threshold dynamics characterized by the basic reproduction number. Using our model, we present illustrative numerical results with a case study in Cape Town, Gauteng, Mpumalanga and Durban communities of South Africa. In addition, the basic model is extended to incorporate time dependent intervention strategies.

1. Introduction

Illicit drug use continues to exert significant toll, with valuable human lives and productive years of many people being lost. An estimated 183,000 drug-related deaths were reported in 2012 [1]. Globally, it is estimated that, in 2012, between 162 million and 324 million people aged between 15 and 64 had used an illicit drug [1]. Illicit drug use is defined as the non-medical use of a variety of drugs that are prohibited by international law [2]. These drugs include amphetamine-type stimulants, cannabis, cocaine, heroin, and other opioids, and MDMA (ecstasy) [3]. The risk of premature mortality and morbidity from illicit drug use is dependent on dose, frequency, and route of administration [2]. Further, the mortality risks of illicit drug consumption increase with increasing frequency and quantity of consumption [4]. The first global burden of death and disability attributable to illicit drug use was first estimated by Donoghoe [5], as part of the global burden of disease project [6]. Donoghoe estimated that illicit drug use was responsible for 10,000 deaths globally in 1990 with about 62% of them in developing countries [6]. Besides causing deaths, illicit drug use has been associated with 50% of mental illness cases [7]. Table 1 illustrates the cumulative prevalence of illicit drug use in Cape Town, South Africa, from 1996 (June) to 2008 (June); the data was adopted from SACENDU [8].

Additional data on recorded cases of illicit drug use in some communities of South Africa, namely, Durban (Table 4), Gauteng (Table 5), and Mpumalanga (Table 6), is presented in Appendix A.

Despite many clinical and theoretical studies [9–20] and tremendous educational campaigns [1, 7, 21, 22], illicit drug use remains a significant threat to public health all over the world. The use of mathematical models to explore the dynamics of drug use has been an interesting topic for a couple of researchers [11–13]. In 2007, White and Comiskey [13] proposed a mathematical model to evaluate the role of treatment and relapse in the dynamics of heroin. Their work revealed among others that relapse of individuals who would have quit heroin use has a significant impact on the generation of new or secondary heroin users. Most recently, Nyabadza et al. [11] constructed a mathematical model to examine the dynamics of crystal meth “Tik” abuse in the presence of drug-supply chains. Using the data from South Africa, their work suggests among others that programs aimed at encouraging light drug users to quit drug use can be more effective to control “Tik” abuse in South Africa.
Our objective in this paper is to formulate a mathematical model for illicit drug use that includes relevant biological and social aspects, accounts for case detection of drug users, and allows optimal control methods to be used. We stress that our model is not for a specific substance abuse. To begin with, we integrate the aforementioned essential components into one SIR-type (Susceptible-Infectious (individuals who are drug users)-Recovered (individuals in rehabilitation)) model to accommodate the diverse dynamics of illicit drug use determined by population-specific parameters such as the rates of light drug user to heavy drug user, light drug user to mental illness, and heavy drug user to mental illness. Our study differs from those in the literature in the fact that mental illness can lead to drug abuse. However, in this paper, we consider that individuals become mentally ill due to illicit drug use. It is worth noting that the stigma attached to substance abuse and mental disorders often hinders early detection, diagnosis, and proper treatment [7]. We are aware that mental illness can lead to drug abuse. However, in this paper, we consider that individuals become mentally ill due to illicit drug use.

After comprehensive mathematical analysis, we extend the basic model to incorporate two control functions: the goal of the first control function $u(t)$ is to reduce the intensity of “social influence,” that is, to weaken the intensity of social interaction between the susceptible population and illicit drug users, while the second $v(t)$ attempts to increase the rate of detection and rehabilitation of illicit drug users. Optimal control theory has found wide-ranging applications in solving biological problems [23] and network structures [24]. Numerical results followed by a brief discussion round up the paper.

### 2. Mathematical Model

#### 2.1. Model Framework

A population of size $N(t)$ is partitioned into subclasses of susceptible individuals $S(t)$ (individuals who are not yet illicit drug users but interact with drug users), light or occasional drug users $I(t)$, heavy drug users $I_a(t)$, mentally ill population $M(t)$ (individuals who suffer mental illness due to drug use), and detected illicit drug users $R(t)$. Thus, $N(t) = S(t) + I(t) + I_a(t) + M(t) + R(t)$. We assume a constant size population with recruitment and non-illicit-related death rate at time $t$ given by $\mu$. Susceptible individuals acquire illicit drug use habits at rate

$$\lambda = \beta (I + \kappa I_a).$$

The parameter $\beta$ measures the strength of interactions between the susceptible individuals and illicit drug users, that is, the “influence” of $I$ and $I_a$ on $S$; $\kappa \geq 1$ is a modification factor which accounts for the increased likelihood of heavy illicit drug users $I_a$ to influence more new drug users compared to light drug users $I$. The model takes the form

$$\begin{align*}
\dot{S} &= \mu - \lambda S - \mu S, \\
\dot{I} &= \lambda S - (\alpha + \gamma + \sigma + \mu + \psi) I, \\
\dot{I}_a &= \alpha I - (\rho + \phi + \mu + d) I_a, \\
\dot{M} &= \sigma I + \phi I_a - (\epsilon + \mu + \delta) M, \\
\dot{R} &= \gamma I + \rho I_a + \epsilon M - (\mu + \omega) R,
\end{align*}$$

(2)

where the upper dot represents the derivative of the component with respect to time. The parameter $\alpha$ is the rate at which light drug users become heavy drug users; $\gamma$, $\epsilon$, and $\rho$ denote the rates of detection and rehabilitation of individuals in classes $I$, $M$, and $I_a$, respectively; $\sigma$ and $\phi$ ($\sigma \leq \phi$) denote the rates at which light and heavy illicit drug users develop mental illness; $\psi$ and $d$ model the permanent exit rates of light and heavy users, respectively, due to either cessation or drug use-related death. Individuals in rehabilitation recover at rate $\omega$ and are assumed to permanently exit the model. Mentally ill individuals permanently exit the model at rate $\delta$ due to drug use-related death. Further, we assume that the mentally ill population does not influence the susceptible individuals to become illicit drug users. We study system (2) in the closed set:

$$\Omega = \left\{(S, I, I_a, M, R) \in \mathbb{R}_+^5 : 0 \leq N \leq 1 \right\},$$

(3)

where $\Omega$ is positively invariant with respect to (2). In the absence of drug abuse in the community system (2) admits a drug-free equilibrium point given by

$$\mathbf{E}^0 = \left(S^0, I^0, I^0_a, M^0, R^0\right) = (1, 0, 0, 0, 0).$$

(4)

The basic reproductive number provides an invasion criterion for the initial spread of the drug misuse in a susceptible population. For this case, the reproductive number is given by (see computations in Appendix B)

$$R_d = \frac{\beta (\alpha k + \mu + \rho + \phi + d)}{(\alpha + \gamma + \mu + \sigma + \psi)(\mu + \rho + \phi + d)}.$$

(5)

The threshold quantity $R_d$ gives the average number of new drug abusers generated by a typical light or heavy drug user during his or her lifetime as an illicit drug user.

Theorem 1 follows from computations in Appendix C.

**Theorem 1.** The drug-free equilibrium point $\mathbf{E}^0$ is globally asymptotically stable if $R_d \leq 1$. If $R_d > 1$, then a unique drug persistence equilibrium point $\mathbf{E}^*$ of system (2) exists and is globally asymptotically stable.
Figure 1: Model system (2) fitted to data for individuals seeking treatment due to illicit drug use. The blue circles indicate the actual data and the solid line indicates the model fit to the data.

Table 2: Model parameters and their baseline values obtained after fitting data in Appendix A.

<table>
<thead>
<tr>
<th>Parameter definition</th>
<th>Symbol</th>
<th>Range</th>
<th>Baseline value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recovery rate</td>
<td>( \omega )</td>
<td>0.1–0.9</td>
<td>0.3</td>
<td>Per year</td>
</tr>
<tr>
<td>Recruitment rate</td>
<td>( \mu )</td>
<td>0.02–0.03</td>
<td>0.02</td>
<td>Per year</td>
</tr>
<tr>
<td>Modification factor</td>
<td>( \kappa )</td>
<td>1–1.85</td>
<td>1.25</td>
<td>—</td>
</tr>
<tr>
<td>Transmission rate</td>
<td>( \beta )</td>
<td>0.31–0.36</td>
<td>0.35</td>
<td>Per year</td>
</tr>
<tr>
<td>Detection rate for light drug users</td>
<td>( \gamma )</td>
<td>0.01–0.14</td>
<td>0.1</td>
<td>Per year</td>
</tr>
<tr>
<td>Detection rate for heavy drug users</td>
<td>( \rho )</td>
<td>0.13–0.78</td>
<td>0.35</td>
<td>Per year</td>
</tr>
<tr>
<td>Detection rate for mentally ill population</td>
<td>( \epsilon )</td>
<td>0.54–0.85</td>
<td>0.6</td>
<td>Per year</td>
</tr>
<tr>
<td>Escalation of a light user to a heavy user</td>
<td>( \alpha )</td>
<td>0.01–0.75</td>
<td>0.01</td>
<td>Per year</td>
</tr>
<tr>
<td>Drug use-related death for light drug users</td>
<td>( \psi )</td>
<td>0.01–0.05</td>
<td>0.035</td>
<td>Per year</td>
</tr>
<tr>
<td>Drug use-related death for heavy drug users</td>
<td>( \delta )</td>
<td>0.1–0.18</td>
<td>0.14</td>
<td>Per year</td>
</tr>
<tr>
<td>Drug use-related death for mentally ill population</td>
<td>( d )</td>
<td>0.077–0.23</td>
<td>0.2</td>
<td>Per year</td>
</tr>
<tr>
<td>Proportion of light users who develop mental illness</td>
<td>( \sigma )</td>
<td>0.10–0.33</td>
<td>0.05</td>
<td>Per year</td>
</tr>
<tr>
<td>Proportion of heavy users who develop mental illness</td>
<td>( \phi )</td>
<td>0.08–0.09</td>
<td>0.09</td>
<td>Per year</td>
</tr>
</tbody>
</table>

2.2. Numerical Results. In this section, we estimate the model parameters used in our numerical simulations. In order to carry out the estimation, we made use of the data from SACENDU (see Appendix A). The data has also been used to estimate the best fit for our model.

Figure 1 shows “best” fit to the real data, illustrated in Appendix A. Here, the fitting process involved the use of the least squares-curve fitting method. A Matlab code has been used together with unknown parameters assigned lower and upper bound from which a set of parameter values that produce the best fit have been obtained. With the exception of Gauteng, simulation results in Figure 1 suggest that the prevalence of illicit drug use increases rapidly over the defined time intervals, while for Gauteng, the increase is gradual. Using the lower bounds of parameter values in Table 2, direct calculation shows that \( R_\infty = 6.22 \), which shows that effective methods should be designed to reduce illicit drug use in these communities. Further, from these simulations, one can suggest that system (C.3) can be an essential tool in predicting future illicit drug use in the communities.

2.2.1. Sensitivity Analysis of the Reproductive Number. Sensitivity analysis of model parameters is very important to design and control strategies as well as a direction to future research. Local sensitivity indices allow us to measure the relative change in a state variable when a parameter changes.
In computing the sensitivity analysis, we adopt the approach described by Arriola [25]. The normalized forward sensitivity index of a variable to a parameter is the ratio of the relative change in the variable to the relative change in the parameter. When the variable is a differentiable function of the parameter, the sensitivity index may be alternatively defined using partial derivatives.

**Definition 2.** The normalized forward sensitivity index of a variable, \( u \), that depends differentiably on a parameter, \( p \), is defined as

\[
\Gamma_p^u = \frac{\partial u}{\partial p} \times \frac{p}{u},
\]

(6)

To determine the numerical output for our sensitivity indices, we used the lower bounds of parameters in Table 2. Table 3 presents the model parameters and their sensitivity indices, we used the lower bounds of parameters in Table 2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Sensitivity index</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta )</td>
<td>+1</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>-0.067</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>0.032</td>
</tr>
<tr>
<td>( \mu )</td>
<td>-0.38</td>
</tr>
<tr>
<td>( \rho )</td>
<td>-0.013</td>
</tr>
<tr>
<td>( d )</td>
<td>-0.008</td>
</tr>
<tr>
<td>( f )</td>
<td>-0.0075</td>
</tr>
<tr>
<td>( \psi )</td>
<td>-0.19</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>-0.059</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>-0.19</td>
</tr>
<tr>
<td>( \phi )</td>
<td>-0.0082</td>
</tr>
</tbody>
</table>

Table 3: Sensitivity indices of \( R_d \) to parameters for model (2), evaluated at the baseline parameter values given in Table 2.

The objective functional, \( J \), is used to formulate the relevant optimization problem: finding the most effective strategy that reduces or eliminates the levels of illicit drug use in the community at minimal cost. This minimization goal will be achieved through the implementation of controls \( u(t) \) and \( v(t) \) over the preselected time interval \([0, T]\). Mathematically, this corresponds to the minimization of the functional \( J \) over a set of feasible, \( u(t) \) and \( v(t) \), strategies on \([0, T] \). \( J \) functional is defined as follows:

\[
J(u, v) = \int_0^T \left[ I + I_u^2 + B_v^2 \right] dt,
\]

(8)

where \( A \) and \( B \) are balancing coefficients transforming the integral into dollars expended over a finite time period of \( T \) years. The balancing coefficients account for the relative size and importance preassigned by the modelers to the contributing terms in the objective functional.

We consider state system (7) of ordinary differential equations in \( \mathbb{R}^5 \) with the set of admissible control functions given by

\[
\Gamma = \{(u(t), v(t)) \in L^1(0, T) \mid 0 \leq u(t), v(t) \leq 1, \forall t \in [0, T]\}.
\]

We consider the optimal control problem of determining \( S^*(t) \), \( I^*(t) \), \( I_u^*(t) \), \( M^*(t) \), and \( R^*(t) \), associated with an admissible control pair \((u^*(t), v^*(t)) \in \Gamma \) on the time interval \([0, T]\), satisfying (7), given the initial conditions \( S(0) \), \( I(0) \), \( I_u(0) \), \( M(0) \), and \( R(0) \), and minimizing the cost functional (8); that is,

\[
J(u^*(t), v^*(t)) = \min_I J(u(t), v(t)).
\]

The existence of optimal controls follows from standard results in optimal control theory [26]. Theorem 3 follows from Appendix C.

**Theorem 3.** Problem (7)–(10) with given initial conditions \( S(0), I(0), I_u(0), M(0), \) and \( R(0) \) and fixed final time \( T \) admits a unique optimal solution \((S^*(t), I^*(t), I_u^*(t), M^*(t), R^*(t))\) associated with an optimal control pair \((u^*, v^*)\) on \([0, T]\).

The optimal control pair predicted by (5) represents the optimal intervention strategy, given cost constraints, and can be found by the application of the Pontryagin maximum principle [26].

2.3.1. Optimal Control Numerical Results. State system (7) and the adjoint system of differential equations together with the control characterization above from the optimality system are solved numerically using the "forward-backward sweep method (FBSM)." For a detailed discussion on FBSM, we refer the reader to [23]. We set \( S(0) = 0.97 \), \( I(0) = 0.02 \),
Table 4: Recorded prevalence of illicit drug use for Durban, South Africa.

<table>
<thead>
<tr>
<th>Year</th>
<th>96b</th>
<th>97a</th>
<th>97b</th>
<th>98a</th>
<th>98b</th>
<th>99a</th>
<th>99b</th>
<th>00a</th>
<th>00b</th>
<th>01a</th>
<th>01b</th>
<th>02a</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>27</td>
<td>31</td>
<td>34</td>
<td>35</td>
<td>33</td>
<td>38</td>
<td>35</td>
<td>40</td>
<td>40</td>
<td>41</td>
<td>42</td>
<td>43</td>
</tr>
</tbody>
</table>

Table 5: Recorded prevalence of illicit drug use for Gauteng, South Africa.

<table>
<thead>
<tr>
<th>Year</th>
<th>96b</th>
<th>97a</th>
<th>97b</th>
<th>98a</th>
<th>98b</th>
<th>99a</th>
<th>99b</th>
<th>00a</th>
<th>00b</th>
<th>01a</th>
<th>01b</th>
<th>02a</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>31</td>
<td>32</td>
<td>33</td>
<td>37</td>
<td>40</td>
<td>40</td>
<td>46</td>
<td>48</td>
<td>46</td>
</tr>
</tbody>
</table>

Table 6: Recorded prevalence of illicit drug use for Mpumalanga, South Africa.

<table>
<thead>
<tr>
<th>Year</th>
<th>96b</th>
<th>97a</th>
<th>97b</th>
<th>98a</th>
<th>98b</th>
<th>99a</th>
<th>99b</th>
<th>00a</th>
<th>00b</th>
<th>01a</th>
<th>01b</th>
<th>02a</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>24</td>
<td>24</td>
<td>29</td>
<td>23</td>
<td>30</td>
<td>31</td>
<td>29</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$I_4(0) = 0.01$, and $M(0) = R(0) = 0$. Parameters used are in Table 2.

Numerical results in Figure 2 illustrate the role of optimal intervention strategies in controlling illicit drug use in the community. In Figure 2, the balancing coefficients are fixed within the integral expression (8) and the optimal schedule of the two controls over $T = 20$ years is simulated with the following assumed initial population levels: $S = 0.97, I = 0.02, I_4 = 0.01$, and $M = R = 0$. System (7) is used to determine the resultant population dynamics. From the illustrations, it is clear that optimal intervention strategies provide a more effective approach on the elimination or reduction of illicit drug use and mental health cases. In Figure 2(a), we note that in the presence of two optimal intervention strategies it can require 14 years for the population of light illicit drug users to die off. Although interventions which are not time dependent lead to a decrease in cases of light illicit drug use, implemented for 14 years, they may not yield cessation in illicit drug use cases.

In Figure 2(b), we set the initial population level of heavy drug users at 0.01 (1%) of the total population. Although this might be a high figure in reality, it allows us to observe the level of effectiveness of optimal intervention strategies. It is evident in Figure 2 that the presence of the aforementioned optimal intervention strategies can lead to elimination of heavy drug users even if their population is considerably high.

In Figures 2(c) and 2(d), the initial densities of mentally ill and detected population were set to zero. Results here show that in the presence of time dependent interventions mental illness associated with illicit drug use will remain low and will eventual die off after 14 years of implementing the aforementioned optimal intervention strategies. Further, it is evident that more illicit drug users will be detected when time dependent interventions are implemented.

Simulations in Figure 3 illustrate the feasibility of the control functions $u(t)$ and $v(t)$. If the control profile starts at the upper bound and drops on the final time to the origin, then it is regarded as highly feasible, implying that more effort and resources should be devoted to such an intervention strategy for effective problem-solving or control. Figure 3(a) assumes that parameter $A$ associated with control function $u(t)$ is greater than parameter $B$ associated with control function $v(t)$. This implies that the cost associated with the implementation of control $u(t)$ is greater than the one associated with control $v(t)$. For this case, we observe that Figure 3(a) suggests that a slightly higher effort should be devoted to control $v(t)$, since it is more feasible compared to control $u(t)$. In Figure 3(b), it is evident that if the cost associated with control $v(t)$ is higher than the cost associated with control $u(t)$, then all the optimal intervention strategies are highly feasible.

2.3.2. Efficacy of Optimal Intervention Strategy. The efficacy of an intervention strategy on controlling the generation of new illicit drug users reflects the strength of the strategy to effectively control the drug use problem in the community. In this section, we explore the effectiveness of the aforementioned optimal intervention methods on reducing cumulative population illicit drug users. We define the efficacy function $E(t)$

$$E(t) = 1 - \left[ \frac{I^*(t) + I^*_n(t)}{I(0) + I_n(0)} \right],$$

where $I^*(t)$ and $I^*_n(t)$ denote the optimal solutions associated with the optimal control of the corresponding variable and $I(0)$ and $I_n(0)$ denote the corresponding initial condition. Function (11) measures the proportional decrease in the number of active illicit drug users imposed by the intervention with controls $(u, v)$, by comparing the number of active illicit
drug users at time $t$ with the initial conditions for which there are no controls implemented ($u = v = 0$). By construction, $E(t) \in [0, 1]$ for all time $t$. Thus, the upper bound of $E(t)$ is one.

Figure 4 demonstrates that optimal interventions may be effective to eliminate illicit drug use in the community after 14 years of implementation. It is worth noting that, after 4 years of implementing these methods, the efficacy level would reach 80% mark; this demonstrates that even for a short time horizon optimal intervention strategies can have a significant impact on controlling illicit drug use in the community.

**3. Concluding Remarks**

About 5% (230 million) of the people of the world’s adult population are estimated to have used an illicit drug at least once in 2010 [21]. In this paper, a simple mathematical model to assess the impact of illicit drug use and its adverse health effects is proposed and analyzed. The model is further extended to incorporate two optimal intervention strategies. The goal of the first optimal intervention is tied on reducing the “social influence” between the susceptible and illicit drug users, while the second aims to increase the detection and rehabilitation of illicit drug users. From the illustrations in...
this study it is clear that time dependent intervention strategies can lead to elimination of illicit drug use in the community. Our analysis also suggests that the implementation of these controls should be devoted to both controls since they are highly feasible for a longer period considered in this study. Although the model is simple and has a couple of simplifying assumptions, qualitative conclusions can be reached in rather broad terms from the simulations and analysis, the kind of conclusions that can generate or provide useful insights into value and effectiveness of time dependent control efforts, aimed at elimination or effective control of illicit drug use and its adverse health problems.

Our model is not exhaustive; a new model that incorporates illicit drug use and age or gender can be developed and used to forecast future trends on illicit drug use.

Appendices

A. Recorded Data on Illicit Drug Use

In this section, we present cumulative prevalence of illicit drug use in South Africa, for the following areas: Durban, Gauteng, and Mpumalanga. The data was adopted from Pluedemann et al. (2008) [8]. We obtained cumulative prevalence after summing the prevalences for the following illicit drugs: cannabis, Mandrax, cocaine, heroin, ecstasy, cocaine, over-the-counter and prescription medicines (OTC/PRE), methamphetamine, and other substances/polysubstance abuse. Key: 96b represents data recorded in 1996 from July to December and 97a represents January to June 1997.

B. Computation of the Reproductive Number

Following Van den Driessche and Watmough [27] and using the notation defined therein, the matrices $F$ and $V$ for the new infection terms and the remaining transfer terms are, respectively, given by

$$F = \begin{pmatrix} \beta & \beta k \\ 0 & 0 \end{pmatrix},$$

$$V = \begin{pmatrix} \mu + \alpha + \gamma + \sigma & 0 \\ -\alpha & \mu + \rho + \phi + d \end{pmatrix}. \quad (B.1)$$

Thus,

$$FV^{-1} = \begin{pmatrix} \frac{\beta (\alpha k + \mu + \rho + \phi + d)}{\alpha + \beta + \gamma + \mu + \sigma + \psi} & \frac{\beta}{\mu + \rho + \phi + d} \\ 0 & 0 \end{pmatrix}. \quad (B.2)$$

Therefore, spectral radius is

$$R_d = \frac{\beta (\alpha k + \mu + \rho + \phi + d)}{\alpha + \beta + \gamma + \mu + \sigma + \psi} (\mu + \rho + \phi + d). \quad (B.3)$$

C. Equilibrium States and Their Stability


To explore the global stability of $E_0$, we consider the Lyapunov function:

$$L(t) = \left( \frac{\alpha \beta + k}{2} \right) I(t) + \frac{\alpha \beta}{k} I_a(t). \quad (C.1)$$

Its Lyapunov derivative along the solutions to system (2) is

$$\dot{L} = \left( \frac{\alpha \beta + k}{2} \right) \dot{I} + \frac{\alpha \beta}{k} \dot{I_a} = \frac{\alpha \beta}{k} (I + \beta I_a) \leq \frac{\alpha \beta}{k} (I + \beta I_a) = \frac{(I + \beta I_a)(R_d - 1)}{R_d - 1} \quad (C.2)$$

with $k_1 = (\mu + \alpha + \gamma + \sigma)$ and $k_2 = (\mu + \rho + \phi + d)$. Thus, $L \leq 0$ if $R_d \leq 1$ with $L = 0$ if and only if $I = I_a = 0$. Therefore, by LaSalle’s Invariance Principle [28], every solution to system (2), with initial conditions in $\Omega$, approaches $E_0$ as $t \to \infty$. That is, $(I, I_a) \to (0, 0)$ as $t \to \infty$. Substituting $I = I_a = 0$ in (2) gives $\dot{\theta}^d$ as $t \to \infty$. Thus, $\theta^d$ is globally asymptotically stable in $\Omega$ whenever $R_d \leq 1$.


Solving system (2) gives

$$S^* = \frac{1}{R_d},$$

$$I^* = \frac{\mu}{k_1 R_d} (R_d - 1),$$

$$M^* = \frac{\mu (\sigma k_2 + \alpha \phi)}{k_1 k_2 k_3 R_d} (R_d - 1),$$

$$R^* = \frac{\mu (\sigma k_2 + \alpha \phi)}{k_1 k_2 k_3 k_4} (R_d - 1),$$

with $k_3 = e + \mu + \delta$, $k_4 = \mu + \omega$. \quad (C.3)
Theorem C.1. If $R_d > 1$, system (2) is uniformly persistent; namely, there exists a constant $\xi > 0$ such that
\[
\lim \inf_{t \to +\infty} S(t) > \xi,
\]
\[
\lim \inf_{t \to +\infty} I(t) > \xi,
\]
\[
\lim \inf_{t \to +\infty} I_a(t) > \xi;
\]
here, $\xi$ is independent with initial data in $\Omega$.

Proof. When $+\infty$, from system (2), we have the following limiting system:
\[
\dot{S} = \mu - \beta (I + k_1 S) S - \mu S,
\]
\[
\dot{I} = \beta (I + k_1 S) S - k_1 I,
\]
\[
\dot{I}_a = \alpha I - k_2 I_a.
\]

Observe that we have omitted the last two equations in (2) since the first three equations of system (2) are independent of the variables $M$ and $R$. Further, for ease of notation, we will continue to denote the drug-free equilibrium point of system (C.5) as $\mathcal{E}_0$. Define
\[
X = \{(S, I, I_a) : S \geq 0, I \geq 0, I_a \geq 0\},
\]
\[
X_0 = \{(S, I, I_a) : S \geq 0, I \geq 0, I_a \geq 0\},
\]
\[
\partial X_0 = \partial \Omega.
\]

It then suffices to show that (C.5) is uniformly persistent with respect to $(X, \partial X_0)$. Firstly, from (C.5), it can be verified that both $X$ and $X_0$ are positively invariant. Clearly, $\partial X_0$ is relatively closed in $X$ and system (C.5) is point dissipative. Consider the following set using solutions $(S(t), I(t), I_a(t))$ of system (C.5):
\[
G_0 = \{(S(0), I(0), I_a(0)) : (S(t), I(t), I_a(t)) \in \partial X_0, \forall t \geq 0\}.
\]

Now we prove that
\[
G_0 = \{(S, 0, 0) : S \geq 0\}.
\]

Assume that $(S(0), I(0), I_a(0)) \in G_0$. It suffices to show that $I(t) = I_a(t) = 0$ for all $t \geq 0$. If not, then there exists $t_0 \geq 0$ such that $(I(t_0) > 0, I_a(t_0) = 0)$ or $(I(t_0) = 0, I_a(t_0) > 0)$. For $(I(t_0) > 0, I_a(t_0) = 0)$, we have
\[
\dot{I_a} = \alpha I(t_0) > 0.
\]

It follows that there is $\varepsilon_0 > 0$ such that $I_a(t) > 0$ for $t_0 < t < t_0 + \varepsilon_0$. Clearly, we can restrict $\varepsilon_0 > 0$ small enough such that $I(t) > 0$ for $t_0 < t < t_0 + \varepsilon_0$. This means that $(S(t), I(t), I_a(t)) \notin \partial X_0$ for $t_0 < t < t_0 + \varepsilon_0$, which contradicts the assumption that $(S(0), I(0), I_a(0)) \in G_0$. For other cases, it can also be easily verified that they contradict the assumption that $(S(0), I(0), I_a(0)) \in G_0$. Thus, (C.8) holds.

Note that $\mathcal{E}_0$ is globally asymptotically stable in $\text{Int} \mathcal{G}_0$. Moreover, $\mathcal{E}_0$ is an isolated invariant set in $X$, every orbit in $\mathcal{G}_0$ converges to $\mathcal{E}_0$, and $\mathcal{E}_0$ is acyclic in $\mathcal{G}_0$. By Theorem 4.6 in [29], we only need to show that $W^s(\mathcal{E}_0) \cap X_0 = \emptyset$ if $R_d > 1$.

Now, we prove that $W^s(\mathcal{E}_0) \cap X_0 = \emptyset$. Then, there exists a positive solution $(\bar{S}(t), \bar{I}(t), \bar{I}_a(t))$ with $(\bar{S}(0), \bar{I}(0), \bar{I}_a(0)) \in X_0$, such that $(\bar{S}(t), \bar{I}(t), \bar{I}_a(t)) \rightarrow \mathcal{E}_0$ as $t \rightarrow +\infty$. Since $R_d > 1$, we choose $\eta > 0$ small enough such that
\[
(1 - \eta) R_d > 1.
\]

Thus, when $t$ is sufficiently large, we have $S_0 - \eta \leq \bar{S}(t) \leq S_0 + \eta$, $0 \leq \bar{I}(t) \leq \eta$, $0 \leq \bar{I}_a(t) \leq \eta$, and
\[
\dot{I} = \beta (I + k_1 S) (S_0 - \eta) - k_1 I,
\]
\[
\dot{I}_a = \alpha I - k_2 I_a.
\]

By the comparison principle, it follows that $\bar{I}(t) \rightarrow +\infty$, $\bar{I}_a(t) \rightarrow +\infty$ as $t \rightarrow +\infty$, which contradicts $\bar{I}(t) \rightarrow 0$, $\bar{I}_a(t) \rightarrow 0$ as $t \rightarrow +\infty$. This proves $W^s \cap X_0 = \emptyset$.

Since $W^s(\mathcal{E}_0) \cap X_0 = \emptyset$, $\mathcal{E}_0$ is an isolated invariant set in $X$, and $\mathcal{E}_0$ is acyclic in $\mathcal{G}_0$, and by Theorem 4.6 in [29], hence we are able to conclude that system (C.5) is uniformly persistent with respect to $(X_0, \partial X_0)$. Then, system (2) is uniformly persistent.

We proceed to investigate the global stability of $\mathcal{E}_0$.

Theorem C.2 (see [30]). Let $x \mapsto f(x) \in \mathbb{R}^n$ be a $C^1$ function for $x$ in an open set $D \subset \mathbb{R}^n$. Consider the differential equation
\[
x' = f(x).
\]

Denote by $x(t, x_0)$ the solution to (C.12) such that $x(0, x_0) = x_0$. A set $K$ is said to be absorbing in $D$ for (C.12) if $x(t, K) \subset D$ for each $K \subset D$ and $x \subset \text{Int} \mathcal{G}_n$. The following conditions are satisfied, then the unique equilibrium $\bar{x}$ is globally asymptotically stable:

1. There exists a compact absorbing set $K \subset D$ and (C.12) has a unique equilibrium in $\bar{x}$ in $D$.
2. $\bar{x}$ is locally asymptotically stable.
4. Each periodic orbit of (C.12) in $D$ is asymptotically orbitally stable.

Theorem C.3. The trajectory of any nonconstant periodic solution of system (2), if it exists, is asymptotically orbitally stable with asymptotic phase.

On investigating the global stability of $\mathcal{E}_0$, the equations for the mentally ill individuals and the recovered population can be omitted. Thus, system (2) reduces to
\[
\dot{S} = \mu - \beta (I + k_1 S) S - \mu S,
\]
\[
\dot{I} = \beta (I + k_1 S) S - k_1 I,
\]
\[
\dot{I}_a = \alpha I - k_2 I_a.
\]
We study system (C.13) in the closed set:
\[ \Phi = \{(S, I, I_a) \in \mathbb{R}^3_+ : 0 \leq S + I + I_a \leq N\}, \]  
where \( \Phi \) is positively invariant with respect to (C.13). The corresponding associated second compound matrix [31, 32] is given by
\[
A^{[2]} = \begin{pmatrix}
-\mu - k_1 - \beta (I + \kappa I_a) + \beta S & \alpha & 0 \\
-\mu - k_2 - \beta (I + \kappa I_a) - \beta S & -\beta & \beta (I + \kappa I_a) \\
\beta (I + \kappa I_a) & -k_1 - k_2 + \beta S & 0
\end{pmatrix}.
\]

Then, the second compound system defined along the periodic solution \((S(t), I(t), I_a(t))\) of system (C.13) is given by
\[
\begin{align*}
\dot{X}(t) &= -[\mu + k_1 + \beta (I + \kappa I_a) - S]X(t) \\
&\quad + \beta k S [Y(t) + Z(t)], \\
Y(t) &= \alpha X(t) - [\mu + k_2 + \beta (I + \kappa I_a)] Y(t) \\
&\quad - \beta SZ(t), \\
Z(t) &= \beta (I + \kappa I_a) Y(t) + [\beta S - (k_1 + k_2)] Z(t).
\end{align*}
\]  

Based on Theorem C.2, if we can prove that system (C.13) is asymptotically stable, then the periodic solution is asymptotically orbitally stable with asymptotic phase. We define the Lyapunov function
\[
V(X, Y, Z, S, I, I_a) = \sup \{|X|, |Y|, |Z|\}.
\]

From condition (2), we have that there exists a constant \(c_0 > 0\) such that
\[
V(X, Y, Z, S, I, I_a) \geq c_0 \sup \{|X|, |Y|, |Z|\},
\]  
for any \((X, Y, Z) \in \mathbb{R}^3\) and any periodic solution \((S, I, I_a) \in \Phi\). Let us estimate the right derivative of \(V\) along a solution \((X(t), Y(t), Z(t))\) of system (C.16):
\[
\begin{align*}
D_+ |X(t)| &\leq -[\mu + k_1 + \beta (I + \kappa I_a - S)]|X(t)| \\
&\quad + \beta k S [|Y(t)| + |Z(t)|], \\
D_+ |Y(t)| &\leq \alpha |X(t)| - [\mu + k_2 + \beta (I + \kappa I_a)]|Y(t)| \\
&\quad - \beta S |Z(t)|, \\
D_+ |Z(t)| &\leq \beta (I + \kappa I_a) Y(t) + [\beta S - (k_1 + k_2)] Z(t).
\end{align*}
\]  

We then need to estimate the following two functions:
\[
\begin{align*}
&g_1(t) = -2\mu - (\alpha + \gamma + \psi + \delta) - \beta (I + \kappa I_a) + \beta S \\
&\quad + \beta k I_a S I, \\
&g_2(t) = \alpha I - 2\mu - (\rho + \phi + d) + I' - I_a' - \frac{I'}{I} - \frac{I_a'}{I_a}.
\end{align*}
\]  

From system (C.13), we have
\[
\begin{align*}
\frac{I}{I} &= \beta S + \frac{\beta k I_a S I}{I} - (\mu + \alpha + \gamma + \sigma + \psi), \\
\frac{I_a}{I_a} &= \alpha I_a - (\mu + \rho + \phi + d).
\end{align*}
\]  

Thus,
\[
\begin{align*}
&g_1 = \frac{I}{I} - (\mu + \psi) - \beta (I + \kappa I_a) \leq \frac{I}{I} - \mu, \\
g_2 \leq \frac{I}{I} - \mu.
\end{align*}
\]  

Therefore,
\[
m(Q) \leq \frac{I}{I} - \mu.
\]
Denote the period of the periodic solutions \((S(t), I(t), I_a(t))\) by \(\tau\). We have
\[
\int_0^\tau \left[ \frac{i}{I} - \mu \right] dt \leq \ln I(t)|_0^\tau - \mu \tau = \ln I(t) - \mu \tau < 0. \quad \text{(C.24)}
\]
Thus, system (C.13) is asymptotically stable. Then, the periodic solutions \((S(t), I(t), I_a(t))\) are asymptotically orbitally stable with asymptotic phase.

**Proof of Theorem 3.** System (2) is converted into an equivalent problem, namely, the problem of minimizing the Hamiltonian \(H\) given by
\[
H = I + I_a + Au^2 + Bv^2 + \lambda_1 [\mu - (1 - u) \lambda S - \mu S] + \lambda_2 [(1 - u) \lambda S - (\alpha + \gamma v + \sigma + \mu + \psi) I] + \lambda_3 [\alpha I - (\phi + \rho v + \mu + d) I_a] + \lambda_4 [\sigma I + \phi I_a - (\mu + \epsilon v + \delta) M] + \lambda_5 [v(\gamma + \rho I_a + \epsilon M) - (\mu + \omega) R],
\]
where \(\lambda_i\), for \(i = 1, 2, \ldots, 5\), are the adjoint functions associated with their respective states. Note that, in (C.25), each adjoint function multiplies the right-hand side of the differential equation of its corresponding state function. The first terms in \(H\) come from the integrand of the objective functional. Given an optimal control pair \((u^*, v^*)\) and corresponding states \((S^*, I^*, I_a^*, M^*, \text{and } R^*)\), there exist adjoint functions satisfying
\[
\begin{align*}
\lambda_1(t) &= \mu \lambda_1 + (1 - u) \beta (I + \kappa I_a) (\lambda_1 - \lambda_2), \\
\lambda_2(t) &= -1 + (\psi + \mu) \lambda_2 + \sigma (\lambda_2 - \lambda_3) + \alpha (\lambda_2 - \lambda_4) + \nu \rho (\lambda_2 - \lambda_5) + (1 - u) \beta S (\lambda_1 - \lambda_2), \\
\lambda_3(t) &= -1 + (\mu + d) \lambda_3 + \phi (\lambda_3 - \lambda_4) + \nu \phi (\lambda_3 - \lambda_5) + (1 - u) \beta c (\lambda_1 - \lambda_2) S, \\
\lambda_4(t) &= \mu \lambda_4 + \psi (\lambda_4 - \lambda_5), \\
\lambda_5(t) &= (\mu + \omega) \lambda_5,
\end{align*}
\]
with transversality conditions, \(\lambda_i(T) = 0\), for \(i = 1, 2, \ldots, 5\). Observe that the right-hand side of \(\lambda_1\) differential equation is \(-\partial H/\partial S\), and similarly for the other adjoint functions. The final time boundary conditions (transversality conditions) are zero since there is no dependence on the states at the final time in the objective functional. Furthermore, the optimal controls are characterized by
\[
\begin{align*}
\mathcal{U}^*(t) &= \min \left\{ \max \left\{ 0, \frac{\beta S (I + \kappa I_a) (\lambda_2 - \lambda_1)}{2\lambda} \right\}, 1 \right\}, \\
\mathcal{V}^*(t) &= \min \left\{ \max \left\{ 0, \frac{\nu \phi (\lambda_2 - \lambda_3) I + \rho (\lambda_3 - \lambda_5) I_a + \epsilon (\lambda_4 - \lambda_5) M}{2B} \right\}, 1 \right\}.
\end{align*}
\]

The control characterization for \(u^*\) is obtained on \(\partial H/\partial u = 0\) whenever \(0 < u^*(t) < 1\) and taking bounds into account, similar to the other controls.

**Conflict of Interests**

The authors declare that there is no conflict of interests regarding the publication of this paper.

**References**


