PLoS ONE

Resource Allocation for Epidemic Control in Metapopulations

Martial L. Ndeffo Mbah[1,2]*, **Christopher A. Gilligan**[2]

1 Yale School of Public Health, Yale University, New Haven, Connecticut, United States of America, 2 Department of Plant Sciences, University of Cambridge, Cambridge, United Kingdom

Abstract

Deployment of limited resources is an issue of major importance for decision-making in crisis events. This is especially true for large-scale outbreaks of infectious diseases. Little is known when it comes to identifying the most efficient way of deploying scarce resources for control when disease outbreaks occur in different but interconnected regions. The policy maker is frequently faced with the challenge of optimizing efficiency (e.g. minimizing the burden of infection) while accounting for social equity (e.g. equal opportunity for infected individuals to access treatment). For a large range of diseases described by a simple SIRS model, we consider strategies that should be used to minimize the discounted number of infected individuals during the course of the epidemic. We show that when faced with the dilemma of choosing between socially equitable and purely efficient strategies, the choice of the control strategy should be informed by key measurable epidemiological factors such as the basic reproductive number and the efficiency of the treatment measure. Our model provides new insights for policy makers in the optimal deployment of limited resources for control in the event of epidemic outbreaks at the landscape scale.

Introduction

The management of diseases involves the expenditure of limited resources, which more often than not are outstripped by the demand for controlling all infected individuals [1–3]. This is often the case when disease occurs simultaneously in different but interconnected regions [2,4,5]. Treatment of infection in one region such as a state, city, or hospital may affect the potential for spread to another region when there is movement of individuals between the regions. Seeking to control disease outbreaks in more than one region, poses a dilemma for epidemiologists and health administrators of how best to deploy limited resources, such as drugs or trained personnel, amongst the different regions [6–11]. One common objective is to minimise the numbers of infected individuals and hence to minimize the burden of infection during the course of an epidemic [4,12]. For epidemics of the SIS (Susceptible-Infected-Susceptible) form, in which individuals can be re-infected, Rowthorn et al. [10] showed that rather than targeting the region with most infecteds, as might have been intuitively expected, it is instead optimal to give preference to treating the region with the *lower* levels of infecteds: the remaining regions are treated as residual claimants, receiving treatment only when there is resource left over. The epidemiological intuition underpinning the optimal strategy is understood by noting that since there are only two types of host (susceptible or infected), preferential treatment in a region with low level of infection is equivalent to giving preference to the region with the highest level of susceptibles available for infection. Since, on average an

infected individual infects more than one susceptible, removing infecteds where susceptibles are plentiful reduces the force of infection of the epidemic and so is likely to bring the epidemic under control. But what happens when there are more than two epidemiological classes? For many diseases, reinfection is often preceded by a period of temporary immunity, yielding a third class of 'removed' individuals in the population that complicates the identification of an optimal strategy for control. In this paper, we focus on this much broader class of epidemics described by an SIRS model.

We consider an SIRS-type epidemic in which infected individuals cease to be infectious and move into a temporary immune (R) class, after which they become susceptible once again. This is characteristic of many diseases, such as malaria [13,14], tuberculosis [15] and syphilis [16], in which infecteds (I) recover naturally or after treatment. Infected individuals gain a temporal immunity to the pathogen, after which they rejoin the susceptible class (S) and can be reinfected. We assume that treatment is not used as a prophylactic so that only infected individuals receive treatment. Hence, the proportion of treated individuals is given as $fI \leq I$ $(0 \leq f \leq 1)$.

To address the problem of resource allocation for disease management in multiple regions, we use a combination of optimization methods from economic theory of disease control [17,18] with a metapopulation model from epidemiological theory [19,20]. This enables us to formalize the problem and to derive criteria for optimality so as to minimize the total number of infections over time. Not infrequently, strict criteria for optimiza-

tion identify strategies that may be logistically impractical, for example by requiring a change in pattern of control at a switching time that may be difficult to monitor [17]. Strictly optimal strategies may also be challenged on grounds of social equity, whereby every infected individual does not have an equal chance of being treated [21,22]. Accordingly we assess the tractability of optimal control strategies and consider also how adaptations may be made to balance, optimality, tractability and social equity. For the sake of simplicity, the analysis is initially carried out for two interconnected regions (e.g.cities, towns or states) and the robustness of the results to spatial structure are later tested for two other simple and realistic spatial configurations.

Model

We consider two coupled sub-populations (regions) of susceptible individuals each with a fixed size N, in which an epidemic is described by a simple SIRS compartmental model:

$$\frac{dS_i}{dt} = \sigma N + v R_i - \beta(1-\epsilon)S_i I_i/N - \beta\epsilon S_i I_j/N - \sigma S_i \quad (1)$$

$$\frac{dI_i}{dt} = \beta(1-\epsilon)S_i I_i/N + \beta\epsilon S_i I_j/N - \mu I_i - \eta f_i I_i - \sigma I_i \quad (2)$$

$$\frac{dR_i}{dt} = \mu I_i + \eta f_i I_i - v R_i - \sigma R_i \quad (3)$$

with $i \neq j$ and $i,j = 1,2$. Each sub-population is composed of susceptible (S), infectious (I) and recovered (R) individuals, and are scaled here as proportions. The transmission rate for each sub-population is given by β. The coupling strength between sub-populations is given by $0 < \epsilon < 1$. The infectious period is given by μ^{-1}; v is the rate of loss of immunity, and σ is rate of birth/death. η is a measure of the incremental increase in the recovery rate of treated individuals, and f_i is the proportion of infected individuals in sub-population i that receive treatment. When all infected individuals receive treatment ($f_1 = f_2 = 1$), the basic reproductive number, which is a widely-used epidemiological measure of the intrinsic potential for multiplication of an epidemic, is given by $R_0 = \frac{\beta}{\mu+\sigma+\eta}$. Without treatment R_0 is equal to $\frac{\beta}{\mu+\sigma}$.

Optimal control

We suppose that expenditure on control is subject to a budget constraint $c(f_1 I_1 + f_2 I_2) \leq M$, where c is the cost of treatment per individual. This simple fixed budget constraint is used as a surrogate that encompasses limitations in the amount of drug available and for mobilisation and delivery of resources at the point of infection (limitations in transport or trained personnel). These limit the instantaneous availability of drug. If there are sufficient resources, all infected individuals will be treated. Otherwise, resources are allocated so as to minimize the discounted number of infected individuals in both sub-populations over time. Hence, we choose f_1 and f_2 so as to minimize the following integral

$$J = \int_0^\infty e^{-rt}(I_1 + I_2)dt \quad (4)$$

The discount rate (r) is included to allow for long-term changes, thus giving greater emphasis to control in the short rather than the long term [17]. The optimization approach we adopt is based upon the Hamiltonian method [23], which is a device for minimizing the objective function subject to the economic constraints and the epidemiological dynamics of the model.

We assume that if it were possible to treat all infected individuals, disease eradication would be achieved in the long term ($\frac{\beta}{\mu+\sigma+\eta} < 1$). Using Filippov's theorem [24], it is possible to show that the optimal control problem does have a solution. To solve the problem of optimal deployment of limited resources (i.e., when there are insufficient resources to treat all individuals that may become infected), we use the Pontryagin maximum principle [23](PMP), a mathematical tool widely used to solve optimal control problems for dynamical systems. This method takes into account the influence of current infection on the future evolution of disease as given by the propagation equations (1)–(3). The influence is embodied in the co-state variables that appear in a mathematical expression known as the Hamiltonian (see Materials and Methods). PMP enables us to derive necessary conditions for optimality from which it is possible to build up a set of candidate strategies for optimality from which ultimately it is possible using extensive numerical simulation to identify an optimal solution.

Results

Efficiency maximization

The Pontryagin maximum principle (PMP) was used to derive necessary conditions for optimal resource allocation, when there are insufficient resources to treat all infected individuals. Using these necessary conditions together with exploratory numerical analysis, we identify the following as candidate strategies for optimality (see Materials and Methods):

- preferential treatment of the more infected sub-population - to equalize disease burden within the regions as fast as possible and thereafter to treat each region equally;

- preferential treatment of the less infected sub-population - initially 'sacrificing' the sub-population with the higher level of infecteds

- preferential treatment of the more susceptible sub-population - initially 'sacrificing' the sub-population with the lower level of susceptibles

- a strategy involving at least one switch between preferential treatment of the more infected to either the less infected or the more susceptible sub-population.

Although it is not possible to prove analytically that a given path is optimal, after extensive numerical simulation, we identify the single switch strategy from giving preference to the more infected sub-population to giving preference to the less infected sub-population as the best allocation strategy that minimizes the discounted total numbers of infected individuals in both sub-populations (Figs. 1 & 2). However, attempts to implement the switching strategy are prone to the risk of missing the optimal switching time. This risk is enhanced by the fact that the optimal switching time depends upon the values of epidemiological parameters and the initial levels of infection that are unlikely to be accurately known in advance.

To conclude our analysis on efficiency maximization, we investigate the effect of the rate of loss of immunity v on the best allocation policy, by considering respectively the cases $v \to \infty$ and $v \to 0$ (see Table 1). Using numerical simulation, we compare the candidate strategies for optimality (see Materials and Methods) and show that for very large values of v, the best allocation strategy is always to give 'preference to the more susceptible sub-population'. This observation agrees with the results of Rowthorn
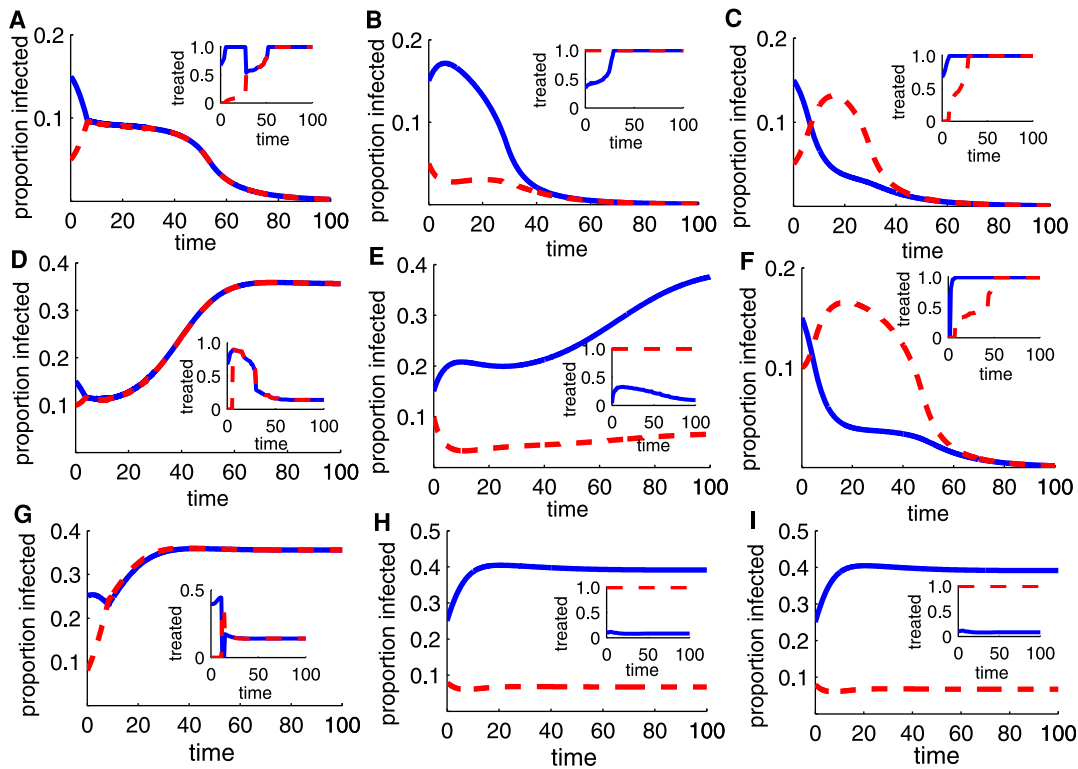
Figure 1. Comparison of disease progress curves for a strategy that gives preferential treatment to the more infected sub-population (A,D,G), preferential treatment to the less infected sub-population (B,E,H) and the most efficient strategy (C,F,I). Disease progress is shown for different control outcomes. (A–C) Progress of disease in two interconnected regions 1 (solid lines) and 2 (dashed lines), with treatment dynamics in insets, showing little differences between control strategies. (D–F) Preferential treatment to the more infected sub-population and preferential treatment to the less infected sub-population diverge markedly from the most efficient strategy. (G–I) Disease continues to increase but markedly less steeply in the region with the lower infestation (region 2), for preferential treatment to the less infected sub-population and most efficient strategy. Disease progress curves are given for three different values of the initial number of infected: (A,B,C), (D,E,F), and (G,H,I). Default parameters are $\eta = 4$ (efficiency of control), $\beta = 6$ (within-region transmission rate), $\epsilon = 0.1$ (coupling strength), $\mu = 2 y^{-1}$ (recovery rate), $v = 0.2 y^{-1}$ (rate of loss of immunity), $\sigma = 0.03 y^{-1}$ (rate of birth/death), $r = 0.05 y^{-1}$ (discount rate), and $M = 0.1$ (fixed expenditure limit).
doi:10.1371/journal.pone.0024577.g001

et al. [10] who show this policy to be the best strategy for the control of an SIS type epidemic. Whereas for very small values of v, a double switch of preference between the more and the less infected region was shown to outperform the other allocation strategies. It is difficult to prove the existence of an upper bound to the number of switches. However, the more switches there are, the harder the implementation of the allocation strategy would be.

The single switch strategy, though the best policy, is not easily implementable. Numerical simulation shows that the second best policy in terms of simplicity and efficiency maximization is either to give preference to the more susceptible sub-population or preference to the less infected sub-population depending on the initial state of the system (Fig. 3). We compare the performance of these policies for different values of the rate of loss of immunity (v) (Fig. 4). For $v \rightarrow \infty$ the two inequitable policies are identical. As the value of v decreases, the difference between the policies increases, with the preferential treatment of the more susceptible sub-population outperforming the preferential treatment of the less infected sub-population. However, when the rate of loss of immunity becomes small ($v/\mu < 0.4$) the relative performance of the two policies becomes highly dependent on the initial state of the system (Fig. 4).

Efficiency and social equity

Since the optimal strategy is very difficult to implement, two robust alternative strategies would be either to give preference to the more susceptible sub-population or to give preference to the less infected sub-population. However, these strategies are likely to be regarded as highly socially inequitable from the perspective of the chance that any infected individual receives treatment. For the initial state of the system satisfying: $I_1, I_2 < 25\%$ and $R_1, R_2 < 25\%$ (which may be regarded as early implementation of control), we consider a widely-advocated, socially equitable strategy comprising

- a pro-rata policy designed to give equal opportunity for any infected individual to receive treatment [22,25].

We compare the performance of this strategy with the three tractable strategies considered above (i.e. not involving switching). We do this for different values of the basic reproductive number R_0 (Fig. 5): R_0 is a widely-used epidemiological measure of the intrinsic potential for multiplication of an epidemic.

Given a threshold value, of the difference between the outcome of a given control strategy and that of the pro-rata strategy, (d%) above which the use of inequitable policies may be justifiable, Fig. 5 shows that there exists a threshold value R^* such that for $R_0 < R^*$, the pro-rata policy performs almost as well as the other policies (e.g. for d = 10% $R^* = 3.5$). In this case, the pro-rata policy is a good compromise in terms of equity, efficiency and simplicity. For $R_0 > R^*$, it would be better to opt for an inequitable policy (e.g. preferential treatment to the more susceptible sub-population). For high values of R_0, the decreasing difference of value
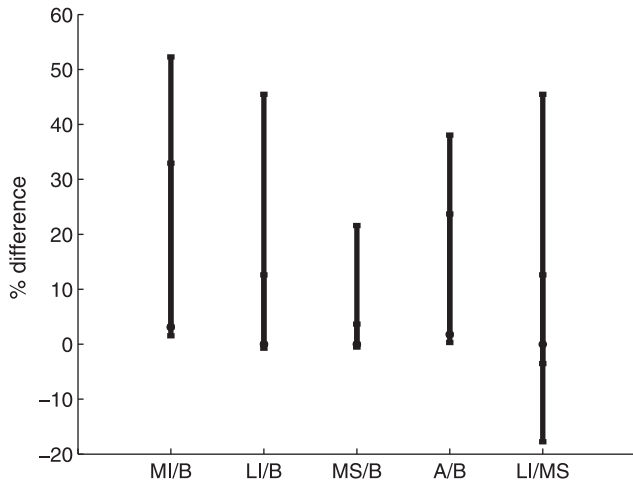
Figure 2. Difference between the outcome of the different policies for the whole range of initial conditions. For a given initial condition, the difference between two strategies (e.g 'preferential treatment to the more infected sub-population' and 'preferential treatment to the less infected sub-population') is computed by $(J_{\text{more infected}} - J_{\text{less infected}})/J_{\text{less infected}}$ in which $J_{\text{less infected}}$ and $J_{\text{more infected}}$ are the values of the discounted burden of infection for the 'preferential treatment to the more infected sub-population' and 'preferential treatment to the less infected sub-population' strategy, respectively. MI, MS, LI, A, and B denote respectively the 'preferential treatment to the more infected sub-population', 'preferential treatment to the more susceptible sub-population', 'preferential treatment to the less infected sub-population', alternative and single switch policy from 'preferential treatment to the more infected sub-population' to 'preferential treatment to the less infected sub-population'. The average difference value is represented by the dot, whereas the top and bottom bars represent respectively the maximum and minimum values. As for the middle bars, they represent respectively the ninety-ninth and first percentiles. The average and percentiles were obtained for 705078 initial conditions (See Materials and Methods).
doi:10.1371/journal.pone.0024577.g002

between the policies is due to the inefficiency of the control measure (drug efficiency) in bringing the epidemic under control (cf. Fig. 5 and 6). We also compare the policies for different value of the coupling strength between the two sub-populations ($0 < \epsilon < 1/2$). We observed that when the coupling strength decreases, the difference between the outcome of the control policies increases with the pro-rata strategy becoming the more less efficient than the optimal strategy (the single switch strategy from giving preference to the more infected sub-population to giving preference to the less infected sub-population). On the other hand, the difference between the outcome of the control policies declines as the coupling strength gets larger. Thus, as the transmission between the sub-population increases the outcome

Table 1. Effect of the rate of loss of immunity (ν) on the best allocation strategy.

Epidemic model	SIRS	SIRS ≡ SIS; $\nu \to \infty$	SIRS ≡ SIR; $\nu \to 0$
Best strategy	'single switch[1]'	'no switch[2]'	'double switch[3]'

[1]single switch of preference from the more infected sub-population to the less infected sub-population.
[2]preference to the less infected sub-population.
[3]double switch of preference between the less infected and the more infected sub-population.
doi:10.1371/journal.pone.0024577.t001

Figure 3. Difference between the outcome of the different policies for the whole range of initial conditions. Notations follow from Fig. 2. The figure shows the difference in efficiency between 'preference to the more infected sub-population' and 'preference to the less infected sub-population' policies. The average value is represented by the dot, whereas the top and bottom black bars represent respectively the maximum and minimum values. The middle blue bars represent respectively the ninety-ninth and first percentiles.
doi:10.1371/journal.pone.0024577.g003

becomes less sensitive to the choice of policy [10] (result not shown here).

To investigate the robustness of the result to spatial structure, we consider two further spatial configurations: 10 identical regions with symmetrical global coupling, and 10 identical regions
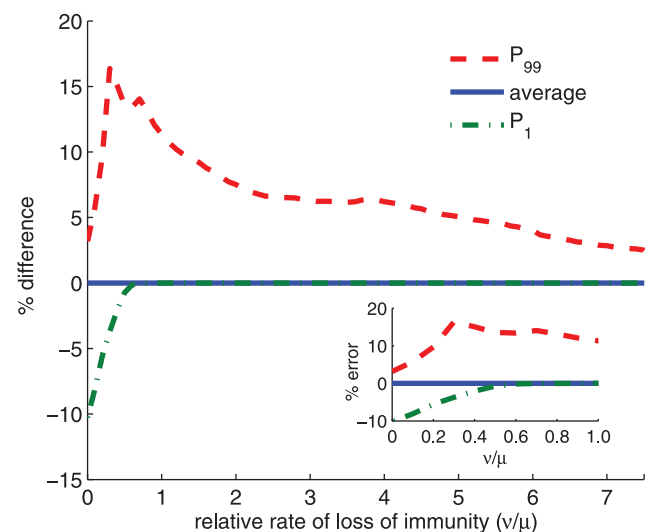
Figure 4. Difference between the outcome of 'preferential treatment to the less infected sub-population' and that of 'preferential treatment to the more susceptible sub-population'. The figure shows the effect of the change of the rate of loss of immunity on the difference between the outcome of the two strategies. The inset shows the effect for values of the rate of loss of immunity (ν) less or equal to the natural recovering rate (μ). Difference is computed as with Fig. 2. The solid line represents the average value, whereas the dashed and dash dotted lines represent respectively the ninety-ninth and the first percentiles. The average and percentiles were obtained for 705078 initial conditions (See Materials and Methods).
doi:10.1371/journal.pone.0024577.g004

Figure 5. Difference between the outcome of selected strategies for different values of R_0. The blue lines represent the difference between the outcome of 'preferential treatment to the more infected sub-population' and that of the pro-rata policy. The green lines represent the difference between the outcome of the pro-rata policy and that of 'preferential treatment to the more susceptible region', and the red lines represent the difference between the outcome of the pro-rata policy and that of 'preferential treatment to the less infected region'. Averages are obtained for 10,000 initial states of the system, generated randomly.
doi:10.1371/journal.pone.0024577.g005

arranged in a circle with each population interacting only with its two nearest neighbours. For small values of R_0, the relative outcome of the control policies is independent of the spatial structure of the system (Fig. 7). But for high values of R_0, the variability in the outcome of the control policies with respect to the
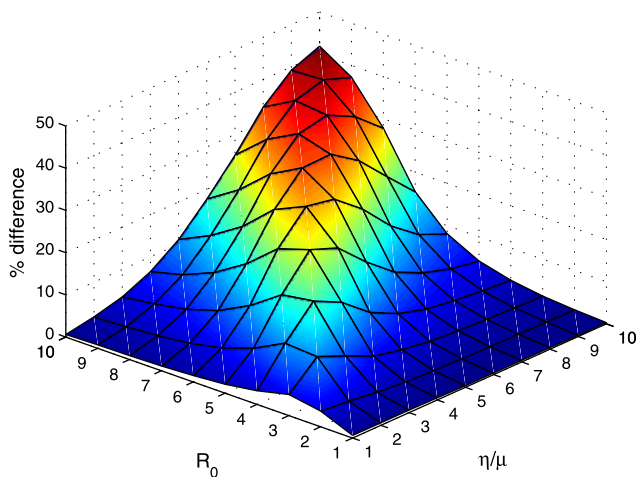
Figure 6. Difference between the outcome of 'preferential treatment to the less infected sub-population' and that of that pro-rata policy for symmetrical global connection between regions. The effect of the efficiency of the treatment measure on the average difference of outcome between 'preferential treatment to the less infected sub-population' and pro-rata policy with respect to the basic reproductive ratio (R_0) is shown. Averages are obtained for 10,000 initial states, generated randomly, of a system of 10 regions. Default parameter values as given in Fig. 1, except $M = 0.5$.
doi:10.1371/journal.pone.0024577.g006

initial state of the system increases with the sparsity of the coupling matrix (Fig. 7). Simulation shows that the threshold value of R_0 increases with the efficiency of the treatment measure (Fig. 6), and decreases for increasing values of the rate of loss of immunity (Fig. 8). Given that the choice of the discount rate affects the relative valuation of the current and future disease, one would expect a correlation between the choice of the discount rate and the value of the percentage error above which social inequity is justifiable.

When pro-rata is not a good candidate strategy in terms of efficiency ($R_0 > R^*$), an alternative strategy for balancing efficiency and social equity may be the use of proportional allocation where a fraction of the resources is allocated pro-rata, while the remaining is allocated so as to maximize efficiency [22]. However, determining what fraction of resource is to be allocated for equity concerns, while retaining a good level of overall efficiency, requires further debate and greater interrogation of epidemiological models with insight from social sciences [22,26].

Discussion

We have addressed the problem of allocation of limited resources for the control of an SIRS-type epidemic in different but interconnected regions. Using a combination of optimization methods from economic theory with a metapopulation model from epidemiological theory for disease management, we have formalized the problem of resource allocation and derived criteria for optimality so as to minimize the discounted number of infected individuals in both sub-populations over time, during the course of the epidemic. Using extensive numerical simulations, we have shown that the best strategy in terms of efficiency maximization is a switching strategy, whereby resources are initially preferentially allocated to the more infected sub-population then to the less infected sub-population. However, this strategy is seldom tractable, due to the fact that the switching time depends upon the value of epidemiological parameters and the initial state of the system, which are unlikely to be accurately known [17,27].

Given that a practical strategy for disease control must account for various factors such as efficiency maximization and social equity amongst others, we have extended previous studies on dynamic resource allocation by investigating how to account for optimality (minimizing the burden of infection), social equity (equal opportunity for infected individuals to access treatment), and simplicity (ease of implementation) in identifying strategies for disease control. We have shown that when faced with the dilemma of choosing between a socially equitable strategy for resource allocation (e.g. a pro-rata allocation strategy) and a purely efficient but inequitable strategy (e.g. by giving preference to the more susceptible sub-population or preference to the less infected sub-population), the decision should be informed by the value of key epidemiological and economic parameters. In particular, we have shown that given a certain percentage of difference between the outcomes of different strategies (i.e. relative discounted number of infections that are not averted under the pro-rata policy) above which the use of an inequitable policy may be justifiable, there exists a threshold value of the basic reproductive number (R_0) below which it is better to adopt a purely socially equitable strategy (pro-rata policy). This threshold value was shown to increases with the efficiency of the treatment measure, and to decrease with the average duration of the period of temporal immunity. The social context of our analysis implies that equal weighting is given to the health of each individual i.e. for the collective good of the entire population.
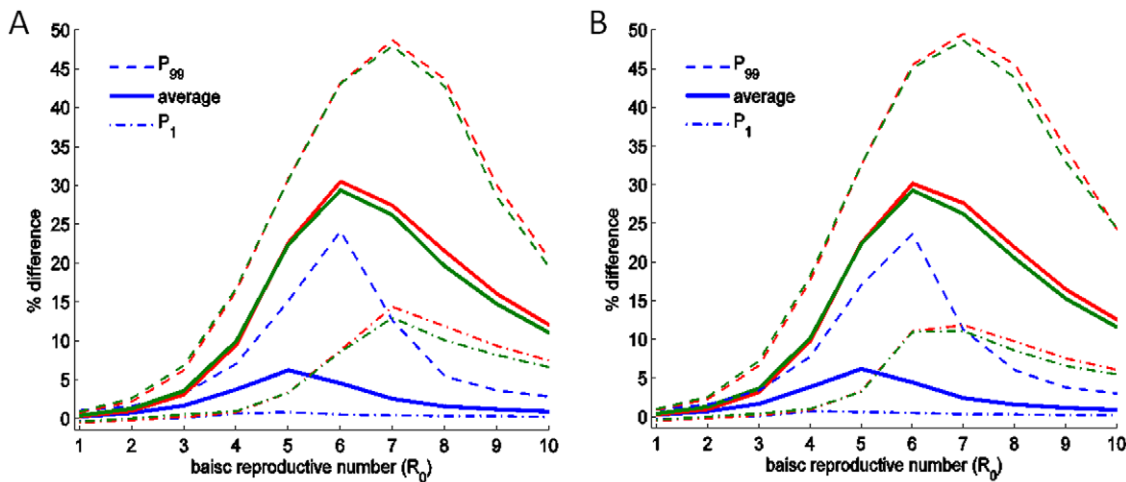
Figure 7. Difference between the outcome of selected policies for different values of R_0 for multiple sub-populations with different coupling between sub-populations. For (A), regions are inter-connected with symmetrical global coupling. For (B), regions are inter-connected only with nearest neighbours. The blue lines represent the difference between the pro-rata policy and the 'preferential treatment to the more infected sub-population' policy. The green lines represent the difference between the 'preferential treatment to the more susceptible sub-population' policy and the pro-rata policy, and the red lines represent the difference between the 'preferential treatment to the less infected sub-population' policy and the pro-rata policy. Averages are obtained for 10,000 initial states, generated randomly, of two systems of 10 regions. Default parameter values as given in Fig. 1, except $M = 0.5$.
doi:10.1371/journal.pone.0024577.g007

Interest in the optimal allocation of resources for epidemic control in structured populations has recently been renewed due to the threat of pandemic influenza [11,28–30]. These studies primarily focus on the optimal deployment of mass vaccination to prevent or mitigate the spread of an outbreak of influenza within a population. Among other things, they show that when vaccine supplies are limited and the public health objective is to minimize infections, it is optimal to target vaccination toward the more epidemiologically important sub-populations (those that
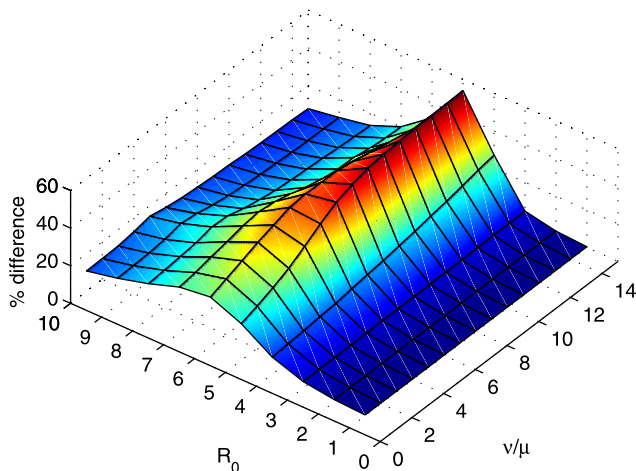
suffer the greatest per capita burden of infection) [11,28–30]. The other sub-populations would thus be indirectly protected through herd immunity [11,28]. These results agree with our analysis which shows that a good control strategy in terms of simplicity and efficiency maximization would be to give preference to the more susceptible sub-population. This sub-population may be regarded as the more epidemiologically important as it is potentially the main contributor to future infections.

Several areas of investigation suggest themselves for future work. Foremost amongst these are allowance for heterogeneity in the size of sub-populations, and the rates of transmission of infection, both of which are recognized to be important factors in metapopulation theory. Further work will also investigate the robustness of the results for different measures for efficiency of control and to uncertainty about the likely values of epidemiological parameters, given that optimal strategies are often very sensitive to the epidemiological parameters [27,31–33], which may not be accurately known before control is implemented.

Materials and Methods

The objective is to minimize the discounted burden of infection during the course of the epidemic

$$J = \int_0^\infty e^{-rt}(I_1 + I_2)dt \qquad (5)$$

subject to the propagation equations (1)–(3) and the following epidemiological and economic constraints:

$$(S_i, I_i, R_i)(0) = (S_{i0}, I_{i0}, R_{i0});$$
$$0 \le f_i \le 1; \ f_1 I_1 + f_2 I_2 = \min(I_1 + I_2, M/c)$$

Each sub-population, of a fixed size N, is composed of susceptible (S), infectious (I) and recovered (R) individuals, and are scaled here as proportions.

Figure 8. Difference between the outcome of 'preferential treatment to the less infected sub-population' and that of the pro-rata policy for symmetrical global connection between regions. The effect of the change of the rate of loss of immunity on the average difference between 'preferential treatment to the less infected sub-population' and pro-rata policy with respect to the basic reproductive ratio (R_0) is shown. Averages are obtained for 10,000 initial states, generated randomly, of a system of 10 regions. Default parameter values as given in Fig. 1, except $M = 0.5$.
doi:10.1371/journal.pone.0024577.g008

Let $A = \{I_1, I_2 : I_1 + I_2 \leq M/c\}$ be the region where there are sufficient resources to treat all infected individuals. Since $S_i + I_i + R_i = 1$, the equations on R_i can be ignored.

When there are more infecteds that can be treated, $c(I_1 + I_2) > M$ and hence $f_1 I_1 + f_2 I_2 = M/c$. The relevant Hamiltonian in this case is

$$H = -e^{-rt}(I_1 + I_2) + m_1 \dot{S}_1 + m_2 \dot{I}_1 + m_3 \dot{S}_2 + m_4 \dot{I}_2 \qquad (6)$$

where m_i are the co-state variables. Since $f_1 I_1 = M/c - f_2 I_2$, the Hamiltonian can be written as

$$
\begin{aligned}
H = &-e^{-rt}(I_1 + I_2) + \\
&m_1(\sigma(1 - S_1) + v(1 - S_1 - I_1) - \beta(1 - \epsilon)S_1 I_1 - \beta\epsilon S_1 I_2) + \\
&m_2(\beta(1 - \epsilon)S_1 I_1 + \beta\epsilon S_1 I_2 - \mu I_1 - \eta(M/c - f_2 I_2) - \sigma I_1) + \quad (7) \\
&m_3(\sigma(1 - S_2) + v(1 - S_2 - I_2) - \beta(1 - \epsilon)S_2 I_2 - \beta\epsilon S_2 I_1) + \\
&m_4(\beta(1 - \epsilon)S_2 I_2 + \beta\epsilon S_2 I_1 - \mu I_2 - \eta f_2 I_2 - \sigma I_2)
\end{aligned}
$$

f_2 (and hence f_1) has to be chosen so as to maximize the Hamiltonian [23]. This yields the following result:

$$
\begin{aligned}
&\text{If } m_2 - m_4 > 0 \text{ then } f_2 I_2 = \min(I_2, M/c), f_1 I_1 = M/c - f_2 I_2, \\
&\text{If } m_2 - m_4 < 0 \text{ then } f_1 I_1 = \min(I_1, M/c), f_2 I_2 = M/c - f_1 I_1
\end{aligned} \qquad (8)
$$

And it must be the case that

$$\dot{m}_i = -\frac{\partial H}{\partial x_i} \qquad (9)$$

where x_i is the corresponding state variable to m_i.

We assume that η satisfies the following condition:

$$\frac{\beta}{\mu + \sigma + \eta} \leq 1 \qquad (10)$$

That η satisfies equation 10 implies that if there are always enough resources to treat all infected individuals, disease will eventually be eradicated in the population. This is justified by the fact that this criterion (equation 10) is equivalent to the basic reproductive ratio R_0 being less than or equal to 1 [34], which here is a necessary and sufficient criterion to prevent invasion of an epidemic. Upon equation 10 any admissible path (disease dynamic curves obtained for a given value of the control functions f_1 and f_2) will either never enter region $A = \{I_1, I_2 : I_1 + I_1 \leq M/c\}$, or enter and never leave (see [10] and [27] for details). Therefore, besides the general transversality conditions $\lim_{t \to \infty} m_i(t) = 0$, there are alternative transversality conditions whenever a path enters region A [23]. We define a function V as follows

$$V(S_1(\tau), I_1(\tau), S_2(\tau), I_2(\tau)) = \int_\tau^\infty e^{-rt}(I_1 + I_2) dt \qquad (11)$$

where the integral is evaluated along the path defined by the propagation equations (1)–(3) when $f_1 = f_2 = 1$, with τ being the time at which the path enter region A. The alternative transversality conditions for a path that enters region A is given by

$$
\begin{aligned}
&m_1(\tau) = \frac{\partial V}{\partial S_1}, \; m_2(\tau) = \frac{\partial V}{\partial I_1} + q_\tau \\
&m_3(\tau) = \frac{\partial V}{\partial S_2}, \; m_4(\tau) = \frac{\partial V}{\partial I_2} + q_\tau
\end{aligned} \qquad (12)
$$

$$\text{and} \quad rV = H(\tau)$$

where q_τ is a multiplier, and $H(\tau)$ is the Hamiltonian evaluated at time τ.

Given an initial state of the system $(S_i, I_i, R_i)(0) = (S_{i0}, I_{i0}, R_{i0})$, the existence of an admissible path which enters region A depends upon the value of the expenditure limit M. When such a path exists, the optimal control problem is equivalent to an optimal timing problem, where the objective is to find the shortest path to reach region A. For such a value of M, a simple application of Filippov's theorem [24] shows that a solution to the optimal control problem exists. This is done using Theorem 10.1 from [24], and the compactness of the set of points $(\hat{S}_i, \hat{I}_i, \hat{R}_i)$ at which admissible paths, starting at (S_{i0}, I_{i0}, R_{i0}), enter region A.

The singular solution

We suppose that there exists an allowable path that satisfies the above maximal conditions on the Hamiltonian, and for which there exists an open interval where we have $m_2 = m_4$. By differentiating $m_2 - m_4$ over that open interval, we obtain

$$
\begin{aligned}
\dot{m}_2 - \dot{m}_4 &= [(m_1 - m_2)\beta(1 - 2\epsilon)]S_1 + m_1 v + \\
&\quad [(m_3 - m_4)\beta(2\epsilon - 1)]S_2 - m_3 v \qquad (13) \\
&= 0
\end{aligned}
$$

From an economical view point, the co-state variables can be interpreted as shadow prices. Thus m_1 and m_2 indicate respectively the marginal benefit to society of increasing by one unit the proportion of susceptible (S_1) and infectious (I_1) individuals of region 1 [35,36]. Because infection is harmful, and increasing the proportion of infectious individuals will result in decreasing the proportion of susceptibles, the shadow price m_2 is negative. Then, $-m_2$ represents the proportion that society is willing to invest for control that will result in reducing the stock of infectious individuals in region 1 by one unit. Moreover $m_1 - m_2$ is positive. The same results hold for m_3 and m_4. Since $m_2 = m_4$ on an open interval, it follows that $\dot{m}_2 - \dot{m}_4 = 0$ on such an interval. Equation 13 is then equivalent to

$$
\begin{aligned}
(\beta(1 - 2\epsilon)(-m_2 + m_1))S_1 + m_1 v = \\
(\beta(1 - 2\epsilon)(-m_2 + m_3))S_2 + m_3 v
\end{aligned} \qquad (14)
$$

From (14), it follows that $m_1 = m_3$ if and only if $S_1 = S_2$. If $S_1 = S_2$ on an open interval, it follows from the previous sentence that we would have $\dot{m}_1 = \dot{m}_3$ on the same open interval. Simple algebra shows that with $m_2 = m_4$ and $m_1 = m_3$ on an open interval, $\dot{m}_1 = \dot{m}_3$ implies that $I_1 = I_2$ on the same interval. Therefore, if there exists an open interval on which $m_2 = m_4$ and $m_1 = m_3$, then $S_1 = S_2$ and $I_1 = I_2$ on the same interval. The control strategy on such an interval would be given by $f_1 I_1 = f_2 I_2 = \frac{M}{2c}$.

Since $\dot{m}_2 = \dot{m}_4$ on an open interval, it follows from equation 9:

$$\dot{m}_1 = -\frac{\partial H}{\partial S_1} = -\frac{\partial H}{\partial I_1}\frac{\partial I_1}{\partial S_1} = \dot{m}_2\frac{\partial I_1}{\partial S_1} = \dot{m}_4\frac{\partial I_1}{\partial S_1}$$

$$= -\frac{\partial H}{\partial I_2}\frac{\partial I_1}{\partial S_1} = -\frac{\partial H}{\partial S_2}\frac{\partial S_2}{\partial I_2}\frac{\partial I_1}{\partial S_1} = \dot{m}_3\frac{\partial S_2}{\partial I_2}(\frac{\partial S_1}{\partial I_1})^{-1}$$

From the symmetry of the system, we have $\frac{\partial S_2}{\partial I_2}(\frac{\partial S_1}{\partial I_1})^{-1} = 1$; then $\dot{m}_1 = \dot{m}_3$ on the open interval. We conclude from the transversality conditions that $m_1 = m_3$. It follows that the singular solution is given by:

$$\text{If } S_1 = S_2, \text{and } I_1 = I_2 \text{ then } \eta_1 I_1 = \eta_2 I_2 = \frac{M}{2c} \qquad (15)$$

which satisfies that following equations:

$$\frac{dS_i}{dt} = \sigma(1 - S_i) + v(1 - S_i - I_i) - \beta(1 - \epsilon)S_i I_i - \beta\epsilon S_i I_j$$

$$\frac{dI_i}{dt} = \beta(1-\epsilon)S_i I_i + \beta\epsilon S_i I_j - \mu I_i - \frac{\eta M}{2c} - \sigma I_i, \qquad (16)$$

$$i = 1,2; \; j = 2,1$$

The singular solution is achieved by preferential treatment of infecteds in the region with the higher prevalence of infecteds (see Eq. 17). The policy is called the MRAP since it involves the Most Rapid Approach Path to the singular solution, in which infection is equalized in both sub-populations.

When $\frac{M}{2c} < \frac{(v+\sigma)}{\eta}(1 - \frac{1}{R_0})$ (where $R_0 = \frac{\beta}{\mu+\sigma}$) equation (16) has two equilibrium points given by

$$S^{*,+} = \frac{\eta M}{2c\beta I^{*,+}} + \frac{1}{R_0},$$

and

$$I^{*,+} = \frac{\left(v + \sigma - \eta\frac{M}{2c}\right)\beta - (\mu+\sigma)(v+\sigma) \pm [\Lambda]^{1/2}}{2(\mu+v+\sigma)\beta}$$

with

$$\Lambda = ((v+\sigma-\eta\frac{M}{2c})\beta - (\mu+\sigma)(v+\sigma))^2 - 2(v+\sigma)(\mu+v+\sigma)\beta\eta\frac{M}{c}.$$

We have $I^+ > I^* > 0$, where (S^*,I^*) is unstable (saddle point), and (S^+,I^+) is stable. Another stable equilibrium point (disease free equilibrium) is reached if the path enters region A.

When $\frac{M}{2c} < \frac{(v+\sigma)}{\eta}\left(1 - \frac{1}{R_0}\right)$, the singular solution may exhibit a saddle-node bifurcation along the bifurcation parameter $\frac{M}{c}$ (see Fig. 9). In other words, when the average proportion of individuals treated individuals in each sub-population $\left(\frac{M}{2c}\right)$ is lower than the epidemiological factor $\frac{(v+\sigma)}{\eta}\left(1 - \frac{1}{R_0}\right)$, the singular solution fails to eradicate the disease, as the infection path converges towards I^+, if control strategy (MRAP) is first implemented when the proportion of infected in both sub-population is above the unstable steady state (dashed line in Fig. 9).
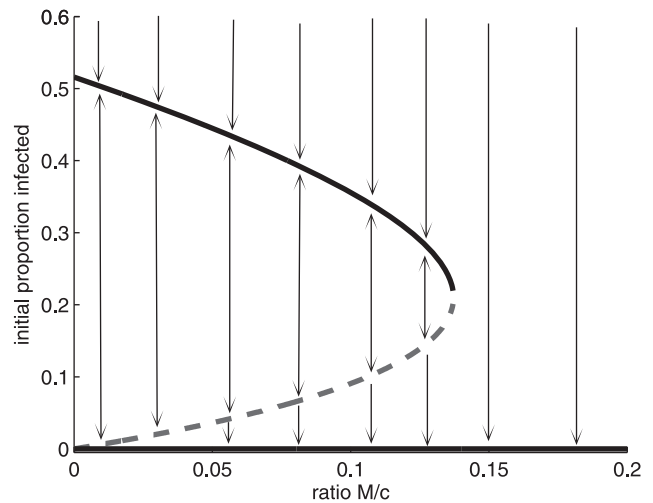
Figure 9. Bifurcation diagram for the singular solution (Eq. 16). The dashed line represents unstable steady states and the solid lines represent stable steady states. Using the most rapid approach path to the singular solution (MRAP), the initial proportion of infected represents the level of infection when the singular solution is first reached. Parameter values are given by Fig. 1.
doi:10.1371/journal.pone.0024577.g009

Candidates for optimality

From the above results, it follows that the optimal control strategy depends on the effect of a marginal change in the value of $m_2 - m_4$. However, this change can only be determine numerically. Using the shadow pricing analogy together with exploratory numerical analysis, we derive some scenarios of practical understanding that can be understood in terms of the co-state variables m_2 and m_4.

From the interpretation of the co-state variables as shadow prices, equation (8) can be interpreted as follows: if increasing the amount of infected individuals in sub-population 1 (sub-population 2) by one unit, would generate more infection in the whole population than an increase of the same amount in sub-population 2 (sub-population 1), then preference in treatment must be given to sub-population 1 (sub-population 2). From equation (15) and equation (8), it follows that an optimal solution is either a switching strategy of preference between sub-population 1 and sub-population 2, or the MRAP (the most rapid approach path to singular solution). The MRAP solution, which is equivalent to 'preferential treatment of the more infected sub-population' is given by the following equation:

$$\begin{cases} \text{If } I_i < I_j, \quad f_j I_j = \min(I_j, M/c) \text{ and } f_i I_i = M/c - f_j I_j \\ \\ \text{If } I_i = I_j \qquad\qquad\qquad\qquad \text{and} \\ \\ (a) \; S_i = S_j, \qquad\qquad f_j I_j = f_i I_i = M/2c \quad (17) \\ \\ (b) \; S_i < S_j, \quad f_i I_i = \min(I_i, M/c) \text{ and } f_j I_j = M/c - f_i I_i, \\ \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{or vice}-\text{versa} \end{cases}$$

As for the switching strategies between sub-population 1 and 2, they can be constructed in an infinite number of ways. Here we

consider two plausible candidate solutions for optimality (based upon exploratory numerical analysis): 'preferential treatment of the more susceptible sub-population', 'preferential treatment of the less infected sub-population'. These strategies are respectively defined by the following equations:

$$\begin{cases} \text{If } S_i < S_j, \quad f_j I_j = \min(I_j, M/c) \text{ and } f_i I_i = M/c - f_j I_j, \\[2mm] \quad \text{If } S_i = S_j \qquad\qquad\qquad\qquad \text{and} \\[2mm] (a)\ I_j = I_i, \quad f_j I_j = \min(I_j, M/c) \text{ and } f_i I_i = M/c - f_j I_j, \\[2mm] (b)\ I_j > I_i, \quad f_j I_j = \min(I_j, M/c) \text{ and } f_i I_i = M/c - f_j I_j, \\[2mm] \qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad orvice-versa \end{cases} \quad (18)$$

and

$$\begin{cases} \quad \text{If } I_i > I_j, \quad f_j I_j = \min(I_j, M/c) \text{ and } f_i I_i = M/c - f_j I_j, \\[2mm] \quad \text{If } I_i = I_j \qquad\qquad\qquad\qquad \text{and} \\[2mm] (a)\ S_i = S_j, \quad f_j I_j = \min(I_j, M/c) \text{ and } f_i I_i = M/c - f_j I_j, \\[2mm] (b)\ S_i < S_j, \quad f_j I_j = \min(I_j, M/c) \text{ and } f_i I_i = M/c - f_j I_j, \\[2mm] \qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad orvice-versa \end{cases} \quad (19)$$

The strategies giving preference to the more susceptible sub-population, preference to the less infected sub-population as well as the single and double switching strategies between one of the above strategies and the MRAP strategy are all candidates for optimality. Moreover, we consider an 'alternative' strategy which consists in the first instance in equalizing the level of infection in both sub-populations as fast as possible. This is done by implementing the strategy giving preference to the more infected sub-population strategy. When equality of the levels of infection is first reached, preference is then given to the more susceptible sub-population. We compare the above strategies. For any value of the initial condition, simulation shows that the smallest value of the objective function (Eq. 5) is obtained with the single switch strategy from giving preference to the more infected sub-population to giving preference to the less infected sub-population. Implementing the single switch strategy is subject to the risk of missing the optimal switching time. We were also able to show that the switching strategy satisfies the Hamiltonian and transversality conditions. We were not able to rule out the possibility that there are other paths, such multiple switching strategies, which outperform the above strategy.

Simulation shows that the optimal switching strategy varies with the rate of loss of immunity (v). For $v \to \infty$ (SIRS equivalent to an SIS model), the best allocation strategy is always to give preference to the less infected sub-population (here giving preference to the less infected sub-population is equivalent to giving preference to the more susceptible sub-population). This observation agrees with Rowthorn et al. [10]. For $v \to 0$ (SIRS equivalent to an SIR model), a double switch of preference between the less and the more infected sub-population outperforms the other allocation strategies.

Details of the numerical explorations

Numerical simulation was done using a fourth order Runge-Kutta scheme with 0.01 time intervals. Experiments were done for different values of the period of integration and time intervals. The accuracy of our method was established up to three decimal places. The state variables were scaled with respect to the fixed sub-population size N.

To compare different control strategies, simulations were done using a large set of initial conditions (the state of the epidemic in each sub-population before resources are first allocated). For every single initial condition, we compared the value of the objective function for each of the control strategies described above. To build the set of initial condition, we proceeded as follows: for each sub-population, we spanned the surface $S + I + R = 1$ using an increment step of 0.02, excluding extreme cases such as $I = 1$ and $R = 1$. By crossing the initial conditions for the two sub-populations, we obtain a set of 705078 initial conditions for the whole system. The optimality of the single switch strategy was shown to hold for all initial conditions.

Comparing the proposed candidates for optimality is not enough to establish the optimality of a given solution. We used the same method as Rowthorn et al. [10]. We consider the paths that eventually reach set A. Any such path crosses the frontier of B at a unique point $(\hat{I}_1, \hat{I}_2)$. At this point, the transversality conditions determine a unique set of shadow prices $(\hat{m}_i)_{i=1}^4$. Taking $(\hat{S}_1, \hat{I}_1, \hat{S}_2, \hat{I}_2, (\hat{m}_i)_{i=1}^4)$ as initial conditions, we can reverse the systems of equations (1) and (9), thus tracking a path backward out of set A. Reversing a second time converts this path into a unique forward path that meets the set B at $(\hat{S}_1, \hat{I}_1, \hat{S}_2, \hat{I}_2)$ and also satisfies the Hamiltonian and transversality conditions. Using various points on the frontier of A, and suitable values of $(\hat{m}_i)_{i=1}^4$, we were not able to find another solution, satisfying the Hamiltonian and transversality conditions, that outperforms the switching strategies (also known as 'bang-bang' solutions [23]).

Author Contributions

Conceived and designed the experiments: MLNM CAG. Performed the experiments: MLNM. Analyzed the data: MLNM. Contributed reagents/materials/analysis tools: MLNM. Wrote the paper: MLNM CAG.

References

1. Lipsitch M, Bergstrom CT, Levin B (2000) The epidemiology of antibiotic resistance in hospitals: Paradoxes and prescriptions. Proc Natl Acad Sci USA 97: 1938–1943.
2. Kiszewski A, Johns B, Schapira A, Delacollette C, Crowell V, et al. (2007) Estimated global resources needed to attain international malaria control goals. Bull World Health Organ 85: 623–630.
3. Monto A (2006) Vaccines and antiviral drugs in pandemic preparedness. J Infect Dis 12: 55–60.
4. Dye C, Gay N (2003) Epidemiology: modeling the SARS epidemic. Science 300: 1884–1885.
5. Sani A, Kroesea D (2008) Controlling the number of HIV infectives in a mobile population. Math Biosci 213: 103–112.
6. May R, Anderson R (1984) Spatial heterogeneity and design of immunization programs. Math Biosci 72: 83–111.
7. Hethcote H, van Ark J (1987) Epidemiological models for heterogeneous populations: proportionate mixing, parameter estimation, and immunization programs. Math Biosci 84: 85–118.
8. Zaric G, Brandeau M (2002) Dynamic resource allocation for epidemic control in multiple populations. IMA J Math Appl Med Biol 19: 235–255.

9. Brandeau M, Zaric G, Ricther A (2003) Resource allocation for control of infectious diseases in multiple independent populations: beyond cost-effectiveness analysis. J Health Econ 22: 575–598.

10. Rowthorn R, Laxminaryan R, Gilligan C (2009) Optimal control of epidemics in metapopulations. JRSoc Interface 6: 1135–1144.

11. Keeling M, White P (2010) Targeting vaccination against novel infections: risk, age and spatial structure for pandemic influenza in great britain. JRSoc Interface;doi: 10.1098.

12. Dushoff J, Plotkin J, Viboud C, Simonsen L, Miller M, et al. (2007) Vaccinating to protect a vulnerable subpopulation. PLoS Med 4: e174.

13. Aron J (1988) Mathematical modeling of immunity to malaria. Mathematical Biosciences 90: 385–396.

14. Filipe J, Riley E, Drakeley C, Sutherland C, Ghani A (2007) Determination of the processes driving the acquisition of immunity to malaria using a mathematical transmission model. PLoS Comput Biol 3: e255.

15. Castillo-Chavez C, Feng Z (1997) To treat or not to treat: the case of tuberculosis. J Math Biol 35: 629–645.

16. Grassly N, Fraser C, Garnett G (2005) Host immunity and synchronized epidemics of syphilis across the united states. Nature 433: 417–421.

17. Forster G, Gilligan C (2007) Optimizing the control of disease infestations at the landscape scale. Proc Natl Acad Sci USA 104: 4984–4989.

18. Goldman S, Lightwood J (2002) Cost optimization in the SIS model of infectious disease with treatment. Top Econ Anal Policy 2: 1–22.

19. Hanski I (1998) Metapopulation dynamics. Nature 396: 41–49.

20. Keeling M, Grenfell TD (2000) Individual-based perspectives on R0. J Theor Biol 203: 51–61.

21. Strosberg M (2006) Allocating scarce resources in a pandemic: Ethical and public policy dimensions. Virtual Mentor (Ethics J Am Med Ass) 8: 241–244.

22. Kaplan E, Merson M (2002) Allocating hiv-prevention resources: balancing efficiency and equity. Am J Pub Health 92: 1905–1907.

23. Seierstad A, Sydsaeter K (1986) Optimal control theory with economic applications. New York, NY, USA: Elsevier North-Holland, Inc.

24. Agrachev A, Sachkov Y (2004) Control theory from the geometric viewpoint. Springer-Verlag, New York, in: encyclopedia of mathematical sciences, vol. 87. edition.

25. HHS (2007) hhs pandemic influenza plan. Technical report, United States Department of Health and Human Services. URL http://www.hhs.gov/pandemic/plan/sup6.html. accessed 2010 March 21.

26. Wu J, Riley S, Leung G (2007) Spatial considerations for the allocation of pre-pandemic influenza vaccination in the united states. Proc R Soc Lond B 274: 2811–2817.

27. Ndeffo-Mbah M, Gilligan C (2010) Optimization of control strategies for epidemics in heterogeneous populations with symmetric and asymmetric transmission. J Theor Biol 262: 757–763.

28. Medlock J, Galvani A (2009) Optimizing influenza vaccine distribution. Science 325: 1705–1708.

29. Wallinga J, van Bovan M, Lipsitch M (2010) Optimizing infectious disease interventions during an emerging epidemic. PNAS 107: 923–928.

30. Goldstein E, Apolloni A, Lewis B, Miller J, Macauley M, et al. (2010) Distribution of vaccine/antivirals and the 'least spread line' in a stratified population. J R Soc Interface 7: 755–764.

31. Tanner M, Sattenspiel L, Ntaimo L (2008) Finding optimal vaccination strategies under parameter uncertainty using stochastic programming. Math Biosci 215: 144–151.

32. Merl D, Johnson R, Gramacy B, Mangel M (2009) A statistical framework for the adaptive management of epidemiological interventions. PLoS ONE 4: e5087.

33. Ndeffo-Mbah M, Forster G, Wesseler J, Gilligan C (2010) Economically optimal timing of crop disease control in the presence of uncertainty: an options approach. JRSoc Interface 7: 1421–1428.

34. Heffernan J, Simth R, Wahl L (2005) Perspectives on the basic reproductive ratio. J R Soc Interface 2: 281–293.

35. Behncke H (2000) Optimal control of deterministic epidemics. Optim Contr Appl Meth 21: 269–285.

36. Dorfman R (1969) An economic interpretation of optimal control theory. Amer Econ Rev 59: 817–831.