Optimal Control of Treatments in a Two-strain Tuberculosis Model

Eunok Jung* Suzanne Lenhart† Zhilan Feng‡

September 5, 2001

Abstract

Optimal control theory is applied to a system of ordinary differential equations modeling a two-strain tuberculosis model. Seeking to reduce the latent and infectious groups with the resistant-strain tuberculosis, we use controls representing two types of treatments. The optimal controls are characterized in terms of the optimality system, which is solved numerically for several scenarios.

1 Introduction

In the absence of an effective vaccine, current control programs for TB have focused on chemotherapy. The antibiotic treatment for an active TB (with drug-sensitive strain) patient requires a much longer period of time and a higher cost than that for those who are infected with sensitive TB but have not developed the disease. Lack of compliance with drug treatments not only may lead to a relapse but to the development of antibiotic resistant TB – one of the most serious public health problems facing society today. A report released by the World Health Organization warns that if countries do not act quickly to strengthen their control of TB, the multidrug resistant strains that have cost New York City and Russia hundreds of lives and more than $1 billion each will continue to emerge in other parts of the world (see WHO, 2000a). The reduction in cases of drug resistant TB can be achieved either by "case holding", which refers to activities and techniques used to ensure regularity of drug intake for a duration adequate to achieve a cure (Chaulet, 1983), or by "case finding", which refers to the identification (through screening, for example) of individuals latently infected with sensitive TB who are at high risk of developing the disease and

*Oak Ridge National Laboratory, P.O. Box. 2008, Bldg. 6012, MS-6367, Oak Ridge, TN 37831-6367 (junge@ornl.gov).

†University of Tennessee, Mathematics Department, Knoxville, TN 37996-1300 and Oak Ridge National Laboratory (lenhart@math.utk.edu).

‡Purdue University, Mathematics Department, West Lafayette, IN 47907-1395 (zfeng@math.purdue.edu). This author's research was partially supported by NSF grant DMS-9974389.

who may benefit from preventive intervention (Reichman, 1993). These preventive treatments will reduce the incidence (new cases per unit of time) of drug sensitive TB and hence indirectly reduce the incidence of drug resistant TB.

Costs for activities to facilitate case holding and case finding may vary depending on many factors. For example, case holding can be very challenging because of the fact that chemotherapy must be maintained for several months to ensure a lasting cure, but patients usually recover their sense of well-being after only a few weeks of treatment and may often stop taking medications (Reichman. 1993). It has been reported by the Centers for Disease Control (CDC) that, in the United States, about 22% of patients currently fail to complete their treatment within a 12-month period and in some areas the failure rate reaches 55% (CDC, 1991). In the past few years, many places in the world have adopted the DOTS (directly observed therapy strategy in which public health nurses, community outreach workers, and others carry most of the responsibility for monitoring the patients during their course of treatment through home visits and administration. Although this program requires a relatively shorter period of time for the treatment, there was only about 24% of all TB patients who were treated through DOTS in 1999 (WHO, 2000b) For case finding, we mainly consider actions for the prevention of disease development with preventive therapy of latently infected persons with sensitive TB. There are several case finding methods. "Active case finding" refers to methods for the identification of TB cases where the first initiative patient/provider contact is taken by health care providers, where as "Passive case finding" refers to methods for the identification of TB cases where the first initiative patient/provider contact is taken by the patient. Another choice of case finding may be targeted screening activities among population groups at high risk of TB (immigrants from high prevalence countries, for example). Different methods have been shown to yield various levels of rewards in resource-poor and resource-rich countries (Nsanzumhire, 1981, Rieder, 1989), and the amount of resources required is also different.

Some past models of tuberculosis, particularly the predictive models attempting to calculate a threshold for the basic reproductive number $\mathcal{R}_0$, have incorporated drug treatment and/or vaccination, and have discussed control of the disease by looking at the role of disease transmission parameters in the reduction of $\mathcal{R}_0$ and the prevalence of the disease (see Blower et al. 1996, Blower and Gerberding, 1998, Blower et al. 1998, Castillo-Chavez and Feng, 1997, 2000). However, these models did not account for time dependent control strategies since their discussions are based on prevalence of the disease at equilibria. Time dependent control strategies have been studied for HIV models (see Kirschner et al, 1997, and Fister et al. 1998). Both approaches of studying control strategies produce valuable theoretical results which can be used to suggest or design epidemic control programs. Depending on a chosen goal (or goals) various objective criteria may be adopted.

In this article we consider (time dependent) optimal control strategies associated with case holding and case finding based on a two-strain TB model developed in Castillo-Chavez and Feng (1997). This model assumes that indi-

viduals in the latent stage develop active TB at a given rate. It also assumes that a proportion of treated individuals with active TB does not finish the treatment, of which a fraction will develop drug resistant TB. We introduce two control mechanisms representing case finding and case holding efforts into this model. The case finding effort is incorporated by adding a control term that identifies and cures a fraction of latent individuals so that the rate at which latent individuals develop the disease will be reduced. The case holding effort is incorporated by adding a control term that may lower the treatment failure rate of individuals with active sensitive TB so that the incidence of acquired drug-resistant TB will be reduced. We choose the reduction in cases of latent and infectious drug-resistant TB to be our main objective together with a lower cost of the controls.

This paper is organized as follows: Section 2 describes a two-strain TB model with two control terms. Our objective functional is also introduced in this section. The analysis of optimal controls is given in Section 3. Section 4 includes some numerical studies of optimal controls and discusses our results.

2 A Two-strain TB Model

Our state system is the following system of six ordinary differential equations from Castillo-Chavez and Feng (Castillo-Chavez and Feng, 1997):

$$\begin{aligned}
\dot{S} &= \Lambda - \beta S\frac{I_1}{N} - \beta^* S\frac{I_2}{N} - \mu S \\
\dot{L}_1 &= \beta S\frac{I_1}{N} - (\mu + k)L_1 - r_1 u_1(t)L_1 + (1-u_2(t))pr_2 I_1 + \beta' T\frac{I_1}{N} - \beta^* L_1\frac{I_2}{N} \\
\dot{I}_1 &= kL_1 - (\mu + d)I_1 - r_2 I_1 \\
\dot{L}_2 &= (1-u_2(t))qr_2 I_1 - (\mu + k')L_2 + \beta^*(S + L_1 + T)\frac{I_2}{N} \\
\dot{I}_2 &= k'L_2 - (\mu + d')I_2 \\
\dot{T} &= u_1(t)r_1 L_1 + (1-(1-u_2(t))(p+q))r_2 I_1 - \beta' T\frac{I_1}{N} - \beta^* T\frac{I_2}{N} - \mu T
\end{aligned} \tag{1}$$

with $S(0), L_1(0), I_1(0), L_2(0), I_2(0), T(0)$ given, where the host population is divided into the following epidemiological classes (state variables):
S: Susceptible
L_1: Latent, infected with typical TB but not infectious
I_1: Infectious with typical TB
L_2: Latent, infected with resistant strain TB but not infectious
I_2: Infectious with resistant strain TB
T: Treated (effectively),
$N = S + L_1 + I_1 + L_2 + I_2 + T$.

We assume that an individual may be infected only through contacts with

infectious individuals. The term Λ is the recruitment rate. Coefficients β and β' are the rates at which susceptible and treated individuals become infected by an infectious individual with typical TB per unit of time. Coefficient β^* is the rate at which an uninfected individual becomes infected by one resistant-TB infectious individual per unit of time. The per-capita natural death rate is μ while the per-capita disease induced death rates are d and d' for the typical TB and resistant TB, respectively. The rates at which an individual leaves the two latent classes by becoming infectious are k and k'. The term $p+q$ is the proportion of those treated infectious individuals, who did not complete their treatment ($p+q \leq 1$).

The control functions, $u_1(t)$ and $u_2(t)$, are bounded, *Lebesque* integrable functions. The "case finding" control, $u_1(t)$, represents the fraction of typical TB latent individuals that is identified and will be put under treatment (to reduce the number of individuals that may be infectious). The coefficient, $1-u_2(t)$, represents the effort that prevents the failure of the treatment in the typical TB infectious individuals (to reduce the number of individuals developing resistant TB). When the "case holding" control, $u_2(t)$ is near 1, there is low treatment failure and high implementation costs.

Our objective functional to be minimized is

$$J(u_1,u_2) = \int_0^{t_f} [L_2(t) + I_2(t) + \frac{B_1}{2}u_1^2(t) + \frac{B_2}{2}u_2^2(t)]dt \tag{2}$$

where we want to minimize the latent and infectious groups with resistant-strain TB while also keeping the cost of the treatments low. We assume the cost of the treatments are nonlinear and take quadratic form here. The coefficients, B_1 and B_2, are balancing cost factors due to size and importance of the three parts of the objective functional. We seek to find an optimal control pair, $u_1^\star$ and $u_2^\star$, such that

$$J(u_1^\star, u_2^\star) = \min_{\Omega} J(u_1,u_2). \tag{3}$$

where $\Omega = \{(u_1,u_2) \in L^1(0,t_f) \mid a_i \leq u_i \leq b_i, i=1,2\}$ and $a_i, b_i, i=1,2$ are fixed positive constants.

In our analysis, we assume the total population N to be constant. To guarantee this, we choose $\Lambda = \mu N, d = d' = 0$.

3 Analysis of Optimal Controls

The necessary conditions that an optimal pair must satisfy come from Pontryagin's Maximum Principle (Pontryagin, 1962). This principle converts (1) - (3)

into a problem of minimizing pointwise a Hamiltonian, H, with respect to u_1 and u_2:

$$H = L_2 + I_2 + \frac{B_1}{2}u_1^2 + \frac{B_2}{2}u_2^2 + \sum_{i=1}^{6} \lambda_i g_i \tag{4}$$

where g_i is the right hand side of the differential equation of the ith state variable. By applying Pontryagin's Maximum Principle (Pontryagin, 1962) and the existence result for the optimal control pairs from (Fleming and Rishel, 1975), we obtain

Theorem 1. *There exists an optimal control pair $u_1^\star$, $u_2^\star$ and corresponding solution, $S^\star$, $L_1^\star$, $I_1^\star$, $L_2^\star$, $I_2^\star$, and $T^\star$, that minimizes $J(u_1, u_2)$ over Ω. Furthermore, there exists adjoint functions, $\lambda_1(t), \ldots, \lambda_6(t)$, such that*

$$\begin{aligned}
\dot{\lambda}_1 &= \lambda_1(\beta\frac{I_1^\star}{N} + \beta^*\frac{I_2^\star}{N} + \mu) + \lambda_2(-\beta\frac{I_1^\star}{N}) + \lambda_4(-\beta^*\frac{I_2^\star}{N}) \\
\dot{\lambda}_2 &= \lambda_2(\mu + k + u_1(t)r_1 + \beta^*\frac{I_2^\star}{N}) + \lambda_3(-k) + \lambda_4(-\beta^*\frac{I_2^\star}{N}) + \lambda_6(-u_1^\star(t)r_1) \\
\dot{\lambda}_3 &= \lambda_1(\beta\frac{S^\star}{N}) + \lambda_2(-\beta\frac{S^\star}{N} - (1-u_2^\star(t))pr_2 - \beta'\frac{T^\star}{N}) + \lambda_3(\mu + d + r_2) \\
&\quad + \lambda_4(-(1-u_2^\star(t))qr_2) + \lambda_6(-(1-(1-u_2^\star(t))(p+q))r_2 + \beta'\frac{T^\star}{N}) \\
\dot{\lambda}_4 &= -1 + \lambda_4(\mu + k') + \lambda_5(-k') \\
\dot{\lambda}_5 &= -1 + \lambda_1(\beta^*\frac{S^\star}{N}) + \lambda_2(\beta^*\frac{L_1^\star}{N}) + \lambda_4(-\beta^*\frac{S^\star + L_1^\star + T^\star}{N}) + \lambda_5(\mu + d') + \lambda_6(\beta^*\frac{T^\star}{N}) \\
\dot{\lambda}_6 &= \lambda_2(-\beta'\frac{I_1^\star}{N}) + \lambda_4(-\beta^*\frac{I_2^\star}{N}) + \lambda_6(\beta'\frac{I_1^\star}{N} + \beta^*\frac{I_2^\star}{N} + \mu)
\end{aligned} \tag{5}$$

with transversality conditions

$$\lambda_i(t_f) = 0, i = 1, \ldots, 6 \tag{6}$$

and $N = S^\star + L_1^\star + I_1^\star + L_2^\star + I_2^\star + T^\star$.

The following characterization holds

$$\begin{aligned}
u_1^\star(t) &= min(max(a_1, \frac{1}{B_1}(\lambda_2 - \lambda_6)rL_1^\star), b_1) \\
&\textit{and} \\
u_2^\star(t) &= min(max(a_2, \frac{1}{B_2}(\lambda_2 p + \lambda_4 q - \lambda_6(p+q)r_2 I_1^\star)), b_2).
\end{aligned} \tag{7}$$

Proof. Corollary 4.1 of (Fleming and Rishel, 1975) gives the existence of an optimal control pair due to the convexity of integrand of J with respect to (u_1, u_2), a *priori* boundedness of the state solutions, and the *Lipschitz* property of the

state system with respect to the state variables. Applying Pontryagin's Maximum Principle, we obtain

$$\frac{d\lambda_1}{dt} = -\frac{\partial H}{\partial S}, \lambda_1(t_f) = 0,$$
$$\cdots$$
$$\frac{d\lambda_6}{dt} = -\frac{\partial H}{\partial T}, \lambda_6(t_f) = 0,$$

evaluated at the optimal control pair and correspoding states, which results in the stated adjoint system (5) and (6) (Kamien and Schwartz, 1991). By considering the optimality conditions,

$$\frac{\partial H}{\partial u_1} = 0, \frac{\partial H}{\partial u_2} = 0$$

and solving for $u_1^\star$, $u_2^\star$, subject to the constraints, the characterizations (7) are derived. □

Due to the a *priori* boundedness of the state and adjoint functions and the resulting *Lipschitz* structure of the ODEs, we obtain the uniqueness of the optimal control for small t_f. The uniqueness of the optimal control pair follows from the uniqueness of the optimality system, which is (1) and (5), (6) with characterizations (7). There is a restriction on the length of the time interval in order to guarantee uniqueness of the optimality system. This smallness restriction on the length on the time interval is due to the opposite time orientations of (1) and (5), (6); the state problem has initial values and the adjoint problem has final values. This restriction is very common in control problems (Fister et. al, 1998, Kirschner et. al, 1997).

Next, we discuss the numerical solutions of the optimality system and the corresponding optimal control pairs, the parameter choices, and the interpretations from various cases.

4 Numerical Results

In this section, we study numerically an optimal treatment strategy of our two-strain TB model. The optimal treatment strategy is obtained by solving the optimality system, consisting of 12 ODEs from the state and adjoint equations. An iterative method is used for solving the optimality system. We start to solve the state equations with a guess for the controls over the simulated time using a forward fourth order Runge-Kutta scheme. Because of the transversality conditions (6), the adjoint equations are solved by a backward fourth order

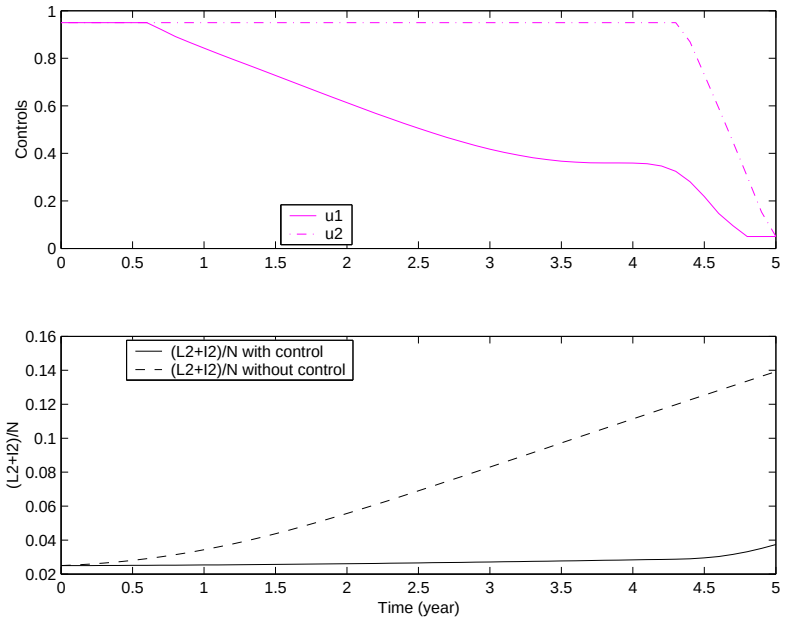

Figure 1: The optimal control strategy for the case of $B_1 = 50$ and $B_2 = 500$.

Runge-Kutta scheme using the current iteration solution of the state equations. Then, update the controls by using a convex combination of the previous controls and the value from the characterizations (7). Repeat this process and stop iterations if the values of unknowns at the previous iteration are very close to the ones at the present iteration.

For the figures presented here, we assume that the weight factor B_2 associated with control u_2, is bigger or equal to B_1, associated with a control u_1. This assumption is based on following facts: The cost associated with u_1 will include the cost of screening and treatment programs and the cost associated with u_2 will include the cost of holding the patients in the hospital or sending people to watch the patients to finish their treatment. Treating infectious TB takes longer (by several months) than treating latent TB. In these three figures, the set of the weight factors, $B_1 = 50$ and $B_2 = 500$, is chosen to illustrate the optimal treatment strategy. Other epidemiological and numerical parameters are presented in Tables 1 and 2, respectively. The initial values of the variables are given in Table 3. We will discuss briefly the cases with different values of B_1 and B_2 later in this section.

Figure 1 shows the optimal treatment strategy for the case of $B_1 = 50$ and $B_2 = 500$. In the top frame, the controls, u_1 (solid curve) and u_2 (dashdot curve), are plotted as a function of time. In the bottom frame, the fractions of individuals infected with resistant TB, $(L_2 + I_2)/N$, with control (solid curve)

and without control (dashed curve) are plotted. Parameters $N = 30000$ and $\beta^* = 0.029$ are chosen. Other parameters are in the Tables 1 and 2. To minimize the total number of the latent and infectious individuals with resistant TB, $L_2 + I_2$, the optimal control u_2 is at the upper bound during almost 4.3 years and then u_2 is decreasing to the lower bound, while the steadily decreasing value for u_1 is applied over the most of the simulated time, 5 years. The total number of individuals $L_2 + I_2$ infected with resistant TB at the final time $t_f = 5$ (years) is 1123 in the case with control and 4176 without control, and the total cases of resistant TB prevented at the end of the control program is 3053.

Figure 2 illustrates how the optimal control strategies depend on the parameter β^*, which denotes the transmission rate of primary infections of resistant TB. The value of β^* is usually given by the product of the number of contacts (with an infectious individuals with resistant TB) per person per unit of time and the probability of being infected with resistant TB per contact. This value varies from place to place depending on many factors including living conditions. In Figure 2, the controls, u_1 (dark color curves) and u_2 (light color curves), are plotted as a function of time for the 4 different values of β^*, 0.0131, 0.0217, 0.0290, and 0.0436. These values for β^* are chosen from (Castillo-Chavez and Feng, 1997). Other parameters are in the Tables 1 and 2. Figure 2 shows that u_1 plays an increasing role while u_2 remains almost the same as β^* decreases. This is an expected result because when β^* is smaller, the new cases of resistant TB arise more from infections acquired from L_1 and I_1 due to treatment failure than from primary infections. In this case, identifying and curing latently infected individuals with sensitive TB becomes more important in the reduction of new cases of resistant TB.

In Figure 3, the controls, u_1 and u_2, are plotted as a function of time for $N = 6000$, 12000, and 30000 in the top and bottom frame, respectively. Other parameters except the total number of individuals and $\beta^* = 0.029$ are fixed for these three cases and presented in the Tables 1 and 2. These results show that more effort should be devoted to "case finding" control u_1 if the population size is small, but "case holding" control u_2 will play a more significant role if the population size is big.

Note that, in general, with B_1 fixed, as B_2 increases, the amount of u_2 decreases. A similar result holds if B_2 is fixed as B_1 increases.

In conclusion, our optimal control results tell whether case holding or case finding efforts should dominate, depending on the population size, cost of implementing treatments controls and the parameters of the model. Following these strategies can effectively lower the number of latent and infectious resistant-strain TB cases.

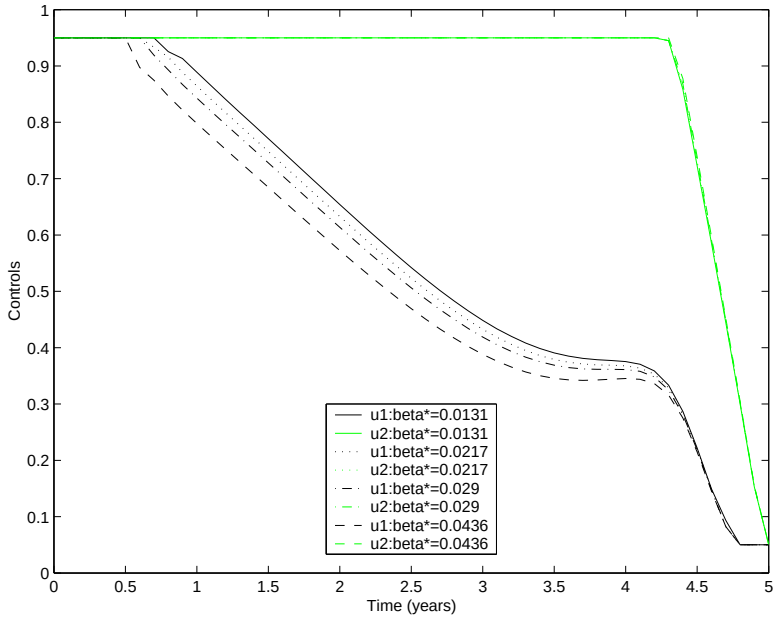

Figure 2: The controls, u_1 (dark color curves) and u_2 (light color curves), are plotted as a function of time for the 4 different values of β^*, 0.0131, 0.0217, 0.0290, and 0.0436.

Parameters	Values
β	13
β'	13
β^*	0.0131, 0.0217, 0.029, 0.0436
mu	0.0143
d	0
d'	0
k	0.5
k'	1
r_1	2
r_2	1
p	0.4
q	0.1
N	6000, 12000, 30000
Λ	μN

Table 1: Parameters and their values

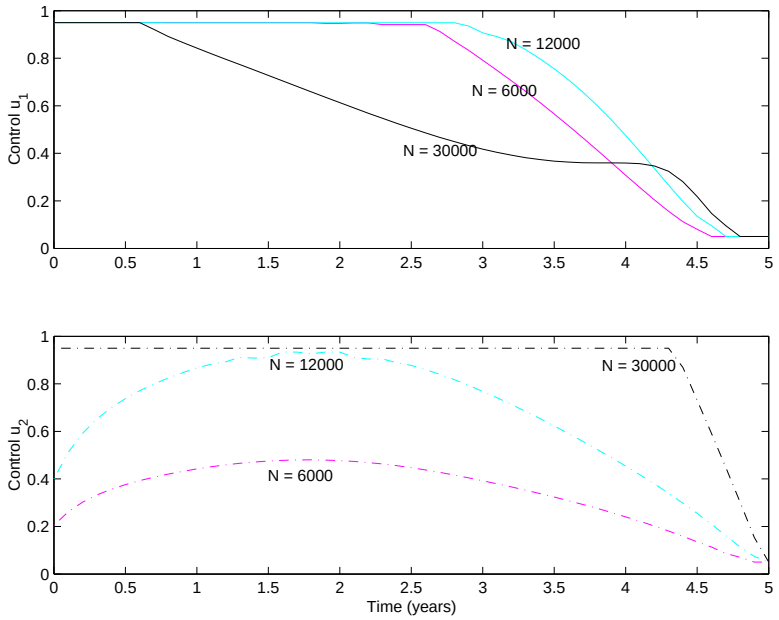

Figure 3: The controls, u_1 and u_2, are plotted as a function of time for $N = 6000$, 12000, and 30000 in the top and bottom frame, respectively.

Computational parameters	Symbol	
Final time	t_f	5 years
Timestep duration	dt	0.1 year
Upper bound for controls		0.95
Lower bound for controls		0.05
Weight factor associated with u_1	B_1	50
Weight factor associated with u_2	B_2	500

Table 2: Computational parameters

Variables	Initial values
S	$(76/120)N$
L_1	$(36/120)N$
I_1	$(4/120)N$
L_2	$(2/120)N$
I_2	$(1/120)N$
T	$(1/120)N$

Table 3: Initial values for individuals

REFERENCES

S.M. Blower, P.M. Small and P.C. Hopewell. 1996. Control strategies for tuberculosis epidemics: new models for old problems. Science 273:497-500.

S.M. Blower, T. Porco and T. Lietman. 1998. Tuberculosis: The evolution of antibiotic resistance and the design of epidemic control strategies. In Mathematical Models in Medical and Health Sciences. Eds Horn, Simonett, Webb. Vanderbilt University Press.

S.M. Blower and J. Gerberding. 1998. Understanding, predicting and controlling the emergence of drug-resistant tuberculosis: a theoretical framework. Journal of Molecular Medicine 76: 624-636.

C. Castillo-Chavez and Z. Feng. 1997. To treat or not to treat: the case of tuberculosis. J. Mathematical Biology, 35, 629-659.

C. Castillo-Chavez and Z. Feng. 1998. Global stability of an age-structure model for TB and its applications to optimal vaccination strategies. Mathematical Biosciences, 151, 135-154.

Center for Disease Control. 1991. Unpublished data. Division of Tuberculosis Elimination. CDC.

P. Chaulet. 1983. Treatment of tuberculosis: case holding until cure. WHO/TB/83. 141. World Health Organization, Geneva.

K.R. Fister, S. Lenhart and J.S. McNally. 1998. Optimizing chemotherapy in an HIV model. Electronic J. Differential Equations, 1-12.

D. Kirschner, S. Lenhart and S. Serbin. 1997. Optimal control of the chemotherapy of HIV. J. Mathematical Biology. 35, 775-792.

H. Nsanzumuhirc et al. 1981. A third study of case-finding methods for pulmonary tuberculosis in Kenya, including the use of community leaders. Tubercle. 62. 79-94.

L.B. Reichman and E.S. Hershfield. 1993. Tuberculosis: a comprehensive international approach. New York.

H.L. Rieder et al. 1989. Epidemiology of tuberculosis in the United States. Epidemiol. Rev. 11. 89-95

W. H. Fleming and R. W. Rishel. 1975. Deterministic and Stochastic Optimal Control. Springer Verlag, New York.

M. I. Kamien and N. L. Schwarz. 1991. Dynamic optimization: the calculus of variations and optimal control. North Holland, Amsterdam.

L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. 1962. The mathematical theory of optimal processes. Wiley, New York.

WHO. 2000a. Press Release WHO/19, 24 March 2000.

WHO. 2000b. Tuberculosis: strategy & operation. www.who.int/gtb/dots.