MATHEMATICAL BIOSCIENCES
AND ENGINEERING
Volume **1**, Number **2**, September **2004**

http://math.asu.edu/˜mbe/

pp. **361–404**

DYNAMICAL MODELS OF TUBERCULOSIS AND THEIR APPLICATIONS

Carlos Castillo-Chavez

Department of Mathematics and Statistics
Arizona State University, Tempe, AZ 85287-1804

Baojun Song

Department of Mathematical Sciences
Montclair State University, Upper Montclair, NJ 07043

(Communicated by Yang Kuang)

Abstract. The reemergence of tuberculosis (TB) from the 1980s to the early 1990s instigated extensive researches on the mechanisms behind the transmission dynamics of TB epidemics. This article provides a detailed review of the work on the dynamics and control of TB. The earliest mathematical models describing the TB dynamics appeared in the 1960s and focused on the prediction and control strategies using simulation approaches. Most recently developed models not only pay attention to simulations but also take care of dynamical analysis using modern knowledge of dynamical systems. Questions addressed by these models mainly concentrate on TB control strategies, optimal vaccination policies, approaches toward the elimination of TB in the U.S.A., TB co-infection with HIV/AIDS, drug-resistant TB, responses of the immune system, impacts of demography, the role of public transportation systems, and the impact of contact patterns. Model formulations involve a variety of mathematical areas, such as ODEs (Ordinary Differential Equations) (both autonomous and non-autonomous systems), PDEs (Partial Differential Equations), system of difference equations, system of integro-differential equations, Markov chain model, and simulation models.

1. **Introduction.** Tuberculosis (TB) is a disease that affects human and animal populations. Ancient Egyptian mummies show deformities consistent with tubercular decay [20, 23]. TB was probably transmitted from animals to humans in areas where agriculture became dominant and animals were domesticated. The growth of human communities probably increased the recurrence of TB epidemics leading to its currently overwhelmingly high levels of endemicity in some developing nations. McGrath estimates that a social network of 180 to 440 persons is required to achieve the stable host pathogen relationship necessary for TB infection to become endemic in a community [58]. Historically the terms phthisis, consumption and white plague were used as synonym for TB.

TB was a "fatal" disease. In earlier times, some physicians refused to visit the late-stage TB to keep their reputation. TB was responsible for at least one billion deaths during the nineteenth and early twentieth century and the leading cause of

2000 *Mathematics Subject Classification.* 92D30.

Key words and phrases. Tuberculosis, dynamical models, prevention and control, global dynamics, bifurcation, HIV, disease emergence and reemergence.

human death for centuries. Today, "only" 3 million deaths worldwide are attributed to TB every year. World Health Organization's (WHO) data shows that most cases of TB are in developing countries. Twenty three counties in East Asia and Africa account for over 80% of all cases around the world [84].

It was not clear how TB was transmitted until Robert Koch's brilliant discovery of the tubercle bacillus in 1882 (Koch also identified the cause of anthrax). He identified *Mycobacterium tuberculosis* as the causative agent of TB. The tubercle bacilli live in the lungs of infected hosts. They spread in the air when infectious individuals sneeze, cough, speak or sing. A susceptible individual may become infected with TB if he or she inhales bacilli from the air. The particles containing *Mycobacterium tuberculosis* are so small that normal air currents keep them airborne and transport them throughout rooms or buildings [83]. Hence, individuals who regularly share space with those with active TB (the infectious stage of the disease) have a much higher risk of becoming infected. These bacilli become established in the alveoli of the lungs from where they spread throughout the body if not suppressed by the immune system.

The hosts' immune responses usually limit bacilli multiplication and, consequently, the spread that follows initial infections. About 10% of infected individuals eventually develop active TB. Most infected individuals remain as latently infected carriers for their entire lives. The average length of the latent period (noninfectious stage) ranges from months to decades. However, the risk of progression toward active TB increases markedly in the presence of co-infections that debilitate the immune system. Persons with HIV co-infections progress faster towards the active TB state than those without them [68].

Most forms of TB can be treated. Effective and widespread treatment for active and latently infected individuals has been available for about five decades. *Streptomycin* is still used today to treat TB but in combination with *pyrazinamide*.

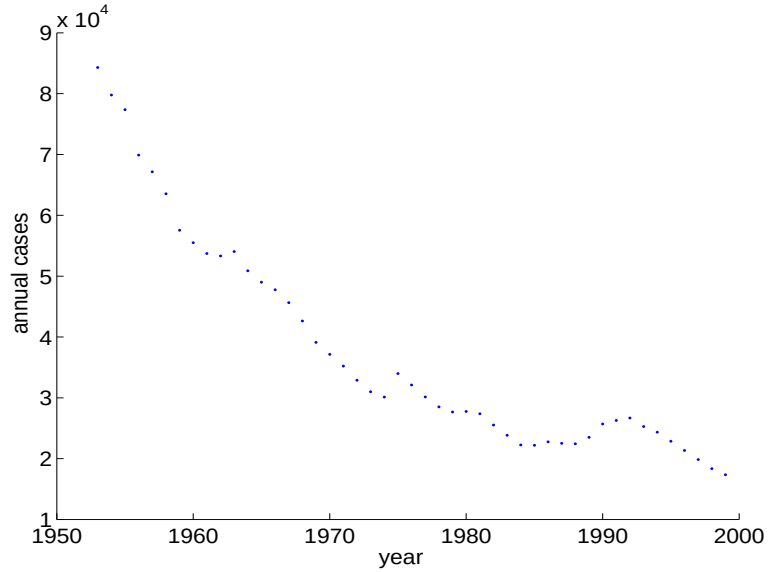


FIGURE 1. Annual new cases of TB in the United States from 1953 to 2000. Data taken from [22].

Isoniazid and *rifampin* are thought to be the most effective in the fight against M. tuberculosis. The widespread introduction of antibiotics reduced mortality by 70% from 1945 to 1955 in the U.S.A. albeit most major reductions in TB mortality rates had already been achieved before their introduction [4, 29, 54]. Latent TB can be handled with *isoniazid* but treatment is effective only if applied for at least six months. Active cases must be treated for nine months with multiple drugs (*isoniazid, rifampin, pyrazinamide*) and complex regimens. Treatment covers over 95% of the cases in the U.S.A. despite its high cost [85]. Antibiotic-resistant strains are easily generated when treatment is not completed. The consequences of incomplete treatment may be serious [15]. Lack of treatment compliance has serious consequences due to its dramatic impact on the evolution of antibiotic resistant strains [49]. The expenses associated with treatment programs for those with active TB are so high that their effective implementation is out of the reach for most developing nations.

As shown in Fig. 1, the mortality associated with TB in the U.S.A. continues to exhibit a downward trend. The annual case rate of TB had been declining steadily but raised slightly in the 1980s and early 1990s in the U.S.A.. The change in this trend had been labeled as a period of TB reemergence. TB reemergence over the past decade and a half has challenged existing prevention and control TB programs in developing nations.

In this paper, we review some of the literature associated with TB models and their theoretical impact–particularly those aspects where the authors or their collaborators have contributed. Some results appearing in this paper by Song and Castillo-Chavez have not been published.

The paper is organized as follows. Section 2 introduces the notation that we try to use throughout the manuscript. Section 3 reviews some of the earliest known TB models. Section 4 deals with the exploration of the impact of various epidemi-

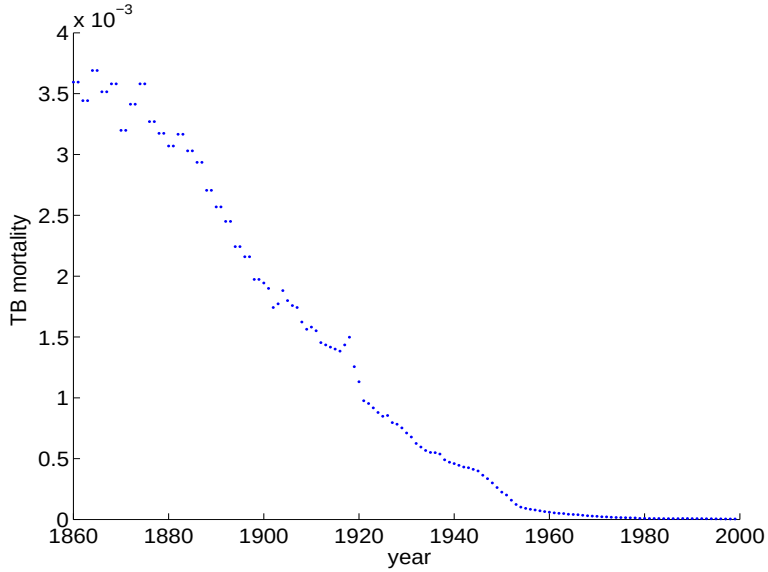


FIGURE 2. TB mortality of the United States from 1860 to 2000. Data taken from [78, 79].

ological factors as well as the role of close and casual contacts on TB dynamics. section 5 looks at the impact of demography on TB dynamics. In section 6, we review some cell-based models for TB transmission at the immune system level. A Markov chain model on TB projections is described in section 7. Models dealing with TB control strategies are discussed in section 8. A model dealing with the role of public mass transportation on TB evolution and control is reviewed in section 9. Finally, a list of challenges associated with modeling TB dynamics is outlined.

2. **Notation.** The population of interest is divided into several compartments (classes, categories, or subpopulations) dictated by the epidemiological stages (host statues). For the most part, in the context of TB, four or five epidemiological stages are identified (see Table 1). We shall do our best to denote these subclasses using uniform symbols as we discuss a multitude of models. Tables 1 and 2 list the definitions and symbols (subpopulations and parameters) that we try to use. If it is necessary to subdivide a population into subpopulations, subscripts will be

TABLE 1. Symbols and definitions of subpopulations

symbol	name	definition
S	Susceptible	not infected but susceptible to infection
E	Exposed	infected but unable to infect others (latent or carrier)
I	Infectious	active-TB infections, i.e., he/she can infect others
T	Treated	treated (from latent-TB or active-TB infection)
V	Vaccinated	possibly reduced susceptibility to TB

used to distinguish them. For instance, I_s and I_r represent the drug-sensitive and drug-resistant infectious TB classes, respectively. *Active TB, case TB, index TB, mature TB, open case,* and *lesion case* all mean active-TB infectious case here.

We use β to measure of the likelihood of transmission or the "force" of infection. However, the meaning of β or its interpretation often changes from model to model.

3. **Early dynamical models.** The first model for the transmission dynamics of TB was built in 1962 by Waaler [81], the chief statistician of the Norwegian Tuberculosis Control Services. Waaler divided the population into three epidemiological classes: noninfected (susceptible), infected non-cases (latent TB), and infected cases (infectious). He formulated the infection rate as an unknown function of the number of infectious individuals. He used a particular linear function to model infection rates in the implementation of his model. The incidence (new cases of infections per unit time) was assumed to depend only on the number of infectious. Furthermore, the equations for the latent and infectious classes were assumed to be uncoupled from the equation for the susceptible class. The central part of this model is given by the following linear system of difference equations:

$$E_{t+1} = E_t + aI_t + eE_t - d_2E_t - gE_t, \tag{1}$$

$$I_{t+1} = I_t + gE_t - d_3I_t - eE_t, \tag{2}$$

TABLE 2. Symbols and definitions of parameters.

symbol	explanation
Λ	recruitment rate
β	transmission rate (meaning varies)
c	average number of contacts per person per unit time
k	per-capita regular progression rate
μ	per-capita natural mortality rate
d	per-capita excess death rate due to TB
r_0	per-capita treatment rate for recently latently-infected
r_1	per-capita treatment rate for latently-infected
r_2	per-capita treatment rate for actively-infected
ω	per-capita progression rate for early latent-TB progression

where the incidence rate aI_t is proportional to the number of infectious; e is the per-capita progression rate from latent-TB to infectious-TB cases; g is the per-capita treatment rate (treated individuals will become members of latent-TB class again.); d_2 is the per-capita death rate of the latent-TB class; and d_3 is the per-capita death rate of the infectious-TB class. Using data from a rural area in south India for the period of 1950 to 1955, Waaler [35] estimated the parameters of this linear model to be $a = 1, e = 0.1, d_2 = 0.014, g = 0.10085, d_3 = 0.07$. Because the eigenvalues all have norms close to 1 (1.04), Waaler predicted that the time trend of TB is unlikely to increase (it may decrease, albeit slowly). This linear model did not model the mechanics of transmission. However, the parameters, estimated from a specific area in India, set useful ranges for the estimation of parameters in developing nations.

Brogger developed a model [10] that improved on Waaler's. Brogger not only introduced heterogeneity (age) but also changed the method used for calculating infection rates. The infection rate in Brogger's model was a combination of linear and nonlinear infection terms. In fact, it was given by the term $\beta S(1 - Z + Z\frac{I}{N})$, where Z was an adjusting parameter used to differentiate between normal infection, superinfection, and direct leaps (within a very short period, an uninfected individual becomes a lesion case or an active-TB case). Two extreme cases were covered in the model: $Z = 1$ making the incidence be $\beta S\frac{I}{N}$, the familiar version of today, and, $Z = 0$ giving an infection rate proportional to the number of susceptibles. The prevalence $\frac{I}{N}$ was used to adjust all flow rates including those from infected to open cases. This was not surprising as Brogger wanted to use prevalence as an indicator of the effectiveness of control policies. His aim was to compare different control strategies that included finding and treating more cases, the utilization of vaccination, and mass roentgenograph. The data of two WHO/UNICEF projects in Thailand from 1960 to 1963 were used to estimate the parameters incorporated into his model. Brogger chose those parameters that "best" fits available data. Control strategies (additional new parameters) were "squeezed" into the model. Simulations

were run and comparisons made to evaluate the value of different strategies of TB control programs. This model did not formulate clearly the relationship between infection rate and prevalence. ReVelle classified this important relationship in 1967.

Using Brogger and Waaler's model as a template, ReVelle introduced the first nonlinear system of ordinary differential equations that models TB dynamics [63, 64]. In modeling the infection rate, he did not follow the typical mass action law, given by the bilinear function βSI of Kermack and McKendrick [47]. It was ReVelle who first, at least in the context of TB dynamics, rigorously explained why the infection rate depends linearly on the prevalence using the probabilistic approach that is common today (homogeneous mixing). The form $\beta S\frac{I}{N}$ for the infection rate is found in most epidemic models used today. Mathematically it is well known that if the total population size N remains constant over time or if it asymptotically approaches a constant then the use of an infection rate proportional to SI does not change the qualitative properties of the model. However, when modeling epidemics for developing countries, as Revell did with his model, $\beta S\frac{I}{N}$ seems a more appropriate form of modeling the infection rate. ReVelle modeled TB dynamics via a system of non-linear differential equations but he ignored population structure. Nine compartments were introduced in Revell's nonlinear model. The total population was governed by the Malthus model because he wanted to apply it to developing nations. Making projects (in Waaler's words "time trend of tuberculosis") was not Revell's main theme. In fact, his main objective seemed to be associated with the evaluation and implementation of control polices and their cost. He developed an optimization model and used it to select control strategies that could be carried out at a minimal cost. It is worth to mention that Waaler also developed a model in 1970 that would minimize the cost of alternative tuberculosis control measures [82].

All dynamical models prior to Ferebee's work were motivated by the study of TB in developing nations. The two data sets used were from Thailand and India. No specific model seems to have been developed for the U.S.A.. Ferebee, associate chief of the research section of the Public Health Service Tuberculosis Program of U.S.A., changed this trend. She set up a discrete model, based upon a set of simple assumptions, to model the dynamics of TB in the U.S.A. [34]. She used the same compartments as Waaler did, that is, susceptible, infected and infectious. The basic time unit was a year. Hence, within one year, new infected people would become infectious and contribute to the pool of new infected, namely, some individuals would move from the susceptible to the infected class and from the infected to the infectious class within a year. Ferebee described her algorithm, methods of estimation of relevant parameters, and the number of infected people in the U.S.A.. The results showed that the number of new cases would decrease, but slowly, if vaccination were not applied to the US population. It is worth mentioning that the estimation of demographic parameters was solely based on the 1963 US data. Despite its shortcomings, this work indeed gave the first rough estimate and forecast of TB cases in the U.S.A.. She stated that her assumptions were checked for consistency with bits and pieces of information obtained from a variety of sources. These assumptions have since played an important role theoretically and practically in the context of TB. In other words, the first US study "defined" an appropriate parameter range. We shall outline underlying assumptions:

i. There are 25 million infected individuals and 125 million susceptible;

ii. The per-capita progression rate from infected to infectious is $k_1 = 1/625$ per year;
iii. Primary infected people exhibit a higher per-capita progression rate in the first year ($k_2 = 1/12$); that is, one out of every 12 new infections will progress to the infectious stage during the first year;
iv. Each new infectious individuals will infect 3 people;
v. No significant additional death is ascribed to tuberculosis.

Her assumption $k_2 >> k_1$ has directed model simulations and constructions. This assumption has been recently incorporated via the inclusion of additional compartments for fast TB and slow TB (see section 4.1).

Earlier mathematical models for TB transmission were developed mostly by statisticians. Their approach followed a pattern: build a mathematical model for TB transmission; with help from a data set, estimate parameters; find numerical solutions; and predict or make inferences about the relative value of alternative control strategies. There was no qualitative analysis of the models, and the long-time behavior (asymptotic properties) of the models was also not studied.

The continuous decline of TB incidence in developed nations and the introduction of effective antibiotics suggested that elimination of active TB in developed nations was possible. This view may have been the main reason why there was almost no theoretical work on TB dynamics from the 1970s to the early 1990s. The story has changed over the last decade because of the reemergence of TB (new outbreaks in the U.S.A. and in many developed nations). In the following sections we shall review some of the most recent models and the theoretical results.

4. **Intrinsic mechanics of transmission.**

4.1. **Slow and fast routes.** The initially exposed individuals (infected individuals) have a higher risk of developing active TB. With time passing, those individuals still face the possibility of progressing to infectious TB, but the rate of progression slows down. In other words, the likelihood of becoming an active infectious case decreases with the age of the infection. Bearing this in mind, several researchers constructed a series of dynamical models for TB progression and transmission in scenarios that took these factors into consideration [6, 7, 8, 19, 32, 60]. We shall review some of this work. In the simplest model (that we know), the population of interest is partitioned into three epidemiological classes: susceptible, latent, and infectious. The infection rate given by βSI (using the mass action law) is divided. A portion $p\beta SI$ gives rise to immediate active cases (fast progression), while the rest $(1-p)\beta SI$ gives rise to latent-TB cases with a low risk of progressing to active TB (slow progression). The progression rate from latent TB to active TB is assumed to be proportional to the number of latent-TB cases, that is, it is given by kE, where k ranges from 0.00256 to 0.00527 (slow progression). The total incidence rate is $p\beta SI + kE$. The version in [6] is given by following system:

$$\frac{dS}{dt} = \Lambda - \beta SI - \mu S, \tag{3}$$

$$\frac{dE}{dt} = (1-p)\beta SI - kE - \mu E, \tag{4}$$

$$\frac{dI}{dt} = p\beta SI + kE - dI - \mu I, \tag{5}$$

where the parameters are defined in Table 2. The qualitative dynamics of model (3–5) are governed by the *basic reproductive number*

$$\mathcal{R}_0 = p\frac{\Lambda}{\mu}\frac{\beta}{\mu+d} + (1-p)\frac{\Lambda}{\mu}\frac{k}{\mu+k}. \tag{6}$$

This dimensionless quantity measures the average number of secondary infectious cases produced by a "typical" infectious individual in a population of susceptibles at a demographic steady state. The first term in (6) gives the new cases resulting from fast progression while the second those resulting from slow progression. Sensitivity and uncertainty analysis were carried out. Simulation results showed that TB dynamics were quite slow for acceptable parameter ranges. Waaler's model also supported slow TB dynamics [81]. Model (3–5) requires that p be known a priori (it is not allowed to change) and, the $\mathcal{R}_0$ derived from the model depends linearly on population size. A model that removes these restrictions is reviewed next.

4.2. **Variable latent period.** Instead of assuming exponential distribution of latency period, Feng studied a model with an arbitrary distribution for the latency period [31]. To describe this model, we let $p(s)$ be a function representing the proportion of those individuals exposed s units of time ago and who, if alive, are still infected (but not infectious) at time s. The removal rate of individuals from the E class into the I class τ units of time after exposure is given by $-\dot{p}(\tau)$. Hence, the total number of exposed individuals from the initial time $t = 0$ to the current time t, who are still in the E class, is given by the integral

$$\int_0^t \beta c S(s)\frac{I(s)}{N(s)}p(t-s)e^{-(\mu+r_1)(t-s)}ds$$

while the number of individuals who develop infectious TB cases from 0 to t, who are still alive and in the I class, is given by the double integral

$$\int_0^t\int_0^\tau \beta c S(s)\frac{I(s)}{N(s)}e^{-(\mu+r_1)(\tau-s)}\left(-\dot{p}(\tau-s)e^{-(\mu+r_2+d)(t-\tau)}\right)dsd\tau.$$

The following system of integro-differential equations is used to model TB dynamics with a variable latent period:

$$\frac{dS}{dt} = \Lambda - \sigma S\frac{I}{N} - \mu S + r_1 E + r_2 I, \tag{7}$$

$$E(t) = E_0(t) + \int_0^t \sigma S(s)\frac{I(s)}{N(s)}p(t-s)e^{-(\mu+r_1)(t-s)}ds, \tag{8}$$

$$\begin{aligned} I(t) = &\int_0^t\int_0^\tau \sigma S(s)\frac{I(s)}{N(s)}e^{-(\mu+r_1)(\tau-s)}\left(-\dot{p}(\tau-s)e^{-(\mu+r_2+d)(t-\tau)}\right)dsd\tau \\ &+ I_0 e^{-(\mu+r_2+d)t} + I_0(t), \end{aligned} \tag{9}$$

$$N = S + E + I,$$

where $\sigma = \beta c$ is the force of infection per infective; r_1 and r_2 are the per-capita treatment rates for the E class and I class, respectively; μ is the natural mortality rate; d is the per-capita death rate due to TB; $E_0(t)$ denotes those individuals in the E class at time $t = 0$ who are still in the latent class at time t; $I_0(t)$ denotes those individuals who are initially in the E class who have moved into class I and still alive at time t; and $I_0 e^{-(\mu+r_2+d)t}$ with $I_0 = I(0)$ represents those individuals who are infectious at time 0 who are still alive in the I class at time t. Mathematically, it is assumed that $E_0(t)$ and $I_0(t)$ have compact support.

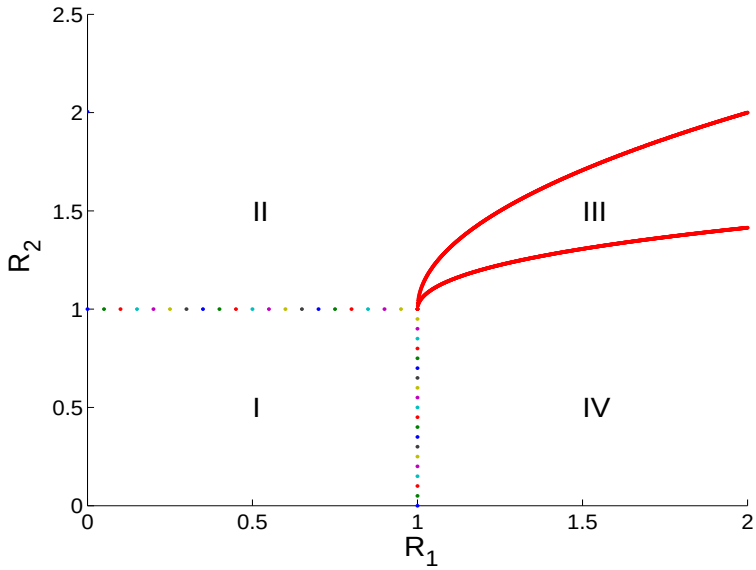


FIGURE 3. Bifurcation diagram for model (10–15) when $q = 0$. In region I, the disease-free equilibrium is globally asymptotically stable; in regions II and IV, one of strains disappears; and III represents the coexistence region.

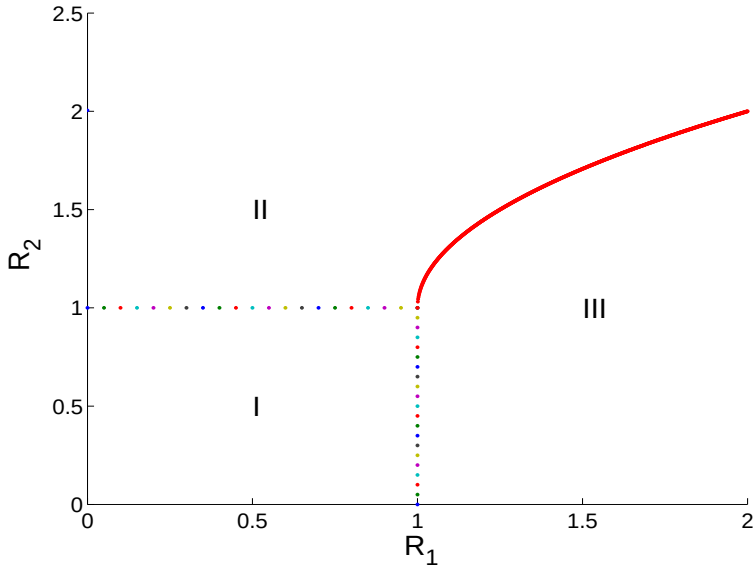


FIGURE 4. Bifurcation diagram for model (10–15) when $q > 0$. The coexistence of the two strains is impossible.

The introduction of an arbitrary distribution of latency period did not change the qualitative dynamics of TB; that is, a forward bifurcation diagram characterizes the dynamics of the last model (7–9).

4.3. **Multiple strains.** Incomplete treatment, wrong therapy, and co-infection with other diseases, for instance, HIV, may give rise to new resistant strains of TB (multiple drug resistant, or MDR strains). Form First Lady Eleanor Roosevelt was one of the victims of MDR TB [24, 62]. Models that include multiple strains of TB have been developed [7, 15, 17]. A recently published two-strain TB model include drug-sensitive and drug-resistant strains [15, 17]. Hence, two subclasses of latent and infectious individuals are required. The subscripts s and r stand for drug-sensitive and drug-resistant types. The model is given by the following set of equations:

$$\frac{dS}{dt} = \Lambda - \beta_s c S \frac{I_s}{N} - \beta_r c S \frac{I_r}{N} - \mu S, \tag{10}$$

$$\frac{dE_s}{dt} = \beta_s c S \frac{I_s}{N} - (\mu + k_s) E_s - r_{1s} E_s + p r_{2s} I_s - \beta_s c E_s \frac{I_r}{N}, \tag{11}$$

$$\frac{dI_s}{dt} = k_s E_s - (\mu + d_s) I_s - r_{2s} I_s, \tag{12}$$

$$\frac{dT}{dt} = r_{1s} E_s + (1 - p - q) r_{2s} - \beta_s c T \frac{I_s}{N} - \beta_r c T \frac{I_r}{N} - \mu T, \tag{13}$$

$$\frac{dE_r}{dt} = q r_{2s} I_s - (\mu + k_r) E_r + \beta_r c (S + E_s + T) \frac{I_r}{N}, \tag{14}$$

$$\frac{dI_r}{dt} = k_r E_r - (\mu + d_r) I_r, \tag{15}$$

$$N = S + E_s + I_s + T + E_r + I_r.$$

It can be seen from Equation (15) that the treatment rate for the I_r class is equal to zero, meaning that TB due to this strain is not treatable by current antibiotics. The proportion of treated infectious individuals who did not complete treatment is $p+q$. The proportion p modifies the rate at which they depart from the latent class; $q r_{2s} I_s$ gives the rate at which individuals develop resistant-TB due to their lack of compliance with TB treatment. The proportion of successfully treated individuals is $1 - p - q$. The dimensionless quantities

$$\mathcal{R}_1 = \left(\frac{\beta_s c + p r_{2s}}{\mu + d_s + r_{2s}} \right) \left(\frac{k_s}{\mu + k_s + d_s} \right)$$

and

$$\mathcal{R}_2 = \left(\frac{\beta_r c}{\mu + d_r} \right) \left(\frac{k_r}{\mu + k_r} \right)$$

give the basic reproductive number of strains j, where $j = 1, 2$. The asymptotic behavior of model (10–15) is determined by $\mathcal{R}_j$. Fig.s 3 and 4 show the bifurcation diagram for this two-strain model. These diagrams shows that naturally resistant and natural types can co-exist, albeit the region of coexistence is small (region IV in Fig. 3). Furthermore, in [15] it was shown that antibiotic-induced resistance results in the substantial expansion of the region of coexistence. In fact, regions IV and III become a single large region of coexistence. In other words, antibiotic-induced resistance guarantees the survival of resistant strains. A two-strain model that incorporates the effects of multiple drug resistance can also be found in [7].

4.4. **Multiple strains and variable latent period.** A model that considers both multiple strains and variable latent period is proposed by Feng [33]. Drug-resistant and drug-sensitive strains are modeled, but only the age of the infection with drug-sensitive strain is considered. They introduce a function $p(\theta)$ as the proportion of the sensitive-strain that are active at infection-age θ to distinguish active TB and inactive TB. The model does not make a difference between active and inactive TB for the drug-resistant strain because after acquiring drug-resistent TB, an individual dies quickly. Consequently for the drug-resistant strain they only count active TB. The total population is divided into three classes: susceptible ($S(t)$), infections with drug-sensitive strain ($I_s(t)$), and infections with drug-resistant strain ($I_r(t)$). Letting $i_s(\theta, t)$ be the infection-age density of infected individuals with the drug-sensitive strain at time t, the model framework takes the following system:

$$\frac{dS}{dt} = b(N)N - \left(\mu + \beta_1 c\frac{I_s^a}{N} + \beta_2 c\frac{I_r}{N}\right) S(t) + (1-r)r_2 I_s^a, \tag{16}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta}\right) i_s(\theta, t) + ((1 - r + qr)r_2 p(\theta) + \mu + d_1)\, i_s(\theta, t) = 0, \tag{17}$$

$$\frac{dI_r}{dt} = \beta_2 c S\frac{I_r}{N} - (\mu + d_2)I_r + qrr_2 I_s^a, \tag{18}$$

$$i_s(0, t) = \beta_1 c S\frac{I_s^a}{N}, \tag{19}$$

where $I_s^a(t) = \int_0^\infty p(\theta) i_s(\theta, t) d\theta$ is the total number of active TB of drug-sensitive strain; $I_s(t) = \int_0^\infty i_s(\theta, t) d\theta$ the total number of infected individuals with drug-sensitive TB (both latent TB and active TB are included); $b(N)$ the per-capita

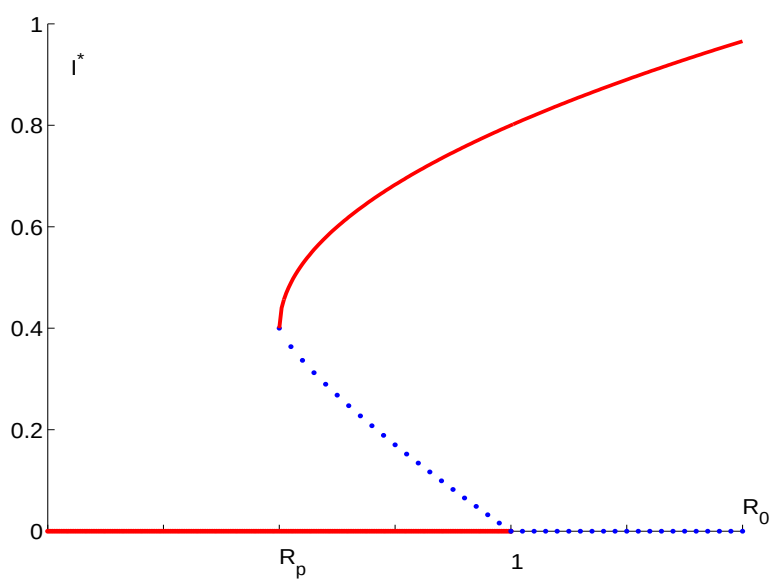


FIGURE 5. Backward bifurcation diagram when exogenous reinfection is included in model (20–23). When $\mathcal{R}_0 < \mathcal{R}_p$, the disease-free equilibrium is globally asymptotically stable. However, when $\mathcal{R}_p < \mathcal{R}_0 < 1$, there are two endemic equilibria. The upper ones are stable, and the lower ones are unstable.

birth rate into the population; $N(t) = S(t) + I_s(t) + I_r(t)$ the total population size; The disease-induced mortalities for drug-sensitive TB and drug-resistant TB are denoted by d_i respectively. It is assumed that a fraction r of the treated individuals with the drug-sensitive strain does not recover due to incomplete treatment, and the remaining fraction $1 - r$ is successfully treated and become susceptible again. It is also assumed that a fraction q of those who do not finish their treatment will generate drug-resistant TB and the remaining fraction $1 - q$ of them will keep as infectious.

Model 16–19) shows "that nonantibiotic-induced coexistence is possible but rare for naturally resistant strains, while coexistence is almost the rule for strains that result from the lack of compliance with antibiotic treatment by TB-infected individuals" [33].

4.5. **Exogenous reinfection.** Immediately after primary infection, an infected individual, on average, has a higher risk of progression (becoming an infectious case). Infected individuals who do not become infectious within a short time period may still develop active TB via exogenous or endogenous reinfection or both. To see the impact of exogenous reinfection on the dynamics of TB, we discuss the model (see [32]) given by the following set of equations:

$$\frac{dS}{dt} = \Lambda - \beta c S \frac{I}{N} - \mu S, \tag{20}$$

$$\frac{dE}{dt} = \beta c S \frac{I}{N} - p\beta c E \frac{I}{N} - (\mu + k)E + \sigma \beta c T \frac{I}{N}, \tag{21}$$

$$\frac{dI}{dt} = p\beta c E \frac{I}{N} + kE - (\mu + r + d)I, \tag{22}$$

$$\frac{dT}{dt} = rI - \sigma \beta c T \frac{I}{N} - \mu T \tag{23}$$

$$N = S + E + I + T.$$

The term $p\beta c E \frac{I}{N}$ models exogenous reinfection, that is, the potential reactivation of TB by continuous exposure of latently-infected individuals to those who have active infections.

The dynamics generated from model (20–23) are "surprising" as they show that $\mathcal{R}_0 = 1$ is not always the key threshold. We introduces a method that deals with bifurcation problem arising from this model as well as more general epidemic models in the following subsection.

4.6. **An approach to determine the direction of the bifurcation at $\mathcal{R}_0 = 1$.** For the most part, in epidemic models, there are two distinct bifurcations at $\mathcal{R}_0 = 1$ forward (supercritical) and backward (subcritical). A forward bifurcation happens when $\mathcal{R}_0$ crosses unity from below; a small positive asymptotically stable equilibrium appears and the disease-free equilibrium losses its stability. On the other hand, a backward bifurcation happens when $\mathcal{R}_0$ is less than unity; a small positive unstable equilibrium appears while the disease-free equilibrium and a larger positive equilibrium are locally asymptotically stable. Epidemiologically, a backward bifurcation "says" that it is not enough to only reduce the basic reproductive number to less than one to eliminate a disease and that when $\mathcal{R}_0$ crosses unity, hysteresis takes place. This phenomenon that probably was first found in epidemiological models by Huang et al. [45] in 1992 in a study of an HIV/AIDS model has now

become common. (Huang et al. identified it as a subcritical bifurcation.) Recent studies with models supporting backward bifurcations include those of Hadeler and Castillo-Chavez in a variable core group model [41]; Dushoff et al. [30] in models for fatal diseases; Feng et al. [32] in a TB model with exogenous reinfection; and Kribs-Zaleta and Velasco-Hernandez [50] in a model with vaccination. Most recently, van den Driessche and Watmough [27, 28] have shown the existence of this behavior in epidemic models with delay while Castillo-Chavez and Huang have done the same for models with age-structure [18].

Center manifold theory has been used to decide the local stability of a nonhyperbolic equilibrium (linearization matrix has at least one eigenvalue with zero real part) [14, 37, 86]. We shall describe a theory that not only can determine the local stability of the nonhyperbolic equilibrium but also t settles the question of the existence of another equilibrium (bifurcated from the nonhyperbolic equilibrium). This theory is based on the general center manifold theory. To describe it, consider a general system of ODEs with a parameter ϕ:

$$\frac{dx}{dt} = f(x, \phi), \quad f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \text{ and } f \in \mathbb{C}^2(\mathbb{R}^n \times \mathbb{R}). \tag{24}$$

Without loss of generality, it is assumed that 0 is an equilibrium for System (24) for all values of the parameter ϕ, that is

$$f(0, \phi) \equiv 0 \text{ for all } \phi. \tag{25}$$

THEOREM 4.1. *Assume*

A1: $A = D_x f(0,0) = \left(\frac{\partial f_i}{\partial x_j}(0,0)\right)$ *is the linearization matrix of System (24) around the equilibrium* 0 *with* ϕ *evaluated at* 0. *Zero is a simple eigenvalue of* A *and all other eigenvalues of* A *have negative real parts;*

A2: *Matrix* A *has a nonnegative right eigenvector* w *and a left eigenvector* v *corresponding to the zero eigenvalue.*

Let f_k *be the* k*th component of* f *and*

$$a = \sum_{k,i,j=1}^{n} v_k w_i w_j \frac{\partial^2 f_k}{\partial x_i \partial x_j}(0,0), \tag{26}$$

$$b = \sum_{k,i=1}^{n} v_k w_i \frac{\partial^2 f_k}{\partial x_i \partial \phi}(0,0). \tag{27}$$

The local dynamics of (24) around 0 *are totally determined by* a *and* b.

i. $a > 0$, $b > 0$. *When* $\phi < 0$ *with* $|\phi| \ll 1$, 0 *is locally asymptotically stable, and there exists a positive unstable equilibrium; when* $0 < \phi \ll 1$, 0 *is unstable and there exists a negative and locally asymptotically stable equilibrium;*

ii. $a < 0$, $b < 0$. *When* $\phi < 0$ *with* $|\phi| \ll 1$, 0 *is unstable; when* $0 < \phi \ll 1$, 0 *is locally asymptotically stable, and there exists a positive unstable equilibrium;*

iii. $a > 0$, $b < 0$. *When* $\phi < 0$ *with* $|\phi| \ll 1$, 0 *is unstable, and there exists a locally asymptotically stable negative equilibrium; when* $0 < \phi \ll 1$, 0 *is stable, and a positive unstable equilibrium appears;*

iv. $a < 0$, $b > 0$. *When* ϕ *changes from negative to positive,* 0 *changes its stability from stable to unstable. Correspondingly a negative unstable equilibrium becomes positive and locally asymptotically stable.*

The results of Theorem 4.1 are summarized in Table 3. Huang et al. [45] and Dushoff et al. [30] initially decomposed the center manifold and found out that the

sign of a can be used to determine the direction of the bifurcation. In doing so, they computed the Taylor expansion of $f(x,\phi)$ only on its state variables. Van den Driessche and Watmough [28] used a Taylor expansion on the state variables and the parameter ϕ in the analysis. Let us sketch our proof below.

Proof. Let $\mathcal{E}^c$ and $\mathcal{E}^s$ be the generalized eigenspaces of A for the zero eigenvalue and all other eigenvalues, respectively. It follows from the center manifold theory that the center manifold W^c is one dimensional, and $\mathbb{R}^n = \mathcal{E}^c \oplus \mathcal{E}^s$. We parameterize the center manifold by $c(t)$ and decompose it into $\mathcal{E}^c$ and $\mathcal{E}^s$, that is,

$$W^c = \{c(t)w + h(c,\phi) : v \cdot h(c,\phi) = 0, |c| \le c_0, c(0) = 0\}, \tag{28}$$

where $c(t) \in \mathcal{E}^c$ and $h(c,\phi) \in \mathcal{E}^s$. Because the center manifold is tangent to $\mathcal{E}^c$ at the origin, $h(c,\phi)$ is a higher order term ($h(c,\phi)$ has at least order 2). It also follows by the invariance of the center manifold under the flow that

$$\frac{d}{dt}(c(t)w + h(c,\phi)) = f(c(t)w + h(c,\phi),\phi). \tag{29}$$

Applying Taylor expansion to the right-hand side of (29) at $(0,0)$ and noticing that $h(c,\phi)$ is of higher order, we obtain that

$$\begin{aligned}
&f(c(t)w + h(c,\phi),\phi)\\
&= f(0,0) + D_x f(0,0)((c(t)w + h(c,\phi)) + D_\phi f(0,0)\phi\\
&+ \frac{1}{2}(I_n \otimes (cw + h(c,\phi))')(D^2_{xx} f(0,0))(cw + h(c,\phi)\\
&+ \phi(D^2_{x\phi} f(0,0))(cw + h(c,\phi))\\
&+ \frac{1}{2}\phi^2(D^2_{\phi\phi} f(0,0)) + \text{ higher order term },
\end{aligned}$$

where D^2_{xx} is the Hessian matrix; I_n is the identity matrix of order n; $\otimes$ is the Kronecker product. Using

$$f(0,0) = D_x f(0,0)c(t)w = D_\phi f(0,0) = D^2_{\phi\phi} f(0,0) = 0$$

and the fact that $ch(c,\phi)$ is of higher order, we simplify the above expansion for f as (higher order terms are dropped)

$$f = (D_x f)h(c,\phi) + \frac{c^2}{2}(I_n \otimes w')(D^2_{xx} f)w + c\phi(D^2_{x\phi} f)w. \tag{30}$$

Multiplying both sides of (29) by v and using the fact that $v \cdot h = 0$ and $vD_x f(0,0) = 0$, we finally obtain the following equation for $c(t)$:

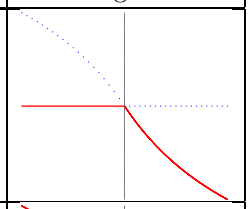
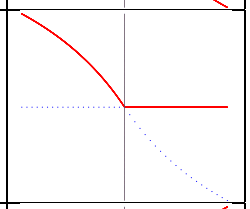
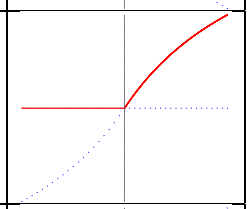
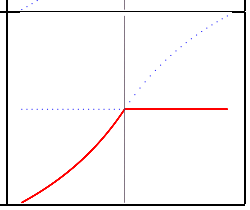
$$\begin{aligned}
\frac{dc}{dt} &= \frac{c^2}{2} v(I_n \otimes w')D^2_{xx} fw + c\phi v D^2_{x\phi} fw\\
&= \frac{c^2}{2}\sum_{k,i,j=1}^{n} v_k w_i w_j \frac{\partial^2 f_k}{\partial x_i \partial x_j} + \sum_{k,i=1}^{n} v_k w_i \frac{\partial^2 f_k}{\partial x_i \partial \phi} c\phi\\
&= \frac{a}{2}c^2 + b\phi c.
\end{aligned}$$

Namely,

$$\frac{dc}{dt} = \frac{a}{2}c^2 + b\phi c \tag{31}$$

Obviously, at $\phi = 0$ a transcritical bifurcation takes place in Equation (31). The bifurcation diagram is shown in Table 3. □

TABLE 3. Summary of Theorem 4.1. In the bifurcation diagrams, the vertical axis represents equilibrium points x^*, and the horizontal axis is the parameter ϕ. Solid lines and dashed lines symbolize stable (S) and unstable (U), respectively.

a and b	stability of 0	stability and sign of x^*	diagram
$a > 0,\ b > 0$	$\phi < 0$, S $\phi > 0$, U	$\phi < 0,\ x^* > 0$, U $\phi > 0,\ x^* < 0$, S	
$a < 0,\ b < 0$	$\phi < 0$, U $\phi > 0$, S	$\phi < 0,\ x^* > 0$, S $\phi > 0,\ x^* < 0$, U	
$a < 0,\ b > 0$	$\phi < 0$, S $\phi > 0$, U	$\phi < 0,\ x^* < 0$, U $\phi > 0,\ x^* > 0$, S	
$a > 0,\ b < 0$	$\phi < 0$, U $\phi > 0$, S	$\phi < 0,\ x^* < 0$, S $\phi > 0,\ x^* > 0$, U	

REMARK 1. *The requirement that w is nonnegative in Theorem 4.1 is not necessary. When some components in w are negative, we still can apply this theorem, but one has to compare w with actual the equilibrium because the general parameterization of the center manifold before the coordinate change is*

$$W^c = \{x_0 + c(t)w + h(c,\phi) : v \cdot h(c,\phi) = 0, |c| \le c_0, c(0) = 0\}$$

provided that x_0 is a nonnegative equilibrium of interest (Usually x_0 is the disease-free equilibrium). Hence, $x_0 - \frac{2b\phi}{a} > 0$ requires that $w(j) > 0$ whenever $x_0(j) = 0$. If $x_0(j) > 0$, then $w(j)$ need not be positive.

COROLLARY 4.1. *When $a > 0$ and $b > 0$, the bifurcation at $\phi = 0$ is subcritical (backward).*

Applying Corollary 4.1 to model (20–23), we give a rigorous proof that model (20–23) undergoes a backward bifurcation. We do it in a rather simple case $\sigma = 1$ though when $\sigma \neq 1$ the arguments are identical. Let $\phi = \beta c$ be the bifurcation

parameter. Introducing $x_1 = S + T$, $x_2 = E$, $x_3 = I$, the System (20–23) becomes

$$\frac{dx_1}{dt} = \Lambda - \phi\frac{x_1x_3}{x_1+x_2+x_3} - \mu x_1 := f_1, \tag{32}$$

$$\frac{dx_2}{dt} = \phi\frac{x_1x_3}{x_1+x_2+x_3} - p\phi\frac{x_2x_3}{x_1+x_2+x_3} - (\mu+k)x_2 := f_2, \tag{33}$$

$$\frac{dx_3}{dt} = p\phi\frac{x_2x_3}{x_1+x_2+x_3} + kx_2 - (\mu+r+d)x_3 := f_3, \tag{34}$$

with $\mathcal{R}_0 = 1$ corresponding to $\phi = \phi^* = \frac{(k+\mu)(\mu+r+d)}{k}$. The disease-free equilibrium is $[x_1^* = \frac{\Lambda}{\mu}, x_2^* = 0, x_3^* = 0]$. The linearization matrix of system (32–34) around the disease-free-equilibrium when $\phi = \phi^*$ is

$$D_x f = \begin{bmatrix} -\mu & 0 & -\phi^* \\ 0 & -(k+\mu) & \phi^* \\ 0 & k & -(\mu+r+d) \end{bmatrix}.$$

It is clear that 0 is a simple eigenvalue of $D_x f$. A right eigenvector associated with 0 eigenvalue is $w = [-\frac{\mu+k}{\mu k}, \frac{1}{k}, \frac{1}{\mu+r+d}]'$, and the left eigenvector v satisfying $v \cdot w = 1$ is $v = [0, \frac{k(\mu+r+d)}{(k+\mu)+(\mu+r+d)}, \frac{(k+\mu)(\mu+r+d)}{(k+\mu)+(\mu+r+d)}]$. Algebraic calculations show that

$$\frac{\partial^2 f_2}{\partial x_2 \partial x_3} = -(1+p)\frac{\phi^*}{x_1^*}, \quad \frac{\partial^2 f_2}{\partial x_3^2} = -2\frac{\phi^*}{x_1^*},$$

$$\frac{\partial^2 f_3}{\partial x_2 \partial x_3} = p\frac{\phi^*}{x_1^*}, \quad \frac{\partial^2 f_2}{\partial x_1 \partial \phi} = 1.$$

The rest of the second derivatives appearing in the formula for a in (26) and b in (27) are all zero. Hence,

$$a = \frac{\phi^*}{x_1^*}\frac{\mu}{k((k+\mu)+(\mu+r+d))}\left(p - \frac{k}{\mu}\left(1 + \frac{k}{\mu+r+d}\right)\right),$$

$$b = v_2 w_3 = \frac{k}{(k+\mu)+(\mu+r+d)} > 0.$$

We collect these results in the theorem below:

THEOREM 4.2. *If $p > p_0 = \frac{k}{\mu}\left(1 + \frac{k}{\mu+r+d}\right)$, the direction of the bifurcation of system (32–34) (or system (20–23)) at $\mathcal{R}_0 = 1$ is backward.*

Model (20–23) exhibits totally different behavior as it supports multiple stable-steady states via a backward bifurcation (subcritical bifurcation). The results are counter-intuitive. that is, when the basic reproductive number $\mathcal{R}_0 = \frac{\beta c}{\mu+r+d}\frac{k}{\mu+k} < 1$, the disease-free equilibrium is only *locally* stable. The system supports multiple endemic equilibria (see Fig. 5). The results of this model reveal how hard it may be to eliminate TB if exogenous reinfection is important. Vynnycky and Fine have also incorporated reinfection in their non-autonomous and linear PDE model (see section 8).

4.7. **Generalized households.** Both close and casual contacts may give rise to infections. Individuals having lengthy exposure may have a substantially higher risk of infection than those having only casual contacts with active-TB cases. For thousands of years, no one knew that TB was an infectious disease. In the nineteenth century, doctors in western Europe and the United States thought that TB was hereditary because it often ran in families. After Doctor Koch identified

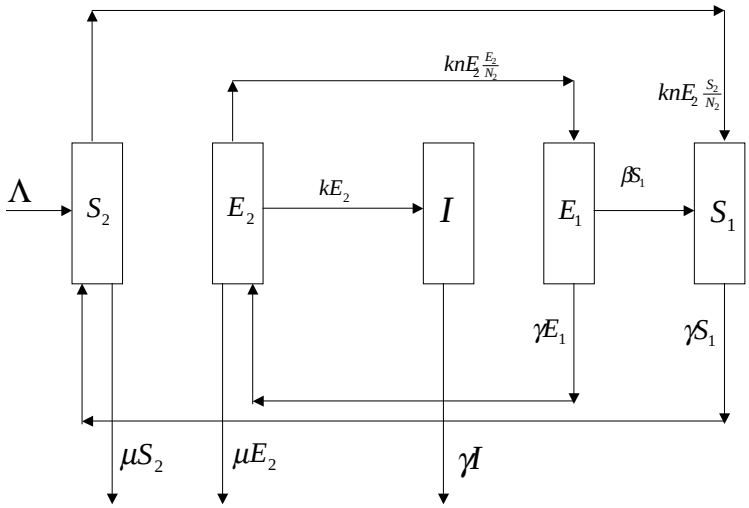


FIGURE 6. Flow chart of the basic cluster model.

the pathogen responsible for TB, it became clear that prolonged contacts with an active-TB case might generate new infections. In fact, now it is known that most new infections are due to close contacts. The case of a teacher-librarian with active TB who infected the children in her classroom but not the children who visited the library [61] supports the incorporation of differences between casual and close contacts in the investigation of TB dynamics at the population level. The report of the Asian tourist who infected at least 6 airline passengers while travelling to and from- and within-the United States further supports this view [49, 46]. In fact, in 1998, WHO declared that flights more than 8 hours may expose passengers and crew to TB infections [59]. Hence, distinguishing the type of contacts in the transmission of TB is important. Aparicio and colleagues proposed a new dynamical model that incorporated close and casual contacts [2]. They focused on the active-TB cases within their social networks (family members, officemates, classmates, any persons who have prolong contacts with an index case). This "new" epidemiological unit, called the generalized household or the epidemiologically active cluster, was used to study the transmission dynamics of TB. We describe the approach next. The population is partitioned into TB-active clusters and TB-inactive clusters. Once one member of an inactive cluster develops active TB, the whole cluster becomes an epidemiologically active-TB cluster. Specific contacts within each epidemiologically active cluster are not modeled, it is assumed that the risk of infection within an epidemiological active cluster just depends on the "life" of the cluster, that is, on the average length of TB's infection period distribution. The flow chart for the basic cluster model is shown in Fig. 6. Letting $N_2 = S_2 + E_2$, the model equations

are

$$\frac{dS_1}{dt} = -(\beta+\gamma)S_1 + \frac{S_2}{N_2}nkE_2, \tag{35}$$

$$\frac{dE_1}{dt} = \beta S_1 - \gamma E_1 + \frac{E_2}{N_2}nkE_2, \tag{36}$$

$$\frac{dI}{dt} = kE_2 - \gamma I, \tag{37}$$

$$\frac{dS_2}{dt} = \Lambda - \mu S_2 + \gamma S_1 - \frac{S_2}{N_2}nkE_2, \tag{38}$$

$$\frac{dE_2}{dt} = \gamma E_1 - (\mu+k)E_2 - \frac{E_2}{N_2}nkE_2. \tag{39}$$

The basic reproductive number for this model is

$$\mathcal{R}_0 = \frac{\beta n}{(\beta+\gamma)}\frac{k}{(\mu+k)}.$$

It can be seen that $\mathcal{R}_0$ depends nonlinearly on the parameter β (risk of infection on an epidemiologically active cluster of size n) and linearly on the average generalized household size, n.

Song et al. [71] further developed the basic cluster model (35–39) via the incorporation of casual contacts. An extended version of the cluster model for the TB transmission that includes the close and casual contacts is given by the following nonlinear system:

$$\frac{dS_1}{dt} = -(p\beta(n)+\gamma)S_1 + \frac{S_2}{N_2}nkE_2 - (1-p)\beta^*\frac{I}{N-n}S_1, \tag{40}$$

$$\frac{dE_1}{dt} = p\beta(n)S_1 - \gamma E_1 + \frac{E_2}{N_2}nkE_2 + (1-p)\beta^*\frac{I}{N-n}S_1, \tag{41}$$

$$\frac{dI}{dt} = kE_2 - \gamma I, \tag{42}$$

$$\frac{dS_2}{dt} = \Lambda - \mu S_2 + \gamma S_1 - \frac{S_2}{N_2}nkE_2 - (1-p)\beta^*\frac{I}{N-n}S_2, \tag{43}$$

$$\frac{dE_2}{dt} = \gamma E_1 - (\mu+k)E_2 - \frac{E_2}{N_2}nkE_2 + (1-p)\beta^*\frac{I}{N-n}S_2. \tag{44}$$

Here β is assumed to depend on the average cluster size n; p denotes the average fraction of time spent by the source case within his/her generalized household; and $1-p$, the average fraction of time spent by this source-case outside the cluster. The rate of infection within clusters becomes $p\beta(n)S_1$, while the rate of infection outside is $(1-p)\beta^*\frac{I}{N-n}(S_1+S_2)$, where N is the total population size, and $N-n$ represents the average total number of individuals outside the cluster. Hence, $(1-p)\beta^*\frac{I}{N-n}S_1$ gives the number of new infections per unit time in the N_1 population, that is, the incidence from S_1 to E_1. The term $(1-p)\beta^*\frac{I}{N-n}S_2$ gives the incidence from S_2 to E_2. There are no new cases of active TB within each epidemiologically active cluster, and consequently, the infection rate is $p\beta(n)S_1$. The generation of new cases from casual contacts is modeled in the typical way. When $p = 1$ and $\beta(n)$ is a constant, the extended cluster model (40–40) becomes the basic cluster model (35–39). The basic reproductive number associated with the generalized cluster

model is

$$\mathcal{R}_0(n) = \left(\frac{p\beta(n)n}{p\beta(n)+\gamma} + (1-p)\frac{\beta^*}{\gamma}\frac{K}{K-n} \right) \frac{k}{(\mu+k)}, \tag{45}$$

where $K = \frac{\Lambda}{\mu}$ is the asymptotic carrying capacity of the total population.

Two specific forms of $\beta(n)$ are discussed below:

Case 1. In the first case, we consider that $\beta(n)$ is a piecewise-defined continuous function that is,

$$\beta(n) = \begin{cases} \beta_0 & \text{for } n \le n_L, \\ \frac{\beta_1}{n} & \text{for } n_L < n < n_M, \end{cases}$$

where n_L is the critical cluster size after which $\beta(n)$ begins to decrease, and n_M is an upper bound for cluster size. Consequently,

$$\mathcal{R}_0(n) = \begin{cases} (\frac{p\beta_0 n}{p\beta_0+\gamma} + (1-p)\frac{\beta^*}{\gamma}\frac{K}{K-n})\frac{k}{(\mu+k)} & \text{for } n \le n_L, \\ (\frac{p\beta_1}{p\frac{\beta_1}{n}+\gamma} + (1-p)\frac{\beta^*}{\gamma}\frac{K}{K-n})\frac{k}{(\mu+k)} & \text{for } n \ge n_L. \end{cases}$$

Since the maximum cluster size n_M is significantly smaller than the carrying capacity K, then $\frac{K}{K-n} \approx 1$, whenever $n < n_L$ $\mathcal{R}_0$ increases linearly with cluster size, and $Q_0 \approx \frac{p\beta_0}{p\beta_0+\gamma}n + (1-p)\frac{\beta^*}{\gamma}$. If $n > n_L$, $Q_0 \approx \frac{p\beta_1 n}{p\beta_1+n\gamma} + (1-p)\frac{\beta^*}{\gamma}\frac{K}{K-n}$ which levels off at the value $p\frac{\beta_1}{\gamma} + (1-p))\frac{\beta^*}{\gamma}$. Hence, an increase in n translates an increase in TB transmission but the increase is non-linear. Initially, this increase is almost linear but as n becomes larger the rate of increase decreases, because the time spent by an infectious individual per contact cancels out increases in cluster size. Hence, $\mathcal{R}_0(n)$ is bounded by a constant value, and this bound limits the size of TB prevalence.

Case 2. The second case assumes $\beta(n)$ to be inversely proportional to n (that is, $\beta(n) = \frac{\beta_1}{n}$). It can be seen from 45 that $\mathcal{R}_0(n)$ is the sum of contributions from both the within and out of cluster new secondary cases of infection. The ratio $E(n)$ of within to between cluster contributions is given by

$$E(n) = \frac{\gamma\beta_1 p}{K\beta^*(1-p)}\frac{n(K-n)}{(p\beta_1+\gamma n)}.$$

Function $E(n)$ increases with n and reaches its maximum value at $n^* = \frac{K}{1+\sqrt{1+\frac{\gamma K}{p\beta_1}}}$. Hence,$n^*$ here defines the *optimal cluster size*; that is, the value maximizes within cluster transmission.

Singular perturbation theory and multiple time scales techniques are used to study the global dynamics of the cluster models [71]. Since the average infectious period (about 3–4 months) is much shorter than the average latent period, which has the same order as the host population, two different time scales can be identified. Time is measured using the average latency period $1/k$ as the unit of time (that is, $\tau = kt$). Variables S_2 and E_2 are re-scaled by Ω, the total asymptotic population size ($\Omega = \frac{\Lambda}{\mu}$); N_1 could be re-scaled by Ω instead, it is re-scaled by the balance factor $\frac{k}{\beta+\gamma}\Omega$. The rescaling formulae are $x_1 = \frac{S_2}{\Omega}$, $x_2 = \frac{E_2}{\Omega}$, $y_1 = \frac{\beta+\gamma}{k}\frac{S_1}{\Omega}$, $y_2 = \frac{\beta+\gamma}{k}\frac{E_1}{\Omega}$, and $y_3 = \frac{\beta+\gamma}{k}\frac{I}{\Omega}$. These new re-scaled variables and parameters are non-dimensional.

The re-scaled model equations are given by

$$\frac{dx_1}{d\tau} = B(1-x_1) + (1-m)y_1 - n\frac{x_1x_2}{x_1+x_2}, \tag{46}$$

$$\frac{dx_2}{d\tau} = (1-m)y_2 - (1+B)x_2 - n\frac{x_2^2}{x_1+x_2}, \tag{47}$$

$$\epsilon\frac{dy_1}{d\tau} = -y_1 + n\frac{x_1x_2}{x_1+x_2}, \tag{48}$$

$$\epsilon\frac{dy_2}{d\tau} = my_1 - (1-m)y_2 + n\frac{x_2^2}{x_1+x_2}, \tag{49}$$

$$\epsilon\frac{dy_3}{d\tau} = x_2 - (1-m)y_3, \tag{50}$$

where $\epsilon = \frac{k}{\beta+\gamma}$, $m = \frac{\beta}{\beta+\gamma} < 1$ and $B = \frac{\mu}{k}$. The terms in right-hand side of system (46–47) all have the same order of magnitude whenever $\epsilon \ll 1$. Therefore, y_1, y_2, and y_3 are fast variables and x_1 and x_2 are slow variables. Hoppensteadt's early theorem [44] help us to show the global stability of the endemic equilibrium when ϵ is small. A Liapunov function and a Dulac function are selected to establish the global stability of the disease-free equilibrium. Hence, a global forward bifurcation (see Fig. 7) characterizes the dynamics of the cluster models.

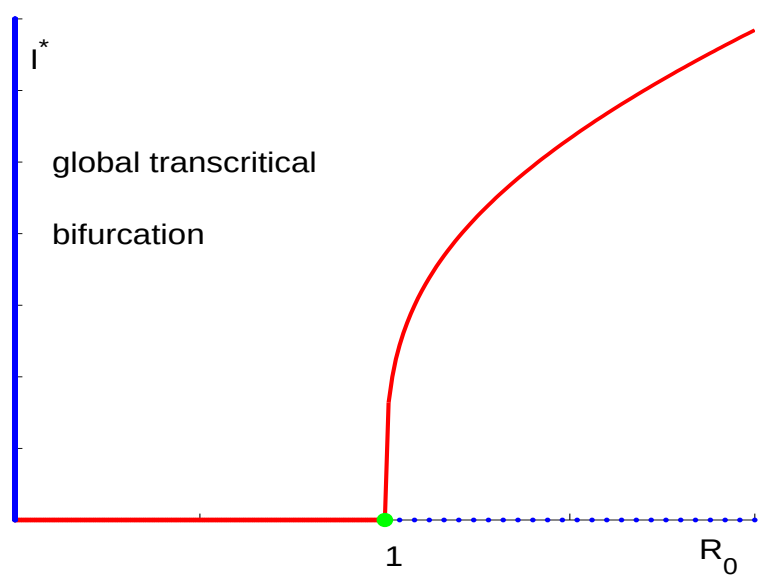


FIGURE 7. Global forward bifurcation diagram for the TB models.

5. **Models with density dependent demography.** Demography plays an important role in the transmission dynamics of TB since the average rate of progression from infected (non-infectious) to active (infectious) TB is very slow. In fact, the average latent period has the same order of magnitude as the life-span of the host population. Two distinct demographic scenarios are studied. In the first, exponential growth is observed over a long time scale, and in the second exponential growth is observed over a short time scale (quasi-exponential growth).

Consequently, three different recruitment rates are used in the study of dynamical models for TB transmission with demography: constant recruitment rate (Λ) [15], linear recruitment rate (rN) [74], and logistic recruitment rate ($rN(1-N/K)$) [74]. The general model is given by

$$\frac{dS}{dt} = B(N) - \beta c S\frac{I}{N} - \mu S, \tag{51}$$

$$\frac{dE}{dt} = \beta c S\frac{I}{N} - (\mu + k + r_1)E + \beta' c T\frac{I}{N}, \tag{52}$$

$$\frac{dI}{dt} = kE - (\mu + d + r_2)I, \tag{53}$$

$$\frac{dT}{dt} = r_1 E + r_2 I - \beta' c T\frac{I}{N} - \mu T, \tag{54}$$

$$N = S + E + I + T,$$

where the recruitment rate $B(N)$ includes the three demographies described above, namely, rN, $rN(1-N/K)$, and Λ. The basic epidemiological reproductive number is given by

$$\mathcal{R}_0 = \left(\frac{k}{\mu + r_1 + k}\right)\left(\frac{\beta c}{\mu + r_2 + d}\right) \tag{55}$$

However, this non-dimensional number is not enough to characterize the dynamics of model (51-54).

5.1. **Linear recruitment rate.** Currently, most deaths caused by TB represent a small proportion of the deaths for most populations, in other words, d is often insignificant. Therefore, a linear recruitment rate $B(N) = rN$ with reasonable r values is likely to support exponential growth on a TB-infected population. The use of a logistic recruitment rate $B(N) = rN(1 - N/K)$ to model the demography in general is also likely to result in logistic growth for the total population N in the presence of TB. To simplify our analysis, we further assume that the infected and reinfected proportions are equal $\beta' = \beta$. The use of the variables, N, E and I, is enough. Hence, model (51–54) reduces to:

$$\frac{dN}{dt} = B(N) - \mu N - dI, \tag{56}$$

$$\frac{dE}{dt} = \beta c(N - E - I)\frac{I}{N} - (\mu + k + r_1)E, \tag{57}$$

$$\frac{dI}{dt} = kE - (\mu + d + r_2)I. \tag{58}$$

We shall consistently use the following compressed notation $m_r = r + r_2 + d, n_r = r + r_1 + k, m_\mu = \mu + r_2 + d, n_\mu = \mu + r_1 + k$, and $\sigma = \beta c$ to simplify the discussions.

First, we study the dynamics of model (56–58) with $B(N) = rN$. That is, it is assumed that the total population exhibits exponential growth in the absence of TB (the net growth rate of the population, in the absence of the disease, is $r - \mu$). Total population size increases exponentially if $r > \mu$, and remains constant if $r = \mu$. The case where $r < \mu$ is trivial. Hence we assume that $r \geq \mu$. In the presence of TB, the total population may (theoretically) decrease exponentially even when $r > \mu$ if d is large enough; that is, technically, a fatal disease can impact population growth (see also May and Anderson, [57]; Busenberg and Hadeler [11]). Realistic examples of situations where a disease has or is likely to to have an impact on the demographic growth can be found in the work on myxomatosis [51] and on

HIV [1, 56]. Three non-dimensional threshold parameters $\mathcal{R}_0$, $\mathcal{R}_1$, and $\mathcal{R}_2$ provide a full characterization of the possible dynamical regimes of system (56–58) under the various demographic regimes. The basic reproductive number

$$\mathcal{R}_0 = \left(\frac{\sigma}{\mu + r_2 + d}\right)\left(\frac{k}{\mu + r_1 + k}\right), \tag{59}$$

gives the average number of secondary cases produced by a typical infectious individual during his/her entire life in a population of mostly susceptibles. It is implied from $\mathcal{R}_0 < 1$ that the infected populations goes to zero, while $\mathcal{R}_0 > 1$ implies that the infected populations grows (initially) exponentially (together with the total population N). In this last case, there are two possibilities: N grows faster than I, or N does not grow faster than I. In the first case, the fraction $u = \frac{I}{N}$ approaches zero as time increases, and the additional threshold parameter

$$\mathcal{R}_1 = \left(\frac{\sigma}{r + r_2 + d}\right)\left(\frac{k}{r + r_1 + k}\right) \tag{60}$$

plays a role; that is, $\mathcal{R}_1$ discriminates between the last two possibilities. If $\mathcal{R}_1 < 1$, then $\lim_{t\to\infty} u = 0$, while $\mathcal{R}_1 > 1$ implies that $\lim_{t\to\infty} u > u^* > 0$. Our assumption $r > \mu$ implies that $\mathcal{R}_0 > \mathcal{R}_1$ is true. If the infectious (I) population changes faster than the total population (N) then (a fatal) disease can drive the population to extinction (even when $\mathcal{R}_1 > 1$). The threshold parameter that decides this last situation is given by

$$\mathcal{R}_2 = \frac{r - \mu}{du^*}, \tag{61}$$

where u^* is a positive constant (independent of μ (see (67)); that is, $\mathcal{R}_2$ determines whether or not the total population size grows exponentially. In fact, population size would decrease exponentially (from TB) only if $\mathcal{R}_2 < 1$. System (56–58) is homogeneous of degree one, and hence, it can support exponential solutions [39, 40]. However, we are interested in the global dynamics rather than the linear (local) stability of the homogeneous systems. Global analysis requires the rewriting of system (56–58) using the projections $u = \frac{I}{N}$, $v = \frac{E}{N}$. The equations for u, v are given by the following quadratic system:

$$\frac{du}{dt} = -m_r u + kv + du^2, \tag{62}$$

$$\frac{dv}{dt} = \sigma u - n_r v + (d - \sigma)uv - \sigma u^2. \tag{63}$$

Note that both u and v are independent of N and μ. The subset

$$\Omega = \{(u, v) \in R_2^+ | u + v \leq 1\}$$

is positively invariant. To further simplify the quadratic system (62–63), we introduce the new variables x and y and re-scale time t. Specifically, we let

$$x = \frac{d}{m_r + n_r}u, \quad y = \frac{kd}{(m_r + n_r)^2}\left(\frac{n_r}{k}u + v\right), \qquad \tau = (m_r + n_r)t. \tag{64}$$

The re-scaled system is

$$\frac{dx}{d\tau} = -x + y + x^2, \tag{65}$$

$$\frac{dy}{d\tau} = x(a_1 + a_2 y + a_3 x), \tag{66}$$

where

$$a_1 = \frac{m_r n_r(\mathcal{R}_1 - 1)}{(m_r + n_r)^2}, \quad a_2 = \frac{d-\sigma}{d}, \quad a_3 = \sigma\frac{n_r - k}{d(m_r + n_r)}.$$

In the new system, Ω becomes

$$\Omega_1 = \left\{(x,y) \in R_2^+ | \frac{n_r}{m_r + n_r}x \le y \le \frac{dk}{(m_r + n_r)^2} + \frac{n_r - k}{m_r + n_r}x\right\}$$

which is positively invariant under the flow of system (65–66). Transformation 64 not only reduces the number of parameters but, more importantly, it fixes the horizontal isocline and decomposes the vertical isocline into a degenerate quadratic curve. Under the standard classification of Ye et al. [87], system (65–66) is *a quadratic system of the second type*. The following two theorems characterize the dynamics of system (65–66), and hence of system (62–63).

THEOREM 5.1. *For system (65–66) with $r > \mu$, the trivial equilibrium $(0,0)$ is globally asymptotically stable if $\mathcal{R}_1 \le 1$. Furthermore there exists a unique positive equilibrium that is globally asymptotically stable if $\mathcal{R}_1 > 1$.*

The standard classification of planar quadratic differential systems rules out the existence of closed orbits or limit cycles. (Other approaches can be used to draw the same conclusion, for example, see Busenberg and van den Driessche [12]; Lin and Hethcote [53]. The full structure of system (56–58) is provided in Theorem 5.2 below:

THEOREM 5.2. *Consider system (56–58) and assume that $r > \mu$:*

i. *If $\mathcal{R}_0 < 1$, then $(\infty, 0, 0)$ is globally asymptotically stable.*
ii. *If $\mathcal{R}_1 < 1 < \mathcal{R}_0$, then (∞,∞,∞) is globally asymptotically stable and $\lim_{t\to\infty}\frac{I}{N} = 0$, $\lim_{t\to\infty}\frac{E}{N} = 0$.*
iii. *If $1 < \mathcal{R}_1 < \mathcal{R}_0$, then*
   (a). *$(0,0,0)$ is globally asymptotically stable and $\lim_{t\to\infty}\frac{I}{N} = u^*$, $\lim_{t\to\infty}\frac{E}{N} = v^*$ when $\mathcal{R}_2 < 1$,*
   (b). *(∞,∞,∞) is globally asymptotically stable and $\lim_{t\to\infty}\frac{I}{N} = u^*$, $\lim_{t\to\infty}\frac{E}{N} = v^*$ when $\mathcal{R}_2 > 1$, where*

$$u^* = \frac{-[d(m_r + n_r) - \sigma(m_r + k)] + \sqrt{\delta}}{2d(\sigma - d)(k\sigma - m_r n_r)}, \tag{67}$$

$$v^* = \frac{m_r(a_2 + a_3 + \Delta^{1/2}) - 2a_2 d u^{*2}}{2a_2 k},$$

$$\delta = [d(m_r + n_r) - \sigma(m_r + k)]^2 + 4d(\sigma - d)(k\sigma - m_r n_r),$$

$$\Delta = (a_2 + a_3)^2 + 4a_1 a_2 > 0.$$

Hence, whenever $\mathcal{R}_0 < 1$ the disease dies out while the total population increases exponentially. Although the disease spreads (that is, the number of infected grows in total number) when $\mathcal{R}_1 < 1 < \mathcal{R}_0$, the proportions $\frac{I}{N}$ and $\frac{E}{N}$ approach zero. One can observe from the case $1 < \mathcal{R}_1 < \mathcal{R}_0$ that disease-induced mortality can lead to the extinction of a population that would otherwise increase exponentially (a fatal disease can indeed regulate a population). Note that $\mathcal{R}_2$ is positive since u^* is positive and independent of μ. We also have established that when $r < \mu$, $(0,0,0)$

is globally asymptotically stable even though $\lim_{t\to\infty} \frac{I}{N} = u^*$ and $\lim_{t\to\infty} \frac{E}{N} = v^*$ when $\mathcal{R}_1 > 1$, and $\mathcal{R}_1 < 1$ implies that $\lim_{t\to\infty} \frac{I}{N} = 0$, $\lim_{t\to\infty} \frac{E}{N} = 0$. Note that $\mathcal{R}_1 < \mathcal{R}_0$ whenever $r > \mu$. Theorem 5.2 provides a complete characterization of the dynamic structure of model (56–58). The global dynamics are shown in Fig. 8. Here, we provide the proof that the disease-free equilibrium is a global attractor.

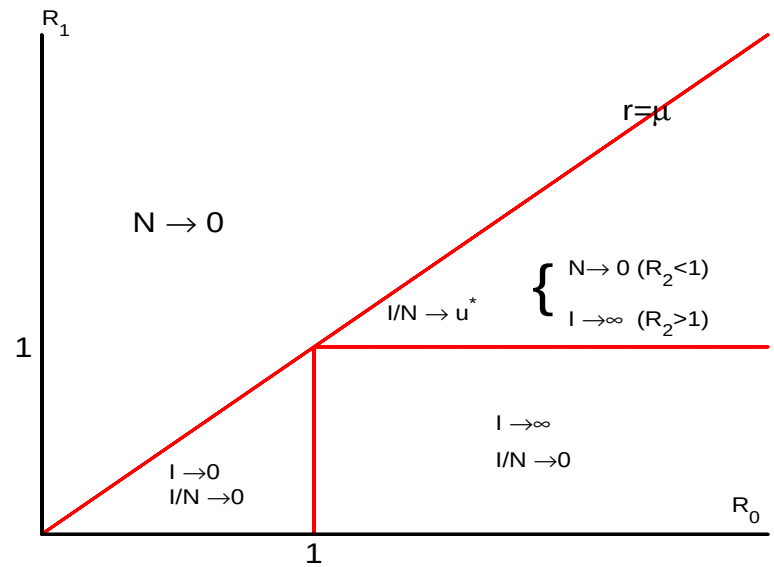


FIGURE 8. Bifurcation diagram for linear recruitment rate.

Proof. The disease-free equilibrium is $(\frac{r-\mu}{r}, 0, 0)$. It is straightforward to show that the endemic equilibrium is unique whenever $\mathcal{R}_0 > 1$ and $\mathcal{R}_2^* > 1$ and the disease-free equilibrium is locally stable whenever $\mathcal{R}_0 \leq 1$. Here, we only need to establish the global stability of the disease-free equilibrium under the assumption $\mathcal{R}_0 \leq 1$. Let $f(t) = \gamma E(t) + 2\sigma I(t)$, where $\gamma = \sqrt{(m_\mu - n_\mu)^2 + 4k\sigma} + m_\mu - n_\mu$. It suffices to show $\lim_{t\to\infty} f(t) = 0$.

$$\begin{aligned}
\frac{df(t)}{dt} &= \gamma \frac{dE(t)}{dt} + 2\sigma \frac{dI(t)}{dt} \\
&\leq \gamma(\sigma I(t) - n_\mu E(t)) + 2\sigma(kE(t) - m_\mu I(t)) \\
&= (2\sigma k - \gamma n_\mu)E(t) + (\gamma\sigma - 2\sigma m_\mu)I(t) \\
&= (-n_\mu + \frac{2\sigma k}{\gamma})\gamma E(t) + (\frac{\gamma}{2} - m_\mu)2\sigma I(t) \\
&= (\gamma E(t) + 2\sigma I(t))\left(\frac{\sqrt{(m_\mu - n_\mu)^2 + 4\sigma k} - (m_\mu + n_\mu)}{2}\right) \\
&= -\frac{(1 - \mathcal{R}_0)}{\sqrt{(m_\mu - n_\mu)^2 + 4\sigma k} + m_\mu + n_\mu} f(t).
\end{aligned}$$

This actually produces a differential inequality on the function $f(t)$; that is,

$$\frac{df(t)}{dt} < -\frac{(1-\mathcal{R}_0)}{\sqrt{(m_\mu - n_\mu)^2 + 4\sigma k} + m_\mu + n_\mu} f(t). \tag{68}$$

It follows that $\lim_{t\to\infty} f(t) = 0$ from $\frac{(1-\mathcal{R}_0)}{\sqrt{(m_\mu-n_\mu)^2+4\sigma k}+m_\mu+n_\mu} > 0$ and $\mathcal{R}_0 < 1$. If $\mathcal{R}_0 = 1$, $f(t)$ no longer decays exponentially, but it still vanishes as time goes to infinite. We estimate the derivative of $f(t)$ again as follows:

$$\begin{aligned}\frac{df(t)}{dt} &= -\sigma\gamma(E+I)\frac{I}{N}\\ &\le -\sigma\gamma(E+I)I\\ &\le -\gamma_1 I f(t), \quad \text{where} \quad \gamma_1 = \frac{min\{2\sigma,\gamma\}}{2} > 0.\end{aligned}$$

Hence, $f(t)$ is a decreasing function and

$$f(t) \le f(0)e^{-\gamma_1 \int_0^t I(t)dt}. \tag{69}$$

If $\lim inf_{t\to\infty} I(t) > 0$, then $\lim_{t\to\infty} f(t) = 0$ by (69), which yields $\lim_{t\to\infty} I(t) = 0$ from the definition of $f(t)$. Hence, $\lim inf_{t\to\infty} I(t) = 0$. It follows that $\lim inf_{t\to\infty} E(t) = 0$ from the fluctuation lemma of Hirsch et al. [43] and Proposition 2.2 by Thieme [77]. Consequently, $\lim inf_{t\to\infty} f(t) = 0$ and $\lim_{t\to\infty} f(t) = 0$ because $f(t)$ is decreasing. □

5.2. **Logistic recruitment rate.** When logistic recruitment rate $B(N) = rN(1-N/K)$ is considered in model (51–54), the dynamics become determined by $\mathcal{R}_0$ and

$$\mathcal{R}_2^* = \frac{r}{\mu + d\frac{k}{\mu+d+r_2+k}\frac{\mathcal{R}_0-1}{\mathcal{R}_0}}.$$

For system (51–54), if $\mathcal{R}_0 \le 1$, the disease-free equilibrium is a global attractor; if $\mathcal{R}_0 > 1$ and $\mathcal{R}_2^* > 1$, there exists a unique endemic equilibrium that is globally asymptotically stable under some assumptions. Hence, a global forward bifurcation describes the dynamics if $\mathcal{R}_2^* > 1$ (see Fig. 7). To show the global stability of the endemic equilibrium, an equivalent monotone system to the original one was identified and a strong version of the Poincaré-Bendixon theorem applied. Further when the disease dies out, the decay is of exponential form with a rate proportional to $\mathcal{R}_0 - 1$. Particularly, interesting dynamics are observed when $1 < \mathcal{R}_0$ and $\mathcal{R}_2^* < 1$. Under these conditions, only the disease-free equilibrium is supported. One cannot study the stability of the disease-free equilibrium by the regular linearization approach since the vector field at the origin is not analytic. Simulations, however, strongly suggest that the disease-free equilibrium is a global attractor. A summary of the results obtained for this model can be found in Table 4. Feng et al. [31] obtained the global forward bifurcation in the case where the recruitment rate is a constant Λ.

6. **Cell-based models: Battle within a host.** The study of TB dynamics within an individual host is also within the realm of epidemics on populations of cells. We review some of these models in a rather superficial way as our work may find applications in immunology.

TABLE 4. Summary of the model with logistic recruitment rate.

expressions	criterion	trajectories
$\mathcal{R}_0 = (\frac{k}{k+\mu+r_1})(\frac{\sigma}{d+\mu+r_2})$	$\mathcal{R}_0 < 1$	$\lim_{t\to\infty} E(t) = \lim_{t\to\infty} I(t) = 0$ $\lim_{t\to\infty} N(t) = \frac{r-\mu}{r}$
$\mathcal{R}_1 = (\frac{k}{k+r+r_1})(\frac{\sigma}{d+r+r_2})$	$\mathcal{R}_1 < 1 < \mathcal{R}_0$	$\lim_{t\to\infty} E(t) = \lim_{t\to\infty} I(t) = 0$ $\lim_{t\to\infty} N(t) = \frac{r-\mu}{r}$
$\mathcal{R}_2^* = \frac{r}{\mu+\frac{dk(1-\mathcal{R}_0^{-1})}{m_\mu+k}}$	$1 < \mathcal{R}_1 < \mathcal{R}_0$ $\mathcal{R}_2^* < 1$	$\lim_{t\to\infty} N(t) = 0$ $\lim_{t\to\infty} E(t) = \lim_{t\to\infty} I(t) = 0$
$N^* = 1 - \mathcal{R}_2^{*-1}$ $E^* = \frac{m_\mu}{m_\mu+k}(1-\mathcal{R}_0^{-1})N^*$ $I^* = \frac{k}{m_\mu+k}(1-\mathcal{R}_0^{-1})N^*$	$1 < \mathcal{R}_1 < \mathcal{R}_0$ $1 < \mathcal{R}_2^*$	$\lim_{t\to\infty} N(t) = N^*$ $\lim_{t\to\infty} E(t) = E^*$ $\lim_{t\to\infty} I(t) = I^*$

6.1. **Co-infection with HIV.** Co-infection of TB and HIV has been considered to be the main cause for the reemergence of TB in the U.S.A.. Here we review a model for the dynamics of TB and HIV at the cell-level. Four populations are considered in a host who has HIV and TB. The model assumes that the compartments are the lymph tissues [48]. Letting $T(t)$ be the armed CD4$^+$ and CD8$^+$ T cells; $M(t)$ the macrophage population; $H(t)$ the HIV population; and $T_b(t)$ the *M. tuberculosis* population, the model equations of Kirschner [48] are

$$\frac{dT}{dt} = 0.5e_1 + 0.5\frac{e_2}{1+H} - \mu_T T + r_T T\left(\frac{H+T_b}{C+H+T_b}\right) - \gamma_1 HT, \tag{70}$$

$$\frac{dM}{dt} = \mu_m(M_0 - M) - \gamma_2 MH + r_m^2 MH + r_M^1 MT_b, \tag{71}$$

$$\frac{dH}{dt} = H(N_1\gamma_1 T + N_2 h_v M) - H(\gamma_3 T + \gamma_4 M) - \mu_h H, \tag{72}$$

$$\frac{dT_b}{dt} = r_{T_b} T_b(K - T_b) - \mu_b T_b - T_b(\gamma_5 T + \gamma_6 M), \tag{73}$$

where e_1 is the source rate of T cells in the absence of infection; e_2 is the rate of change in the T-cell supply; M_0 represents the equilibrium value for the macrophage population; N_1 and N_2 are the numbers of new viral particles produced by an infected T cell and macrophage, respectively; γ_3 is the rate at which CD8$^+$ T cells kill the virus; γ_4 is the rate at which macrophages kill the virus; T cells clear *M. tuberculosis* at rate γ_5; macrophages clear *M. tuberculosis* at rate γ_6.

In constructing this model, the growth rate of T cells is of Michaelis-Menten type, $\frac{H+T_b}{C+H+T_b}$, the growth of TB is of Logistic type, $r_{T_b}T_b(K-T_b)$, and all interactions are characterized by the mass action law. Model (70–73) undergoes local transcritical and Hopf bifurcations. Bistable equilibria are also found in some ranges of parameter space. The selection of μ_{T_b} and r_{T_b} as bifurcation parameters helps establish the existence of a local transcritical bifurcation. This cell-based model supports the hypothesis that the presence of TB worsens the clinical picture of HIV-infected individuals and that HIV helps activate TB infections. It also strengthens the view that TB treatment for HIV/TB infected individuals is essential.

6.2. **Multiple levels of macrophages.** Another cell-based model involving macrophages, T cells, and *M. tuberculosis* was proposed by Mathematical and Theoretical Biology Institute students in 1996 [42]. The function modeling the engulfing of bacteria by macrophages is characterized by three distinct stages: M_0, normal stage (before encountering bacilli), M_1 inactive stage (containing bacilli but inactive), and M_2 active state (activated by the help of T cells). *M. tuberculosis* is measured by the bacilli density in the blood of a host. The equations are

$$\begin{aligned}
\frac{dM_0}{dt} &= \Lambda_m - \mu_m M_0 - \epsilon M_0 T_b,\\
\frac{dM_1}{dt} &= \epsilon M_0 T_b - \alpha T M_1 - (\mu_m + g) M_1,\\
\frac{dM_2}{dt} &= \alpha T M_1 - \mu_m M_2,\\
\frac{dT}{dt} &= \Lambda_T + \alpha T M_1 - \mu_T T,\\
\frac{dT_b}{dt} &= b T_b - \mu_{T_b} - \epsilon M_0 T_b + (\mu_m N + gL) M_1,
\end{aligned}$$

where μ_m is the natural death rate of the macrophage population and g the bursting rate that results from the proliferation of bacilli inside the cytoplasm of a macrophage. If an inactive macrophage dies of natural causes, it releases N bacilli (on average); if it dies of bursting, then g bacilli are released. Usually, we assume $g \geq N$. Encounters of T cells with inactive macrophages stimulate the replication of T cells, which is modeled by the term rM_1T. An exponential growth rate for bacilli is assumed in the absence of macrophages. The total dynamical analysis for this model has not been completed.

7. **Geographic distribution.** A stochastic Markov chain model was proposed to study the impact of various geographic factors. Debanne et al. [25] constructed a multivariate Markovian model to project the distribution of TB cases across the U.S.A. by state, race and ethnic group. In their simulations, the transition probabilities from susceptible to infected are estimated via the annual risk for infection. Similar to the basic reproductive number, a contagious parameter is used to define the average number of susceptible infected by an infectious case [75, 76]. The effect of HIV is considered as an exogenous input. Because there is a high degree of uncertainty for the value of most parameters (only lower and upper bounds of parameters are available), a sensitivity analysis is conducted. The selection of parameters are used mostly to calibrate the model to data. The initial state (1980) was computed using inverse iteration (difference equations) back to the time when the actual data were available. The same model framework was applied to different racial and ethnic groups and different states by choosing appropriate parameters. Simulations results showed that TB cases change from state to state as well as among racial and ethnic groups. These simulations were put together and gave a picture of TB trends for the whole nation. Results showed that from 1980 to 2010 TB cases in the U.S.A. experienced an intermediate increase followed by a continuing decline. For instance, by 2010 the TB rate would be 4.6 per 100,000. This figure is higher than the expected case rate required to meet the CDC's goals.

8. **Control strategies.** A variety of dynamical models are proposed to study TB vaccination and elimination strategies. Our new results are included in subsections 8.3 8.4, and 8.5.

8.1. **Vaccination.** A TB vaccine called BCG (Bacillus of Calmette and Guerin) has been available for many decades, but its effectiveness in preventing TB is controversial [67]. Results of field trials of the vaccine have differed widely. Some indicates protection rates as high as 70% to 80%, while others indicates the vaccine is completely ineffective in the prevention of TB. We review three models with vaccinations, including an age-dependent model.

Two vaccine models, preexposure and postexposure, are proposed by Lietman and Blower [52]. In constructing their models, the susceptible and latent classes are divided into unvaccinated and vaccinated subclasses, specified by subscripts u and v, respectively. Preexposure vaccinations only vaccinate newborns at a proportion C with probability of success q. The immunity provided by vaccination is not permanent, leading to an expected effective vaccination time $\bar{\omega}$ that should be less than the average life-span. The formula for fast progression is the same as in model (3–5) and slow progression rates in this model are assumed to be dependent on the number of infectious via two general functions $k_u(I)$ and $k_v(I)$. The model equations are

$$\frac{dS_u}{dt} = (1 - Cq)\Lambda - \beta_u S_u I + \bar{\omega} S_v - \mu S_u, \tag{74}$$

$$\frac{dS_v}{dt} = Cq\Lambda - \beta_v S_v I - \bar{\omega} S_v - \mu S_v, \tag{75}$$

$$\frac{dE_u}{dt} = (1 - p_u)\beta_u S_u I - k_u(I)E_u - \mu E_u, \tag{76}$$

$$\frac{dE_v}{dt} = (1 - p_v)\beta_v S_v I - k_v(I)E_v - \mu E_v, \tag{77}$$

$$\frac{dI}{dt} = p_u\beta_u S_u I + p_v\beta_v S_v I + k_u(I)E_u + k_v(I)E_v - (d + \mu + r_2)I. \tag{78}$$

Postexposure vaccinations only apply to the individuals with latent TB. A third subclass of the latent-TB class is added to represent the waned state of vaccination for those with latent TB. The model equations for postexposure vaccinations are

$$\frac{dS_u}{dt} = \Lambda - \beta S I - \mu S_u, \tag{79}$$

$$\frac{dE_u}{dt} = (1 - p)\beta S I - k_u(I)E_u - \mu E_u - \phi E_u, \tag{80}$$

$$\frac{dE_v}{dt} = \phi E_u - \bar{\omega} E_v - k_v(I)E_v - \mu E_v, \tag{81}$$

$$\frac{dE_w}{dt} = \bar{\omega} E_v - k_w(I)E_w - \mu E_w, \tag{82}$$

$$\frac{dI}{dt} = p\beta S I + k_u(I)E_u + k_v(I)E_v + k_w(I)E_w - (d + \mu + r_2)I. \tag{83}$$

An interesting property of models (74–78) and (79–83) lies in their ability to identify distinct vaccination strategies for different nations that wish to eliminate TB. For developed countries, where the prevalence of latent TB is low, only a preexposure vaccine with treatment of active TB would be necessary. For developing nations

where the prevalence of latent TB is high, a combination of preexposure and post-exposure vaccine with treating active TB would be the most effective strategy. One must with care deal with these results, particularly for developing countries with rapidly expanding populations, as the models have no real demography.

BCG vaccination is age-dependent. The impact of BCG vaccination may be better assessed with the use of age-structured models. In fact, age has been included in most simulation models [10, 25, 34]. A linear partial differential equation model for TB transmission is proposed by Vynnycky and Fine [80]. They include both endogenous and exogenous reinfection in the model, plus the effects of BCG vaccination. There is no contact structure for the general population in this model. The infection and reinfection rates are modeled as explicit time-dependent linear functions (piecewise linear function). Numerically, the Euler method is applied to solve the corresponding linear non-autonomous PDE model. Parameters are estimated using data from across the world. Some parameters are estimated from the best fit required by the use of least squares methods. Their target populations are those of England and Wales, where HIV is not a significant factor. Hence, HIV is excluded from the model. Results from their model showed that ARI for age groups 0-15 years, 16-19 years and 20 years above are 4%, 9% and 14%, respectively.

To search for an optimal strategy for TB vaccination, another dynamical model including age structure, contact structure, and vaccination is proposed by Castillo-Chavez and Feng [16]. The model includes five subpopulations: susceptible (s), vaccinated (v), latent (l), infectious (i), and treated (j). Their model is given by the following system of PDEs:

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)s(t,a) = -\beta(a)c(a)B(t)s(t,a) - \mu(a)s(t,a) - \psi(a)s(t,a), \tag{84}$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)v(t,a) = \psi(a)s(t,a) - \mu(a)v(t,a) - \delta\beta(a)B(t)v(t,a), \tag{85}$$

$$\begin{aligned}\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)l(t,a) &= \beta(a)c(a)B(t)(s(t,a)+\sigma j(t,a)+\delta v(t,a))\\ &\quad - (k+\mu(a))l(t,a),\end{aligned} \tag{86}$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)i(t,a) = kl(t,a) - (r+\mu(a))i(t,a), \tag{87}$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)j(t,a) = ri(t,a) - \sigma\beta(a)C(a)B(t)j(t,a) - \mu(a)j(t,a), \tag{88}$$

$$\begin{aligned}&p(t,a') = \frac{c(a')n(t,a')}{\int_0^\infty c(u)n(t,u)du},\\ &s(t,0)=\Lambda,\quad v(t,0)=l(t,0)=i(t,0)=j(t,0)=0,\\ &s(0,a)=s_0(a),\ v(0,a)=v_0(a),\ l(0,a)=l_0(a),\\ &i(0,a)=i_0(a),\ j(0,a)=j_0(a).\end{aligned}$$

Parameter Λ is the birth rate; $\mu(a)$ is the age specific per-capita natural death rate; $\beta(a)$ is the age-specific (average) probability of becoming infected through contacts with infectious individuals; $c(a)$ is the age-specific per-capita contact/activity rate; σ and δ represent the reduction in risk of infection due to treatment and vaccination, respectively ($0 \le \sigma,\ \delta \le 1$); k is the per-capita rate of progression to active TB; r is the per-capita treatment rate; and $\psi(a)$ is the per-capita vaccination rate. In

addition, $p(t,a,a')$ gives the "probability" that an individual of age a has a contact with an individual of age a' given that the individual of age a had a contact with a member of the population. The infection force is $B(t) = \int_0^\infty \frac{i(t,a')}{n(t,a')} p(t,a,a')da'$; $p(t,a,a') = \frac{c(a')n(t,a')}{\int_0^\infty c(u)n(t,u)du}$ describes the contact structure corresponding to proportionate mixing [13]; and $n(t,a) = s(t,a) + v(t,a) + l(t,a) + i(t,a) + j(t,a)$ is the total population density.

The basic reproductive number $\mathcal{R}_0$ can be calculated following the next generation operator approach [26]. Let $S(\xi)$ denote demographic steady state of the population in the absence of disease; $A(\tau,\xi,\eta)$ denote the expected infectivity of an individual who was infected τ units of time ago while at age η; (that is, $A(\tau,\xi,\eta)$ denotes the average infectivity that can be exercised on an uninfected individual at age ξ provided the uninfected population finds itself at the steady demographic state $S(\xi)$)). Under the special assumption of proportionate-mixing [13]; that is, $A(\tau,\xi,\eta) = f(\xi)g(\tau,\eta)$, the basic reproductive number $\mathcal{R}_0$ is

$$\mathcal{R}_0 = \int_\Omega \int_0^\infty g(\tau,\eta)S(\eta)f(\eta)d\tau d\eta. \tag{89}$$

The demographic steady state (the state where the infection is absent) of the system is given by the following nonuniform age-distribution:

$$\begin{aligned} n(a) &= \Lambda e^{-\mu a},\ s(a) = \Lambda e^{-\mu a}\mathcal{F}_\psi(a), \\ v(a) &= \Lambda e^{-\mu a}\big(1 - \mathcal{F}_\psi(a)\big),\ l = i = j = 0, \end{aligned}$$

where

$$\mathcal{F}(a) = e^{-\int_0^a \mu(s)ds}, \quad \mathcal{F}_\psi(a) = e^{-\int_0^a \psi(b)db}.$$

The probability that an individual of age $\alpha + \tau$, who was infected τ units of time ago, is still in class i, is given by

$$\gamma(\tau,\alpha) = \int_0^\tau ke^{-(\mu+k)u}e^{-(r+\mu)(\tau-u)}du =: \mathcal{K}(\tau)e^{-\mu\tau}$$

where

$$\mathcal{K}(\tau) = \frac{k}{r-k}(e^{-k\tau} - e^{-r\tau}). \tag{90}$$

Thus the expected infectivity is given by

$$A(\tau,a,\alpha) = \beta(a)c(a)p(\alpha+\tau)\frac{\gamma(\tau,\alpha)}{n(\alpha+\tau)},$$

where $p(a) = c(a)n(a)/\int_0^\infty c(u)n(u)du$. Letting $A(\tau,a,\alpha) = f(a)g(\tau,\alpha)$, where $f(a) = \beta(a)c(a)$ and $g(\tau,\alpha) = p(\alpha+\tau)\gamma(\tau,\alpha)/n(\alpha+\tau)$, (89) leads to the computation of the reproductive number associated with the vaccination strategy ψ. Namely,

$$\mathcal{R}_0(\psi) = \int_0^\infty \int_0^\infty p(\alpha+\tau)\beta(\alpha)c(\alpha)\mathcal{K}(\tau)\mathcal{V}_\psi(\alpha)d\tau d\alpha,$$

where $\mathcal{V}_\psi(a) = \mathcal{F}_\psi(a) + \delta(1 - \mathcal{F}_\psi(a)) < 1$, and $\mathcal{K}(\tau)$ is given by (90). Hence, in the absence of a vaccine; that is, whenever $\psi(a) = 0$ the above formula reduces to

$$\mathcal{R}_0 = \int_0^\infty \int_0^\infty p(\alpha+\tau)\beta(\alpha)c(\alpha)\mathcal{K}(\tau)d\tau d\alpha.$$

Whenever the basic reproductive number $\mathcal{R}(\psi) < 1$, the disease-free steady state is *globally* stable provided that the total susceptible population has reached the demographic steady state $n(a)$.

Two optimization problems are proposed: if the goal is to bring $\mathcal{R}_0(\psi)$ to less than a pre-assigned value then find the vaccination strategy $\psi(a)$ that minimizes the total cost associated with this goal (reduced prevalence to a target level); if the budget is a fixed cost then find the vaccination strategy $\psi(a)$ that minimizes $\mathcal{R}_0(\psi)$; that is, that minimizes the prevalence. They found two possible optimal strategies, one-age strategy vaccinating the susceptible population at exactly age A; and two-age strategy vaccinating part of the susceptible population at exactly age A_1 and the remaining susceptible population at a later age A_2. The optimal strategies (one or two ages) also depends on the data, particularly on the type of cost function used.

8.2. **Treatment of latent TB.** Earlier models (prior to the 1970s) targeted the evaluation of control strategies of TB [10, 63, 81, 82], such as vaccination strategies. However, these "optimal" strategies have not worked well toward the elimination of TB globally or even regionally. The reasons behind the lack of success of these policies are debatable. Either these strategies have not been applied by the policy-makers or they are not truly "optimal". For instance, although 12% of GNP of the U.S.A. is spent on health care [5], the amount spent on prevention is very small. WHO estimates that if the amount of aid spent on TB treatment programs could be increased from 15 to 100 million dollars yearly, then 1.2 million deaths could be avoided every year [65]. That is, the death toll would be reduced by over 30%. This leaves the far-reaching question of what are the best strategies for the complete elimination of TB. Part of this work shows that the focus should include control measures in the latent-TB class. The reason is simple. The huge pool of latent-TB patients is a timebomb or reservoir of infection [62]. Forces or new diseases that compromise the immune system may lead to new TB outbreaks, as it has occurred with HIV. Mathematical models that stress the importance of treating individuals with latent TB are introduced by Blower [88] and Castillo-Chavez and Song [73]. Adding an early latent class and long-term latent class into the model (3–5) generates the following system:

$$\frac{dS}{dt} = \Lambda - \beta SI - \mu S, \tag{91}$$

$$\frac{dE_1}{dt} = \beta SI - (\mu + \omega + r_0)E_1, \tag{92}$$

$$\frac{dE_2}{dt} = (1-p)\omega E_1 - (k + \mu + r_1)E_2, \tag{93}$$

$$\frac{dI}{dt} = p\omega E_1 + kE - (\mu + d + r_2)I. \tag{94}$$

Early latent-TB individuals progress to active TB at the rate $p\omega$ and to long-term latent TB at the rate $(1-p)\omega$; long-term latent-TB individuals develop active TB at the rate k; and treatment rates are r_0, r_1, and r_2 for early latent TB, long-term latent TB, and active TB, respectively. This two-stage latent TB model distinguishes the risk of progression to active TB for primary infections and long-term infections, which is consistent with the fact that $p = 0.05 >> k = 0.00256$. (One has seen this in Ferebee's assumptions in section 3.) Model (91–94) allows one to explore the role of treating early latent-TB cases. Results from this two-stage

latent TB model show that treatment of 25% of early latent-TB cases together with treatment of 80% of active-TB cases may result in the elimination of TB.

8.3. **Elimination of TB in the U.S.A..** A comprehensive and executable model that leads to TB elimination in the U.S.A. is proposed by Song [72, 73]. Collecting public TB and demographic data for the U.S.A. for the past half century, a model of non-autonomous systems of ordinary differential equations is used to fit U.S.A. tuberculosis incidence over the past five decades. It is shown that tuberculosis in the U.S.A. may be controlled to the point of meeting CDC's criterion of one case per million, but only by the year 2020. This goal may be accomplished only if at least 20% of latently-infected individuals are treated. The effect of HIV/AIDS after 1983 is included in the analysis via a variation of our model that incorporates a function that accelerates TB progression over a window of time. It is shown that TB's case rate may be controlled despite increases in the rate of TB progression due to HIV. From the census and projection data, the total population size $N(t)$ is included in the model as an external input. Two latently-infected classes are introduced, primary latent/exposed class (L_1) and a permanently latent class (L_2). A non-autonomous ODE model with two latent classes and the incorporation of HIV stands

$$\frac{dL_1}{dt} = \beta\left(N(t) - L_1 - L_2 - I\right)\frac{I}{N(t)} - (\mu(t) + k + r_1 + p + A(t))L_1, \tag{95}$$

$$\frac{dL_2}{dt} = pL_1 - (\mu(t) + r_2 + A(t))L_2, \tag{96}$$

$$\frac{dI}{dt} = kL_1 + A(t)(L_1 + L_2) - (\mu(t) + d(t) + r_3)I, \tag{97}$$

where $A(t)$ has the form:

$$A(t) = \begin{cases} \alpha_1(t-1983)^{\alpha_2}e^{-\alpha_3(t-1983)^{\alpha_4}} & \text{if } 1983 < t, \\ 0 & \text{otherwise.} \end{cases}$$

Here α_i is constant to be determined via simulations. $N(t)$, now assumed to be independent of the disease, is a known "external" input to the epidemiological model. The values of $N(t)$ are in fact input from extrapolated published U.S.A. demographic data. The transmission rate β is assumed to be a constant; k, TB's activation rate, is also assumed to be constant; r_i $(i = 1, 2, 3)$, the treatment rates defined before, are also assumed to be constant; p is the rate at which primary latent-TB cases become permanent latent-TB cases; and $\mu(t)$ and $d(t)$ are functions of time. Estimates for some of these parameters are listed in Table 5. Fig. 9 shows

TABLE 5. Estimated parameters of TB transmission for the U.S.A..

parameter	β	c	k	r_1	r_2	r_3	p
estimation	0.22	10	0.01	0.05	0.05	0.65	0.1

that for the selected parameter ranges, the progression rate function $A(t)$ fits the data very well. That is, the fit successfully captures the past history of TB in the U.S.A.. The values of $r_1 = r_2 = 0.05$ means that in the past only 5% of latently-infected individuals got treated per year. The treatment of 100% of active-TB cases per unit time ($r_3 = 1$, instead of 0.65) is insufficient to reach CDC's goal (see

Fig. 10). However, treatment of more latently-infected individuals, for instance, raising r_1 and r_2 to 20% per year, would help reach CDC's target of 1/1,000,000 in a more reasonable period of time (see Fig.s 10 and 11). Fig. 12 illustrates the effect of HIV on the control of TB. It is clear that HIV delays the achievement of CDC's goal but has no permanent impact on

the long-term persistence of TB. However, prolonging TB's "survival" enhances the likelihood of its evolution, a situation that is not explored here. We have introduced the impact of HIV/AIDS on TB progression during the past two decades via a temporary perturbation on the distribution of TB progression times. Our selection of this perturbation is driven by our desire to fit active-TB data since our primary goal is to look at not the coevolution of co-infections but at HIV/AIDS co-infections on the ability of the U.S.A. to meet CDC's target by 2010. Our model suggests that if emphasis is placed on treating at least 20% of the latently-infected individuals then CDC's target may be met by 2020. Our model also shows that re-emergence of diseases that compromise the immune system (or recurrent outbreaks) would make it very difficult to control TB unless treatment emphasis is put on the earlier (non-detectable) stages of TB disease.

Theoretically sufficient conditions for TB extinction and persistence are derived in terms of upper limits and lower limits of the mortality functions for the non-autonomous model. We place this mathematical result in subsection 8.5

8.4. **Regression approach.** To partially back up our conclusions from a statistical viewpoint, we use regression to study the trend of new TB cases each year. We let the number of new cases be the response variable, denoted by Y, and time be the predictor, denoted by X. As can be seen from Fig. 1, the scatter plot of annual new cases Y versus year X appears to be exponential. Intuitively, a logarithmic transformation is taken on the response variable. The quadratic regression

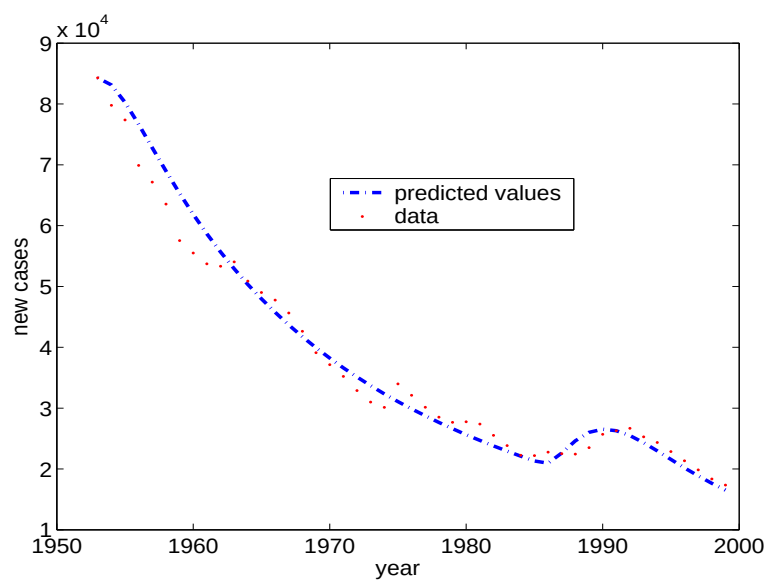


FIGURE 9. New cases of TB and data.

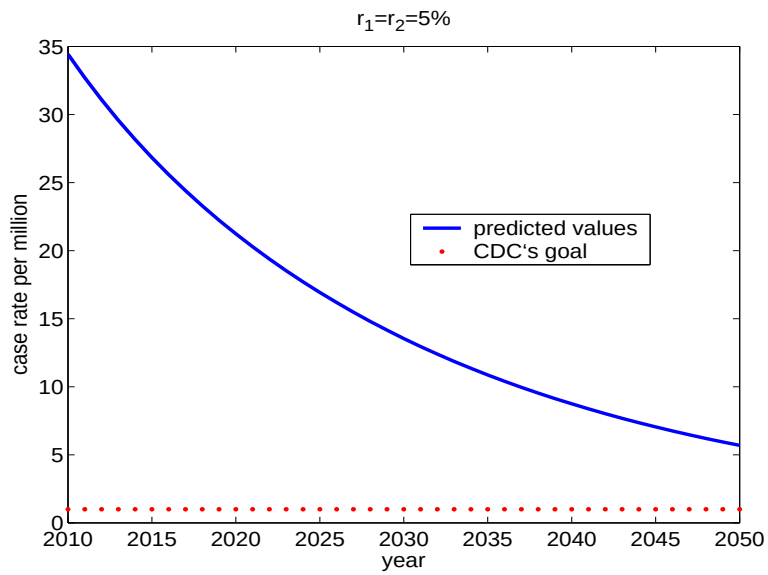


FIGURE 10. $r_1 = r_2 = 5\%$. CDC's "TB elimination" cannot be achieved by 2020.

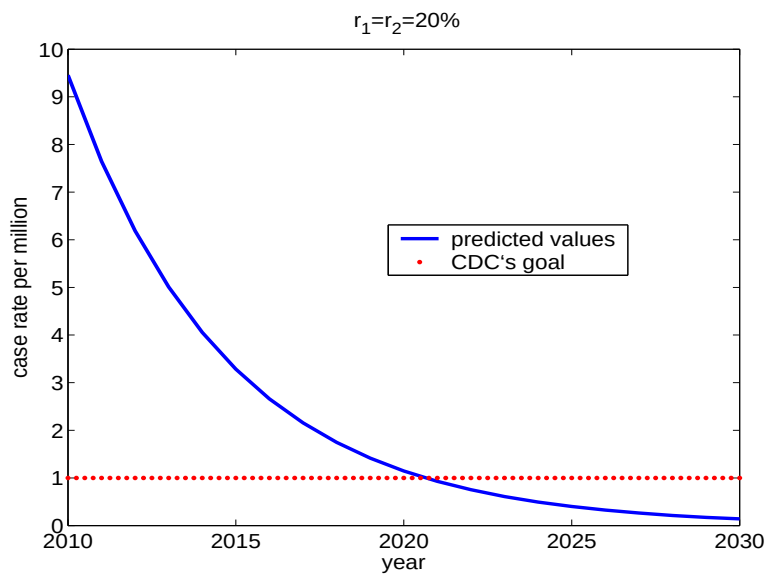


FIGURE 11. $r_1 = r_2 = 20\%$. CDC's "TB elimination" can be achieved by 2020.

is turned out to be the best fit. The regression equation

$$\log Y = 11.3970 - 0.0597X + 0.0006X^2, \tag{98}$$

is the best fit. Fig. 13 shows the fitted curve, 90% confidence bands, and 90% prediction bands.

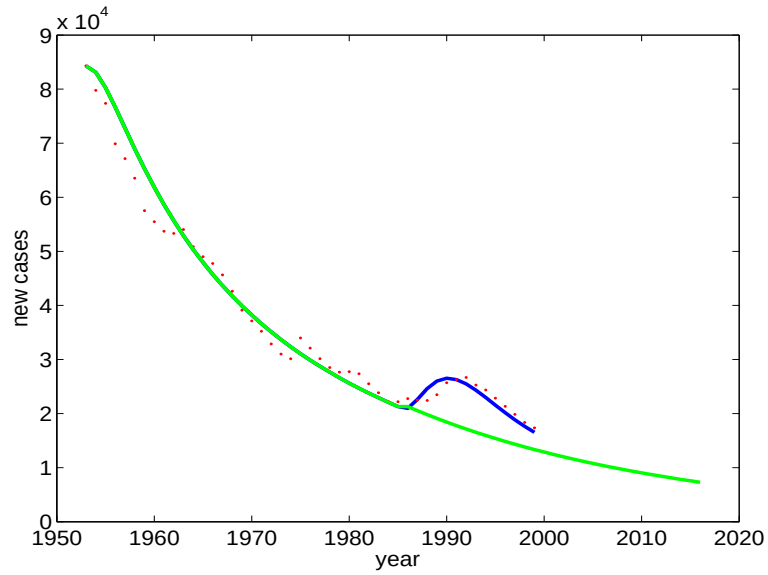


FIGURE 12. Impact of HIV. The lower curve represents no HIV effect; the upper curve represents the case rate when HIV is included; both are the same before 1983. Dots represent real data.

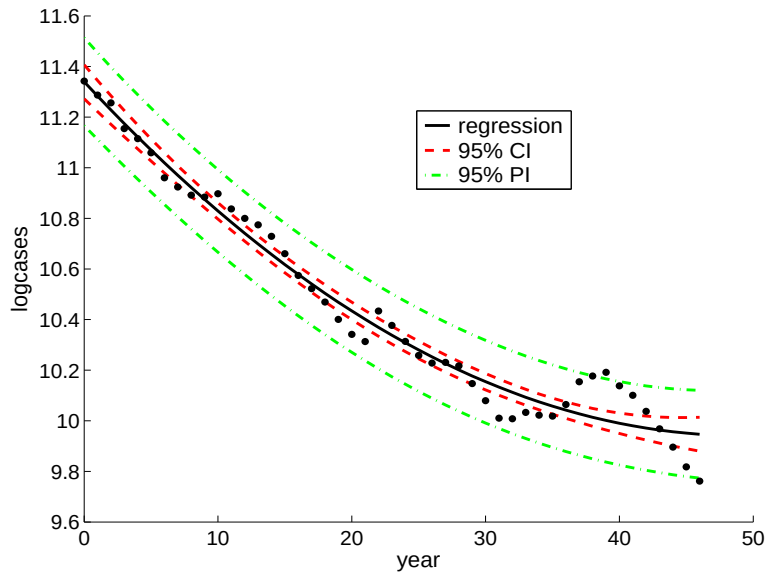


FIGURE 13. Quadratic regression $\log Y$ vs. X. Dots are the real data. Confidence bands and prediction bands are included.

The quadratic regression model turns out to be appropriate after the regression assumptions are verified. Consequently, we can use Equation (98) to predict the number of new cases in the near future. The results shows that the predicted case

rate in 2010 is $7.1659/1,000,000$. The same rate from our deterministic model (95–97) is $5.2952/1,000,000$. Both figures draw the same conclusion; that is, CDC's goals for TB elimination are unrealistic within the proposeed time horizon.

8.5. **Asymptotic behavior of the non-autonomous model.** An asymptotic analysis of model (95–97) is carried out and the results are discussed [72] in this section since they play a role in the parameterization of the model. The analysis helps establish a criterion for disease persistence (that is, a threshold condition) which must be met by the parameterized model. The long-term behavior of our system is determined by the asymptotic property of the functions $N(t)$, $\mu(t)$, and $d(t)$. The following theorem characterizes such behavior.

THEOREM 8.1. *Assume that* $\lim inf_{t\to\infty}\mu(t) = \mu_\infty$, $\lim inf_{t\to\infty}d(t) = d_\infty$, *and* $\lim sup_{t\to\infty}\mu(t) = \mu^\infty$, $\lim sup_{t\to\infty}d(t) = d^\infty$.

i. *If* $R_\infty = \left(\frac{k}{k+\mu_\infty+r_1+p}\right)\left(\frac{\beta}{\mu_\infty+d_\infty+r_3}\right) \le 1$ *then* $\lim\limits_{t\to\infty} L_1(t) = L_2(t) = \lim\limits_{t\to\infty} I(t) = 0$.
ii. *If* $R^\infty = \left(\frac{k}{k+\mu^\infty+r_1+p}\right)\left(\frac{\beta}{\mu^\infty+d^\infty+r_3}\right) \ge 1$ *then* $\lim sup_{t\to\infty}L_1(t) > 0$, $\lim sup_{t\to\infty}L_2(t) > 0$, *and* $\lim sup_{t\to\infty}I(t) > 0$.

Proof. We will use the following equalities, which are straightforward in real analysis. Whenever $\lim A$ exists, the following limit equalities hold:

$$\begin{aligned}
&\lim sup(A+B) = \lim A + \lim supB,\\
&\lim inf(A+B) = \lim A + \lim infB,\\
&\lim sup(AB) = \lim A \lim supB,\\
&\lim inf(AB) = \lim A \lim infB.
\end{aligned}$$

The proof is based on the fluctuation lemma [43] and its extension by Thieme (see Theorem 2.3 in [77]). Applying Theorem 2.3 from [77] to Equations (95) and (97) directly, one obtains that $0 \le \beta I^\infty - (\mu_\infty + k + r_1 + p)L_1^\infty$ and $kL_1^\infty \le (\mu_\infty + d_\infty + r_3)I^\infty$. It follows that $\beta I^\infty \ge \frac{\mu_\infty+k+r_1+p}{k}(\mu_\infty + d_\infty + r_3)I^\infty$; that is, $I^\infty\left(\frac{k}{k+\mu_\infty+r_1+p}\frac{\beta}{\mu_\infty+d_\infty+r_3} - 1\right) = I^\infty(R_\infty - 1) \ge 0$. Since $R_\infty < 1$ and $I(t)$ is bounded, it follows that $I^\infty = 0$. A similar argument results in $L^\infty = 0$. The first part of the theorem is proved.

It is not difficult to show that $\lim sup_{t\to\infty}L_1(t) = 0$ if and only if $\lim sup_{t\to\infty}I(t) = 0$. For instance, the fact that $\lim sup_{t\to\infty}I(t) = 0$ implies $\limsup\limits_{t\to\infty} L_1(t) = 0$ is verified below. From Equation (95), we obtain

$$\frac{dL_1}{dt} \le \beta I - (\mu(t) + k + r_1 + p)L_1.$$

It follows from the comparison principle that

$$L_1(t) \le \frac{L_1(0) + \int_0^t \beta I(s) e^{\int_0^s(\mu(\tau)+k+r_1+p)d\tau}ds}{e^{\int_0^t(\mu(\tau)+k+r_1+p)d\tau}}.$$

Hence,

$$\begin{aligned}
\lim sup_{t\to\infty}L_1(t) &\le \lim sup_{t\to\infty}\frac{\beta I(t)e^{\int_0^t(\mu(\tau)+k+r_1+p)d\tau}}{(\mu(t)+k+r_2)e^{\int_0^t(\mu(\tau)+k+r_1+p)d\tau}}\\
&= \lim sup_{t\to\infty}\frac{\beta I(t)}{\mu(t)+k+r_1+p} = 0.
\end{aligned}$$

The same argument can be used to show that $\lim sup_{t\to\infty} L_1(t) = 0$ implies $\lim sup_{t\to\infty} I(t) = 0$. It is also clear that $\lim sup_{t\to\infty} L_1(t) = 0$ implies $\lim sup_{t\to\infty} L_2(t) = 0$ from Equation (96). Two case are handled separately:

Case 1: If $kL_1(t) \geq (\mu(t)+k+r_1+p)I(t)$ holds for all $t > 0$, then $\frac{dI}{dt} > 0$ directly implies $\limsup_{t\to\infty} I(t) > 0$;

Case 2: If $kL_1(t) < (\mu(t) + k + r_1 + p)I(t)$ holds for all $t > 0$, then $\limsup_{t\to\infty} I(t) > 0$. The case 2 is proved by contradiction. Suppose $\lim sup_{t\to\infty} I(t) = 0$, then

$$\begin{aligned}
\frac{dL_1}{dt} &\geq \beta \frac{I(t)}{N(t)} \left(N(t) - \frac{\mu(t) + d(t) + r_3}{k} I - L_2 - I \right) \\
&\quad - \frac{(\mu(t) + d(t) + r_3)(\mu(t) + k + r_1 + p)}{k} I \\
&= \left(\beta - \frac{(\mu(t) + d(t) + r_3)(\mu(t) + k + r_1 + p)}{k} \right) I + o(I) \\
&\geq \left(\beta - \frac{(\mu^\infty + d^\infty) + r_3)(\mu^\infty + k + r_1 + p)}{k} \right) I + o(I) \\
&= \frac{(\mu^\infty + d^\infty) + r_3)(\mu^\infty + k + r_1 + p)}{k} (R^\infty - 1)I + o(I) > 0, \text{ for } t >> 1.
\end{aligned}$$

This implies that $\limsup_{t\to\infty} L_1(t) > 0$, which contradicts the assumption $\lim sup_{t\to\infty} I(t) = 0$. Trajectories of system (95–97) cannot intercept $kL_1(t) = (\mu(t) + d(t) + r_3)I(t)$ infinitely many times if $I(t) \to 0$ as $t \to \infty$, because $\frac{dL_1}{dt} > 0$ when $kL_1(t) = (\mu(t) + d(t) + r_3)I(t)$ whenever $R^\infty > 1$. Therefore, $\limsup_{t\to\infty} I(t) = 0$ implies that either $kL_1(t) > (\mu(t) + d(t) + r_3)I(t)$ or $kL_1(t) \leq (\mu(t) + d(t) + r_3)I(t)$ holds eventually. Hence, $\limsup_{t\to\infty} I(t) > 0$. □

When $\mu(t) \equiv \mu$ and $d(t) \equiv d$ (both constant), $R^\infty = R_\infty = \mathcal{R}_0$ gives the classical basic reproductive number

$$\mathcal{R}_0 = \beta \left(\frac{1}{\mu + d + r_3} \right) \left(\frac{k}{k + \mu + r_1 + p} \right)$$

where β is the effective contact rate; $\frac{1}{r_3+\mu+d}$ is the effective infectious period; and $\frac{k}{r_1+p+\mu+k}$ is the proportion of primarily-infected individuals who make it to the active stage.

The theorem provides conditions for differentiation of the two important biological states: disease elimination or persistence for this non-autonomous system. These thresholds are helpful not only in verifying the reasonableness of published parameters but also in the selection of reasonable ranges of unknown parameters. Our model generalizes the results established for related autonomous systems by Feng et al. [31] and Song et al. [74].

9. **TB transmitted by public transportation.** In Argentina, the TB incidence rate in the 1990s was 42/100,000 of the population, but this value was misleading since in the inner city of Buenos Aires it was 160/100,000 (four times higher than the national average). Buenos Aires has 12 million people, and 9.2 million passengers are carried by the bus system, accounting for 82% of the movement in public transportation [19]. A specific model targeting the population (in Buenos Aires)

that incorporates the impact of public transportation (buses) has been developed by Castillo-Chavez et al. [19].

The city is divided into N neighborhoods. Each neighborhood is further subdivided according to whether or not an individual frequently takes a bus. Type I individuals are those who seldom take buses or do not take them at all while type II individuals are those who frequently take buses. An individual of any type falls into one of four epidemiological groups at any time susceptible (S), infected but not infectious (E), infectious (I), and treated (T).

Type I and type II individuals have different levels of activity, which are modeled by the contact rates C_i^1 and C_i^2, where i indexes the neighborhood. The mixing structure of the population is driven by the bus system, which depends on the average time spent on the bus by the type I members of each neighborhood. To describe the model, let $Q_i^j = S_i^j + E_i^j + I_i^j + T_i^j$ be the total number of individuals of type-j in the ith neighborhood, $j = 1, 2$. The following parameters are required:

a_i = per-capita average contact rate of type I individuals in neighborhood i;
b_i = per-capita average contact rate of type II individuals in neighborhood i;
σ_i = per-capita rate of getting off the bus by type II individuals in neighborhood i;
ρ_i = per-capita rate of boarding a bus by type II individuals in neighborhood i.

Then,

$\frac{1}{\rho_i}$ = average time on a bus by a type II person;
$\frac{\rho_i}{\rho_i+\sigma_i}$ = probability of staying in the bus (type II person);
$\frac{\sigma_i}{\rho_i+\sigma_i}$ = probability of staying off the bus (type II person).

Proportionate mixing is assumed [13]. The mixing probabilities are calculated using the above definitions.

$P_i^{11} = \frac{a_i Q_i^1}{a_i Q_i^1 + b_i \frac{\sigma_i}{\rho_i+\sigma_i} Q_i^2}$ is the mixing probability of type I individuals within the same neighborhood i;
$P_i^{12} = \frac{b_i \frac{\sigma_i}{\rho_i+\sigma_i} Q_i^2}{a_i Q_i^1 + b_i \frac{\sigma_i}{\rho_i+\sigma_i} Q_i^2}$ is the mixing probability between type I and type II individuals within the same neighborhood i;
$P_i^{21} = \frac{a_i Q_i^1}{a_i Q_i^1 + b_i \frac{\sigma_i}{\rho_i+\sigma_i} Q_i^2} \left(\frac{\sigma_i}{\rho_i+\sigma_i}\right)$ is the mixing probability between type II and type I individuals within the same neighborhood i;
$P_i^{22} = \frac{b_i \frac{\sigma_i}{\rho_i+\sigma_i} Q_i^2}{a_i Q_i^1 + b_i \frac{\sigma_i}{\rho_i+\sigma_i} Q_i^2} \left(\frac{\sigma_i}{\rho_i+\sigma_i}\right)$ is the mixing probability between type II individuals in the ith neighborhood;
$P_{ij}^{22} = \dfrac{b_j \frac{\sigma_j}{\rho_j+\sigma_j} Q_j^2}{\displaystyle\sum_{i=1}^{N} \left(b_i \frac{\sigma_i}{\rho_i+\sigma_i} Q_i^2\right)} \left(\frac{\rho_i}{\rho_i+\sigma_i}\right)$ the mixing probability between type II individuals in the ith neighborhood and type II individuals in the jth neighborhood.

These mixing probabilities satisfy $P_i^{11} + P_i^{12} = 1$ and $P_i^{21} + P_i^{22} + \sum_{j=1}^{N} P_{ij}^{22} = 1$.

The recruitment rate Λ_i^j vary across neighborhoods and types. The natural mortality rate μ is assumed to be identical for all neighborhoods and all epidemiological groups. Treatment rates r_i, progression rate k_i, and the loss of immunity rate (treated person becomes susceptible again) α_i depend on the neighborhood

but not on the types. The model equations are

$$\frac{dS_i^j}{dt} = \Lambda_i^j - B_i^j(t) - \mu S_i^j + \alpha_i T_i^j, \tag{99}$$

$$\frac{dE_i^j}{dt} = B_i^j(t) - (\mu + k_i)E_i^j, \tag{100}$$

$$\frac{dI_i^j}{dt} = k_i E_i^j - (\mu + r_i + d)E_i^j, \tag{101}$$

$$\frac{dT_i^j}{dt} = r_i E_i^j - (\mu + \alpha_i)T_i^j, \tag{102}$$

$$Q_i^j = S_i^j + E_i^j + I_i^j + T_i^j,$$

where $i = 1, 2, 3, \ldots, N$, and $j = 1, 2$. The superscripts refer to the type and the subscripts index neighborhoods. The equations of type I and type II populations and the population across neighborhoods are coupled by the incidence rates $B_i^1(t)$ and $B_i^2(t)$, where

$$B_i^1(t) = \beta_i C_i^1 S_i \left(P_i^{11} \frac{I_i^1}{Q_i^1 + Q_i^2 \frac{\sigma_i}{\rho_i + \sigma_i}} + P_i^{12} \frac{I_i^2 \frac{\sigma_i}{\rho_i + \sigma_i}}{Q_i^1 + Q_i^2 \frac{\sigma_i}{\rho_i + \sigma_i}} \right),$$

$$B_i^2(t) = \beta_i C_i^2 S_i^2 \left(P_i^{21} \frac{I_i^1}{Q_i^1 + Q_i^2 \frac{\sigma_i}{\rho_i + \sigma_i}} + P_i^{22} \frac{I_i^2 \frac{\sigma_i}{\rho_i + \sigma_i}}{Q_i^1 + Q_i^2 \frac{\sigma_i}{\rho_i + \sigma_i}} + \sum_{j=1}^{N} P_{ij}^{22} \frac{I_j^2 \frac{\rho_j}{\rho_j + \sigma_j}}{Q_j^2 \frac{\rho_j}{\rho_j + \sigma_j}} \right).$$

This research found that the larger the difference of prevalence between neighborhoods, the larger the basic reproductive number. After estimating the relevant parameters, it was found that on average 100 people enter and leave the bus hourly, and that one TB infection per 1,000 travelers was generated per hour of travel. Using another model they found bus travel could be responsible for about 30% of new cases of TB [9, 19]. They also found that variations in TB transmission were most sensitive to transmission within the transportation system.

10. **Questions and conclusions.**

10.1. **An old prediction.** In the context of TB control, the importance of $\mathcal{R}_0$ was established in 1937, more than two decades before the introduction of first dynamical model for TB [36]. W. H. Frost, an epidemiologist at John Hopkins University, addressed the fundamental role of the reproductive number in the following way:

> "However, for the eventual eradication of tuberculosis it is not necessary that transmission be *immediately* and *completely* prevented. It is necessary only that the rate of transmission be held permanently below the level at which a given number of infection spreading (*i.e.*, open) cases succeed in establishing an equivalent number to carry on the succession. If in successive periods of time, the number of infectious hosts is continuously reduced, the end-result of this diminishing ratio, if continued long enough, must be extermination of tubercle bacillus."…. "This means that under present conditions of human resistance and environment the tubercle bacillus is losing ground, and that the eventual eradication of tuberculosis requires only that the present balance against it be maintained."[36]

TB is a slow disease, and therefore, *the basic reproductive number*$\mathcal{R}_0$ plays a fundamental role in the study its dynamics and control. Its role (transcritical bifurcation) goes well with the observed downward trend of TB mortality and incidence rates. In fact, it suggests that $\mathcal{R}_0$ is "moving" downward as parameters (naturally) change. Mathematically, the results are not surprising, because the models used are modifications of the frameworks developed by Kermack and McKendrick [47] and Ross [66].

10.2. **Challenging questions.** We have collected a number of dynamical models, the results, and insights that have generated in the study of TB dynamics. The historical evolution of dynamical models of TB follows a common pattern in biology from linear to nonlinear, from one strain to multiple strains, from homogenous to heterogeneous, from deterministic to stochastic, from empirical to theoretical (and vice versa). The models are given by system of difference equations, differential equations (ODEs and PDEs), integro-differential equations, and Markov chains. Although we have seen a remarkable progress in the development of a theoretical framework for the study of the dynamics of TB and other epidemiological processes, there many interesting and challenging topics and questions remain. A partial list includes

i. *Immigration*
All models "essentially" assume closed populations, ignoring the effects of immigration. Immigration is probably the critical factor in the generation of new TB cases. In the U.S.A. alone, over 40% of total new cases have been among immigrants in the past few years.

ii. *Race and ethnicity*
There is evidence showing that case rates of TB are different among different groups of people. These changes may be related to variations in susceptibility to the tubercle bacilli. We have seen a complex Markov chain model that takes this into account. However, more work is required if we are to better understand the role of race and ethnicity in disease dynamics.

iii. *Genetics*
The major reduction on TB mortality rates was achieved long before the introduction of antibiotics. Can this reduction be explained, at least partially, by the evolution of human susceptibility? Recent work by Aparicio et al. [4] provide a good start. The work on HIV and genetics by Hsu Schmitz [69, 70] suggests valuable approaches.

iv. *Sanitarium*
The sanitarium waxed and waned historically. It played a critical role in isolating and curing active-TB cases when antibiotics were not available. Models that incorporate the role of isolation on TB control are rare. Frost concluded that it could delay the number of cases [36].

v. *Global dynamics*
Theoretically, we characterized the global dynamics of a few models. For most models, the characterization of their global dynamics remains an open question. Multiple strain models, like (10–14) and models with fast and slow progression, like (3–5) should be further analyzed.

vi. *Time dependence*
The case of time-dependent coefficients is not only more realistic but often necessary [73]. Time-dependent parameters lead to the study of non-autonomous

models. To study the long-term behavior of models with time-dependent coefficients, threshold values (like the basic reproductive number) need to be further developed. Obviously, the classic approach for computing the basic reproductive number is not helpful [26]. We found upper and lower limits for the persistence and eradication of TB, but it seems hard to get the sharp threshold explicitly. The methods of averages by Ma et al. [55] may be helpful in the study of this problem. Time-dependent models provide a useful way of connecting parameters to data. The work of Aparicio et al. [4] shows this conclusively.

vii. *Mean latent period*
The distribution of the latent period is unknown as well as its mean. Knowledge of the shape of this distribution seems critical for control. The results of Feng et al. [31] have shown that it may not have an important qualitative role, but it certainly plays a critical quantitative role.

REFERENCES

[1] R. M. Anderson, R. M. May and A. R. Mclean (1988). Possible demographic impact of AIDS in developing countries, *Nature*, **332**:228–234.

[2] J. P. Aparicio, A. F. Capurro, and C. Castillo-Chavez (2000). Transmission and dynamics of tuberculosis on generalized households, *J. Theor. Biol.*, **206**:327–341.

[3] J. P. Aparicio, A. F. Capurro and C. Castillo-Chavez (2002). Frequency dependent risk of infection and the spread of infectious diseases, in *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: An Introduction*, C. Castillo-Chavez with S. M. Blower, P van den Driessche, D. Kirschner, and A. A. Yakubu (Eds.), IMA, Vol.125, Springer-Verlag, New York, 341–350.

[4] J. P. Aparicio, A. F. Capurro, C. Castillo-Chavez (2002a). Long-term dynamics and reemergence of tuberculosis, in: *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: An Introduction*, C. Castillo-Chavez with S. M. Blower, P. van den Driessche, D. Kirschner, A. A. Yakubu (Eds.), IMA, Vol.125, Springer-Verlag, New York, pp. 351–360.

[5] B. R. Bloom and C. J. L. Murray (1992). Tuberculosis: Commentary on a reemergent killer, *Science*, **257:**1055–1064.

[6] S. M. Blower, A. R. McLean, T. C. Porco, P. M. Small, P. C. Hopwell, M. A. Sanchez,and A. R. Moss (1995). The intrinsic transmission dynamics of tuberculosis epidemics, *Nature Medicine*, **1(8):**815–821.

[7] S. M. Blower, P. M. Small, and P. C. Hopwell (1996). Control strategies for tuberculosis epidemics: New models for old problems, *Science*, **273**:497–500.

[8] S. M. Blower and J. L. Gerberding (1998). Understanding, predicting and controlling the emergence of drug-resistant tuberculosis: A theoretical framework, *J.Mod. Med.*,**76:**624–636.

[9] D. Brand (1999). Tuberculosis spreads via crowded city buses, biomathemtician finds,*Cornell Chronicle*, Vol. 30, Number 20, February 4, 1999.

[10] S. Brogger (1967). Systems analysis in tuberculosis control: A model, *Amer. Rev. Resp. Dis.*, **95**:419–434.

[11] S. Busenberg and K. P. Hadeler (1990). Demography and epidemics, *Math. Biosci.*, **101:**41–62.

[12] S. Busenberg and P. van den Driessche (1990). Nonexistence of periodic solutions for a class of epidemiological models, in *Differential Equations Models in Biology, Epidemiology and Ecology*, S. Busenberg and M. Martelli (Eds.), Lecture Notes in Biology **92**, Springer-Verlag, Berlin-Heidelberg-New York, 71–79.

[13] S. Busenberg and C. Castillo-Chavez (1991). A general solution of the problem of mixing sub-populations, and its application to risk-and age-structured epidemic models for the spread of AIDS, *IMA J. of Mathematics Applied in Med. and Biol.*, **8:**1–29.

[14] J. Carr (1981), *Applications Centre Manifold Theory*, Springer-Verlag, New York.

[15] C. Castillo-Chavez and Z. Feng (1997), To treat or not to treat: The case of tuberculosis, *J. Math. Biol.*, **35:**629–656.

[16] C. Castillo-Chavez and Z. Feng (1998). Global stability of an age-structure model for TB and its application to optimal vaccination strategies, *Math. Biosci.*, **151**:135–154.

[17] C. Castillo-Chavez and Z. Feng (1998). Mathematical models for the disease dynamics of tuberculosis, in *Advances in Mathematical Population Dynamics: Molecules, Cells and Man*, O. Arino, D. Axelrod, and M. Kimmel (Eds.), World Scientific, 629–656.

[18] C. Castillo-Chavez and W. Huang (2002). Age-structure core group model and its impact on STD dynamics, in:. *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods and Theory*, C. Castillo-Chavez with S. M. Blower, P. van den Driessche, D. Kirschner, and A. A. Yakubu (Eds.), IMA, Vol. 126, Springer-Verlag, New York, 261–274.

[19] C. Castillo-Chavez, A. F. Capurro, M. Zellner, and J. X. Velasco-Hernandez (1998). El transporte pblico y la dinmica de la tuberculosis a nivel poblacional. *Aportaciones Matemaqticas, Serie Comunicaciones*, **22**:209–225.

[20] A. J. E. Cave (1939). *British Journal of Tuberculosis*, **33:**142.

[21] CDC (1989). A strategic plan for the elimination of tuberculosis in the United States,*MMWR*, **38:** (suppl. No. S-3) 1–25.

[22] CDC (1999). *Surveillance reports reported tuberculosis in the United States, 1999.* http://www.cdc.gov/nchstp/tb/surv/surv99/surv99.htm.

[23] A. Cockburn and E. Cockburn (1980). *Mummies, disease, and ancient Cultures*, Cambridge.

[24] T. M. Daniel (2000). *Captain of death: The story of tuberculosis*, University of Rochester Press, Rochester, New York.

[25] A. M. Debanne, R. A. Bielefeld, G. M. Cauthen, T. M. Daniel, and D. Y. Rowland (2000). Multivariate Markovian modeling of tuberculosis: Forecasts for the United States, *Emerging Infectious Diseases*, **6(2)**:148–157.

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

[27] P. van den Driessche and J. Watmough (2000). A simple SIS epidemic model with a bickward bifurcation, *J. Math. Biol.*, **40:**525–540.

[28] P. van den Driessche and J. Watmough (2001). Reproductive numbers and sub-threshold endemic equilibria for compartment models of disease trsmission, *Math. Biosci.*, **180:**29–48.

[29] R. Dubos and J. Dubos (1952). *The White Plague: tuberculosis, man, and society*, Little and Brown, Boston.

[30] J. Dushoff, W. Huang, and C. Castillo-Chavez (1998). Backward bifurcations and catastrophe in simple models of fatal diseases,*J. Math. Biol.*, **36**:227–248.

[31] Z. Feng, W. Huang, and C. Castillo-Chavez (2001). On the role of variable latent periods in mathematical models for tuberculosis, *Journal of Dynamics and Differential Equations*, **13(2):**425–452.

[32] Z. Feng, C. Castillo-Chavez,and A. F. Capurro (2000). A model for tuberculosis with exogenous reinfection, *Theor. Pop. Biol.*, **57**:235–247.

[33] Z. Feng, M. Iannelli, and F. Milner (2002). A two-strain TB model with age-structure, *SIAM J. Appl. Math.*, **62(5)**:1634–1656.

[34] S. H. Ferebee (1967). An epidemiological model of tuberculosis in the United States, *NTA Bulletin* , January, 4–7.

[35] J. A. Frimodt-Möller (1960). Tuberculosis study in a south India rural populations, 1950–1955, *Bull. World Health Organ.*, **22**:413.

[36] W. H. Frost (1937). How much control of tuberculosis, *American Journal of Public Health and the Nation's Health,* **27(8):**759–766.

[37] J. Gukerhamer and P. Homes (1983). *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, Springer-Verlag, New York.

[38] J. Guckenheimer (1996). Towards a global theory of singularly perturbed systems, *Progress in Nonlinear Differential Equations and Their Applications*, **19:**214–225.

[39] K. P. Hadeler and K. Ngoma (1990). Homogeneous models for sexually transmitted diseases, *Rocky Mountain, J. Math.*, **20**:967–986.

[40] K. P. Hadeler (1992). Periodic solutions of homogeneous equations, *J. Differential Equations*, **95**:183–202

[41] K. P. Hadeler and C. Castillo-Chavez (1995). A core group model for disease transmission, *J. Math. Biosci.* **128**:41–55

[42] C. Herrera, S. Lima, R. Munoz, G. Ramos, A. Rodriguez, and C. Salzberg (1996). A model describing the response of immune system to Mycobacterium tuberculosis, Department of Biometrics Technical Report Series # BU-1364-M, Biometrics Department, Cornell University.

[43] W. M. Hirsch, H. Hanisch, and J. P. Gabriel (1985). Differential equation models for some parasitic infections: methods for the study of asymptotic behavor, *Comm. Pure Appl. Math.*, **38:**733–753.

[44] F. Hoppensteadt (1974). Asymptotic stability in singular perturbation problems. II: Problems having matched asymptotic expansion solutions, *J. Differential Equations*, **15:**510–521.

[45] W. Huang, K. L. Cook,and C. Castillo-Chavez (1992). Stability and bifurcation for a multiple-group model for the dynamics of HIV/AIDS transmission, *SIAM J. Appl. Math.*, **52(3):**835–854.

[46] T. A. Kenyon, S. E. Valway, W. W. Ihle, et al. (1996). Transmission of multidrug-resistant Mycobacterium tuberculosis during a long airplane flight, *New England Journal of Medicine*, **334:**933–938.

[47] W. O. Kermack and A. G. McKendrick (1927). A contribution to the mathematical theory of epidemics, *Proc. R. Soc.*, **A,115:**700–721.

[48] D. Kirschner (1999). Dynamics of co-infection with M. tuberculosis and HIV-1, *Theoretical Population Biology*, **55:**94–109.

[49] G. Kolata (1995). First documented cases of TB passed on airliner is reported by U.S.A., *The New York Times*, March 3.

[50] C. M. Krisb-Zeleta and J. X. Velasco-Hernandez (2001). A simple vaccination model with multiple endemic states,*Math. Biosci.*, **164(2)**:183–201.

[51] S. A. Levin and D. Pimentel (1981). Selection of intermediate rates of increase in parasite-host systems, *Am. Nat.*, **117**:308–315.

[52] T. Lietman and S. M. Blowe (2000). Potential impact of tuberculosis vaccines as epidemic control agents, *Clinical Infectious Diseases*, **30(Suppl. 3):** S316–322.

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

[54] A. M. Lowell, L. B. Edwards, and C. E. Palme (1969). *Tuberculosis*, Harvard University Press, Cambridge, MA.

[55] Z. Ma, B. Song, and T. G. Hallam (1989). The threshold of survival for systems in a fluctuating environment, *Bull. Math. Biol.*, **51(3):**311–323.

[56] R. M. May and R. M. Anderson (1989). Possible demographic consequence of HIV/AIDS epidemics: II, assuming HIV infection does not necessarily lead to AIDS, in *Mathematical Approaches to Problems in Resource Management and Epidemiology*, C. Castillo-Chavez, S. A. Levin, and C. A. Shoemaker (Eds.), Lecture Notes in Biomathematics **81**, Springer-Verlag, Berlin, 220–248.

[57] R. M. May and R. M. Anderson (1985). Endemic infections in growing populations, *Math. Biosci.*, **77**:141–156.

[58] J. W. McGrath (1988). Social network of disease spread in the lower Illinios valley: a simulation approach, *Am. J. Phys. Anthropol.*, **77:**483–496.

[59] News in Brief (1999). WHO warning for air passengers, *Lancet*, **353**:305.

[60] T. C. Porco and S. M. Blower (1998). Quantifying the intrinsic transmission dynamics of tuberculosis, *Theoretical Population Biology*, **54:**117–132.

[61] J. Raffalli, K. A. Sepkowitz, and D. Armstrong (1996) Community-based outbreaks of tuberculosis, *Arch. Intern. Med.*, **156**:1053–1060.

[62] L. B. Reichman and J. H. Tanne (2002). *Timebomb: The global epidemic of multi-drug resistant tuberculosis*, McGraw-Hill, New York.

[63] C. S. ReVelle, W. R. Lynn, and F. Feldmann (1967). Mathematical models for the economic allocation of tuberculosis control activities in developing nations, *Am. Rev. Respir. Dis.*, **96**:893–909.

[64] C. S. ReVelle (1967). *The economics allocation of tuberculosis control activities in developing nations*, Cornell University.

[65] Reuters (1993). Agency cities urgent need to fight increase in TB. *The New York Times*, November 16, C8.

[66] R. Ross (1911). *The prevention of malaria*, Murry, London.

[67] A. A. Salyers and D. D. Whitt (1994). *Bacterial pathogenesis*, Washington D.C.

[68] P. A. Selwyn, D. Hartel, and V. A. Lewis et al. (1989). A prospective study of the risk of tuberculosis among intravenous drug users with human immunodeficiency virus infection. *N. Engl. J. Med.*, **320:**545–550.

[69] S-F. Shu Schmitz (2000a). A mathematical model of HIV transimission in homosexuals with genetic heterogenity, *Journal of Theoretical Medicine*,bf 2:285–296.

[70] S-F. Shu Schmitz (2000b). Treatment and vaccination against HIV/AIDS in homosexuals with genetic heterogenity, *Math. Biosci.*,**167:**1–18.

[71] B. Song, J. P. Aparicio, and C. Castillo-Chavez (2002). Tuberculosis models with fast and slow dynamics: The role of close and casual contacts, *Math. Biosci.*, **180:**187–205.

[72] B. Song (2002). Dynamical epidemical models and their applications, Cornell University.

[73] B. Song and C. Castillo-Chavez (2001). Tuberculosis control in the U.S.A.: A strategy to meet CDC's goal, Department of Biometrics Technical Report Series # BU-1562-M, Biometrics Department, Cornell University.

[74] B. Song, C. Castillo-Chavez, and J. P. Aparicio (2002). Global dynamics of tuberculosis models with density dependent demography, in *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods, and Theory.*, C. Castillo-Chavez with S. M. Blower, P. van den Driessche, D. Kirschner, and A. A. Yakubu (Eds.), IMA, Vol. 126, Springer-Verlag, New York, 275–294.

[75] K. Styblo (1985). The relationship between the risk of tuberculosis infection and the risk of developing infectious tuberculosis, *Bulletin of the International Union against Tuberculosis*, **60:**117–119.

[76] K. Styblo (1991). Epidemiology of tuberculosis, Selected Papers, **24**, Royal Netherlands Tuberculosis Association, The Hague.

[77] H. R. Thieme (1993). Persistence under relaxed point-dissipativity (with application to an endemic model), *SIAM J. Math. Anal.*, **24(2):** 407–435.

[78] U.S. Census Bureau (1975). *Historical statistics of the United States: Colonial times to 1970*, Government Printing Office Washington, D. C. .

[79] U.S. Census Bureau (1999). *Statistical Abstracts of the United States*,119th edition, Washington, D. C.

[80] E. Vynnycky and P. E. M. Fine (1997). The natural history of tuberculosis: The implications of age-dependent risks of disease and the role of reinfection, *Epidemiol. Infect.*, **119**:183–201.

[81] H. T. Waaler, A. Gese, and S. Anderson (1962) The use of mathematical models in the study of the epidemiology of tuberculosis, *Am. J. Publ. Health*, **52:**1002–1013.

[82] H. T. Waaler and M. A. Piot (1970). Use of an epidemiological model for estimating the effectiveness of tuberculosis control measures, *Bull. World Health. Org.*, **43:**1–16.

[83] W. F. Wells (1995). *Aerodynamics of droplet nuclei, airborne contagion, and air hygiene*, Harvard University Press, Cambridge.

[84] WHO (2001). *Global Tuberculosis Control, WHO report 2001*, Geneva.

[85] WHO (2000). *Global Tuberculosis Control, WHO report 2000*, Geneva.

[86] S. Wiggins (1990). *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Springer-Verlag, Berlin.

[87] Y. Ye (Ed.) (1986) *Theory of limit cycles*, Translations of mathematical monographs By American Mathematical Society, **66:** 245–260.

[88] E. Ziv, C. L. Daley, and S. M. Blower (2001) Early therapy for latent tuberculosis infection, *Am. J. Epidemiol.*, **153(4):**381–385.

Received on Feb. 27, 2004. Revised on June 20, 2004.

E-mail address: `chavez@math.la.asu.edu`
E-mail address: `songb@mail.montclair.edu`