J. Math. Biol. (1997) 35: 629–656

Journal of Mathematical Biology

To treat or not to treat: the case of tuberculosis

Carlos Castillo-Chavez, Zhilan Feng

Biometrics Unit, Cornell University, Ithaca, NY 14853, USA

Received 22 May 1995; received in revised form 28 February 1996

Abstract. Incomplete treatment of patients with infectious tuberculosis (TB) may not only lead to relapse but also to the development of antibiotic resistant TB – one of the most serious health problems facing society today. In this article, we formulate one-strain and two-strain TB models to determine possible mechanisms that may allow for the survival and spread of naturally resistant strains of TB as well as antibiotic-generated resistant strains of TB. Analysis of our models shows that non-antibiotic co-existence is possible but rare for naturally resistant strains while co-existence is almost the rule for strains that result from the lack of compliance with antibiotic treatment by TB infected individuals.

Key words: Tuberculosis – Antibiotic resistance – Epidemiology – Co-existence – Dynamical systems

Introduction

Infectious diseases like measles, influenza, chicken pox, and rubeola have several features in common; for example, they cause recurrent epidemic outbreaks and transmission rates depend strongly on age-dependent contact rates. The etiological agents of these communicable diseases are viruses from different families but all capable of generating similar epidemiological responses at the level of the individual (symptoms). Common responses include relatively short latent periods, followed by also relatively short infectious periods and permanent immunity after recovery. It is not completely clear when individuals become infectious (that is, capable of transmitting the disease) as some may become infectious while symptomless. Effective vaccines have been developed for these communicable diseases (to some degree only influenza remains the major challenge, as the family of viruses responsible for the "flu" experiences continuous minor and major genetic changes). A dense

and mature literature associated with the use of mathematical models to study communicable diseases such as measles, influenza, rubeola, and chicken pox is already in place (see Hethcote 1976; Dietz 1979; Hethcote, Stech, and van den Driescsche 1981; Anderson 1982; Anderson and May 1982, 1991; Dietz and Schenzle 1985; Dietz 1985; Anderson and May 1983; Schenzle 1984; Hethcote and Van Ark 1987; Castillo-Chavez et al. 1988, 1989; Feng 1994; Feng and Thieme 1995). The situation of tuberculosis, despite its fundamental role in the development of bacteriology and modern epidemiology, is paradoxically different.

Tuberculosis (TB) is a bacterial disease with about one third of the world human population as its reservoir (Bloom 1994; Miller 1993). It is one of the oldest recorded human diseases (it seems clear that TB has afflicted animal populations before the origin of the human species). Evidence that supports human cases of TB as well as its role in human mortality goes back for centuries (petrified bones 8000 B.C., Hindu texts from 2000 B.C. and mummified reliquiae from Egypt and pre-Columbian America including an Incan child 700 A.D.). TB was so devasting that it became the motivating force in the development of the fields of bacteriology, modern epidemiology, and public health. TB or TB associated symptoms appear to have been the source of inspiration for Frascatorius' theory of contagion (18th century). However, a search for a cause without a clear understanding of the sources and nature of disease, naturally led to what Ayvazain (Ayvazain 1993) calls "centuries of nonscientific chaos."

The situation changed when Villeman (19th century) used animal models to establish TB as a specific infection due to an inoculable agent (Reichman and Hershfield 1993). On March 29, 1882, Robert Koch presented to the Berlin Physiologic Society the results of his research on the causes of disease. Koch's fundamental research identified the mechanisms for disease transmission and the agents responsible for some diseases, including the etiological agent of TB (Reichman and Harshfield 1993). Koch's research opened new doors and eventually led to the discovery by various investigators of other bacteriological disease agents including the bacilli for typhoid, glanders, and diphtheria.

Despite its sociological and historical importance, the study of the spread of TB using statistical and mathematical models has not received enough attention. In fact we have observed only an extremely limited use of mathematical models in the study of the transmission dynamics of TB in human populations (personal communication with Blower during the meeting of Mathematical Modeling of Tuberculosis in 1995). Tuberculosis is caused by *Mycobacterium tuberculosis*. The disease is most commonly transmitted from a person suffering from infectious (active) tuberculosis to other persons by infected droplets created when the person with active TB coughs or sneezes. Among generally healthy persons, infection with TB is highly likely to be asymptomatic. Data from a variety of sources suggest that the life time risk of developing clinically evident TB after being infected is approximately 10%, with 90% likelihood of the infection remaining latent (Hopewell 1994).

Individuals who have a latent TB infection are not clinically ill nor capable of transmitting TB (Miller 1993). At greater ages, the immunity of persons who have been previously infected may wane, and they may be then at risk of developing active TB as a consequence of either exogenous reinfection (i.e., acquiring a new infection from another infectious individual) or endogenous reactivation of latent bacilli (i.e., re-activation of a pre-existing dormant infection) (Styblo 1991; Smith 1994).

The epidemiology of TB disease is not simple. For the purpose of this article we only provide a superficial view which we believe is sufficient for a rough understanding of the dynamics of TB transmission at the population level. General sources of information on TB dynamics suggest that TB is hard to transmit. Transmission (it is said) occurs only when there is prolonged close contact between a susceptible person and a person who has an active case of TB. Nonetheless, under the right conditions a single person with active TB can infect many other people (Salyers 1994). For example, it seems that about 13 persons were infected with TB per year by one source of infection in a Netherlands community in the period 1921–1938 (Styblo 1991). However, it is not clear that TB is in fact hard to transmit. Recent documented cases of TB transmission during lengthy plane trips (Kolata 1995; MMWR 1995) seem to indicate that transmission may be highly facilitated in a modern society. It is not at all unlikely that the risk of infection may be quite high in public places where there are actively infected TB individuals present. Recently mathematical models have been developed to estimate the probability of transmission of TB in close public environments. These models support the view that the acquisition of TB infection may not be as difficult as previously thought (Edward Nardell 1995). A naive look at the fact that one third of the world population is actually infected suggests that either the tubercle bacillus is easy to acquire, or that in many parts of the world exposure and re-exposure to TB is extremely persistent, or both. Current epidemiological studies strongly support the claim that exposed individuals are unable to transmit the tubercle bacillus and that only individuals with "active" TB are capable of spreading this bacteria. Therefore, exposed individuals provide a tremendously large reservoir for the tubercle bacillus but as latent carriers of this bacillus they are uncapable of transmission. What are the epidemiological consequences of this situation in a world where populations become closer and closer? Here lies one of the central issues associated with the study of TB dynamics.

Exposed TB individuals may remain in this latent stage for variable periods of time (in fact, many die without ever developing active TB). Apparently, the longer that we carry this bacteria the less likely we are to develop active TB unless our immune system becomes seriously compromised by other diseases. Consequently, age of infection as well as chronological age are important factors in disease progression. How important are these factors as predictors or measures of spread at the population level? Because it has been estimated that 10% of those infected with TB actually develop active TB during their life time then the 10% rule has become a useful measure for rough and immediate public health measures. This rule is useful but at the same time

it is also superficial. It is well known that TB progression is not uniform but in fact is closely linked to various other factors such as nutritional status and/or access to decent medical care and living conditions (Bloom 1994). The good news is that latent and active TB can be treated with antibiotics. The bad news is that its treatment has side effects (sometimes quite serious) and takes a long time. Carriers of the tubercle bacillus who have not developed TB disease can be treated with a single drug INH; unfortunately, it must be taken religiously for 6–9 months. Treatment for those with active TB requires the simultaneous use of three drugs for a period of at least 12 months. Lack of compliance with these drug treatments (a very serious problem) not only may lead to a relapse but to the development of antibiotic resistant TB – one of the most serious public health problems facing society today.

TB remains the leading cause of death by an infectious disease in the world. TB is also the most prevalent infection in the world (Bloom 1994; Miller 1993). As stated before, a third of the world's population is a carrier of tuberculosis and is at risk for developing active TB. It is estimated that there are between 8 and 10 million new cases per year, of which about 3 million people die (Kochi 1991). In the United States, the estimated total number of TB infections lies between 10–15 million persons (Miller 1993). However, dramatic increases in the incidence of TB (new cases per year) have occurred within the United States over the past few years. From 1985 to 1991, the number of reported cases of TB has increased 18% with 26,283 cases reported in 1991 (Kent 1993). In 1991, a large California prison with 5,421 inmates and 1,500 staff members had 18 cases of active TB (Salyers 1994). Against the backdrop of an increasing incidence of TB in the United States there is a second problem, namely that of multi-drug resistant TB (MDR-TB). Resistant-TB develops when the treatment of a TB patient is inadequate or incomplete, thereby allowing some of the stronger/resistant bacilli to survive and prosper. Outbreaks of MDR-TB in the United States have begun to alarm doctors and public health officials. Over 80% of the patients in these outbreaks have died, often within weeks of being diagnosed as having tuberculosis. These problems are compounded by economics, as the cost of treating a patient with MDR-TB can exceed $250,000: nearly 100 times the cost of treating most other TB cases (Press release WHO/89 Nov. 1994). The emergence of the HIV epidemic has dramatically increased the risk of developing clinical TB in infected persons, substantially increasing TB rates globally (Miller 1993).

A TB vaccine called BCG (Bacillus of Calmette and Guérin) has been available for many decades. The BCG vaccine is cheap, costing about 10 cents per dose but its effectiveness in preventing TB is controversial (Salyers 1994). Results of field trials of the vaccine have differed widely, some indicating protection rates as high as 70% to 80%, others indicating the vaccine was completely ineffective in preventing TB (Salyers 1994). Potential problems associated with the generalized use of the BCG vaccine in some populations are closely associated to the fact that vaccinated individuals will test positive for TB. It becomes therefore nearly impossible to be able to detect the

prevalence of a disease in a population (like the Argentinian population) where most individuals are vaccinated.

The purposes of this paper are quite specific. We formulate a basic transmission model to study the dynamics of TB in as simple a setting as possible. The advantage of this approach is that using this simple setting, we are able to fully analyze the effects of basic epidemiological factors such as the latent and infectious periods on the dynamics of TB on a homogeneously mixing population. This model then becomes our basic structure to begin the study of the effects of resistant-TB on the same population. The difficulties in treating these infectious patients and their role in spreading drug-resistant bacilli to others is incorporated in our model.

This paper is organized as follows: Sect. 1 introduces a "simple" TB model. We compute its basic reproductive number and study its role on the dynamics and stability properties of this model. In Sect. 2 we introduce a two-strain TB model and study its dynamics under two distinct assumptions. First, we assume that we are only dealing with two competing strains and we find that co-existence is possible but "rare". Secondly we assume that the second strain is the result of antibiotic resistance and find that co-existence is common. Our mathematical results are based on quasi-steady state approximations. The time scales involved in the process (length of latent period, average life span of individuals) are not small enough to properly support this assumption. However, our extensive numerical simulations of the full model support our analytical conclusions. Our numerical bifurcation analysis of the full model also support our analysis. Section 3 is devoted to the discussion of numerical simulations that support or complete our analytical results. Our simulations involve the construction of bifurcation diagrams that support the results of Sect. 2. Section 4 details some of our current efforts and extensions including the incorporation of distributed delays, re-infection, and the effects of age-dependent contact rates. An appendix collects some of the mathematical details.

1 A simple TB model

In this section we introduce a simple model for the transmission of TB. The host population is divided into the following epidemiological class or subgroups: Susceptibles (S), Latent (L, infected but not infectious), Infectious (I), and (effectively) Treated (T) individuals. N denotes the total population. Using Fig. 1a we formulate the following model for TB:

$$\begin{aligned}
\frac{d}{dt}S &= \Lambda - \beta c S\frac{I}{N} - \mu S \\
\frac{d}{dt}L &= \beta c S\frac{I}{N} - (\mu + k + r_1)L + \beta' c T\frac{I}{N} \\
\frac{d}{dt}I &= kL - (\mu + d)I - r_2 I
\end{aligned} \tag{1.1}$$

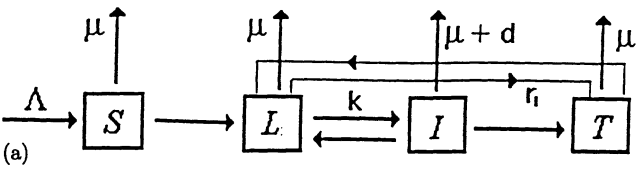


Fig. 1a. A diagram for one strain TB disease transmission. S stands for susceptible, L – exposed (latent), I – infectious, T – treated. Λ is the recruitment rate, μ is the per capita death rate, d is the disease-induced death rate (per capita), r_1 and r_2 denote the treatment rates for latent and infectious individuals, respectively. Individuals in S and T classes can be infected only through contacts with infectious individuals

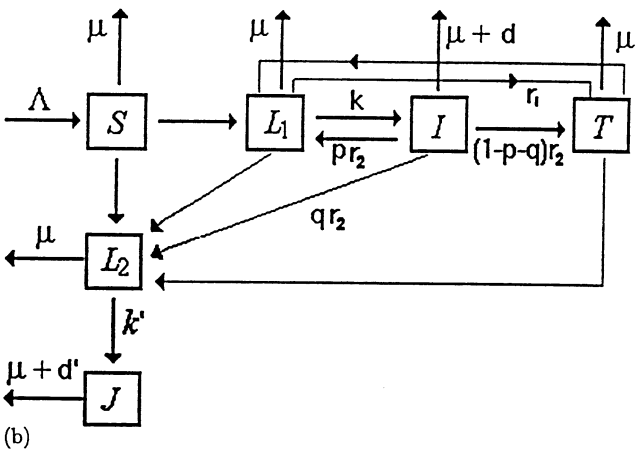


Fig. 1b. A diagram for two-strains TB transmission. L_1 and L_2 denote individuals exposed to typical TB and antibiotic resistant TB, respectively. J stands for infectious individuals with resistant TB. $p + q$ is the proportion of those treated infectious individuals who did not complete their treatment. The proportion p modifies the rate that departs from the latent class, and hence $qr_2 I$ gives the rate at which individuals develop resistant-TB because they did not complete the treatment of active TB. $p \geqq 0$, $q \geqq 0$ and $p + q \leqq 1$. $\Lambda, \mu, d, r_1, r_2$ have the same meanings as in Fig. 1a and d' is the disease (resistant TB) induced death rate.

$$\frac{d}{dt}T = r_1 L + r_2 I - \beta' c T \frac{I}{N} - \mu T$$

$$N = S + L + I + T\ .$$

Λ is the recruitment rate, β and β' are the probabilities that susceptible and treated individuals become infected by one infectious individual per contact per unit of time, respectively; c is the per-capita contact rate; μ is the per-capita natural death rate; k is the rate at which an individual leaves the latent class by becoming infectious; d is the per-capita disease induced death rate, and r_1 and r_2 are per-capita treatment rates. We assumed that an individual may be infected only through contacts with infectious individuals.

The basic reproductive number for (1.1) is

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

that is, the basic reproductive number is given by the product $\beta c/(\mu + r_2 + d)$, the average number of susceptibles infected by one infectious individual during his or her *effective* infectious period and $k/(\mu + r_1 + k)$, the fraction of the population which survives the latent period. Therefore $\mathscr{R}_0$ gives the number of secondary infectious cases produced by an infectious individual during his or her effective infectious period when introduced in a population of susceptibles.

There are two possible equilibria, i.e., the disease-free equilibrium denoted by E^0 and the endemic equilibrium denoted by E^* if $\mathscr{R}_0 > 1$ (the expression of E^* is given in Appendix).

System (1.1) is similar to some existing SEIR, SEI or SIRS models with different assumptions on either immunity or latent periods, or the disease-induced death rate (see, for example, Gao and Hethcote (1992); Pugliese (1991); Greenhalgh (1992)). Our model is sort of like a SEIS model with individuals (latent or infectious) going back to the "susceptible" class by treatment and with variable population size due to an extra mortality in the infectious class.

If we assume that the infection probabilities per contact for the treated class is the same as that of the susceptible class, i.e., $\beta' = \beta$, then the dynamics of System (1.1) is qualitatively similar to that of a *SEIS* model. We have established the following result:

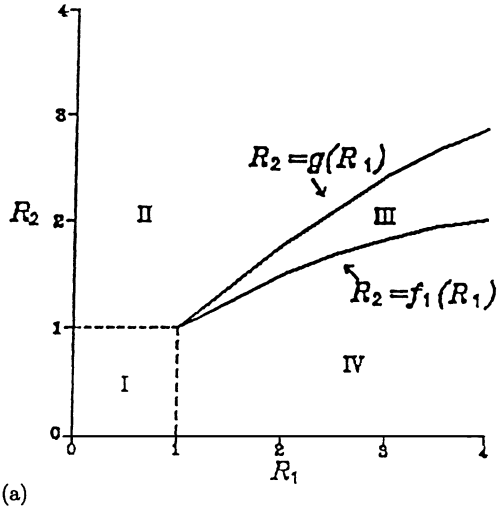
Theorem 1. *If $\mathscr{R}_0 < 1$ then the disease-free equilibrium is globally asymptotically stable (g.a.s.). While if $\mathscr{R}_0 > 1$ then the unique endemic equilibrium is locally asymptotically stable (l.a.s.).*

A proof of Theorem 1 can be found in the Appendix.

From the proof of Theorem 1 it is easy to see that if there is no treatment, that is, if $r_1 = r_2 = 0$ in System (1.1), then the qualitative dynamics are identical to those of the model with positive treatment rates ($r_1 > 0$, $r_2 > 0$). However, because $\mathscr{R}_0$ is a decreasing function of r_1 and r_2, and because the value of $\mathscr{R}_0$ decreases when we increase the treatment rates, then the disease levels (quantitative dynamics) are different. Treatment – as expected – reduces prevalence while increasing the fraction of non-infected individuals (see the expression for the endemic equilibrium in the proof of Theorem 1 found in the Appendix). We observe that no quasi-steady state approximation is needed for the proof of Theorem 1.

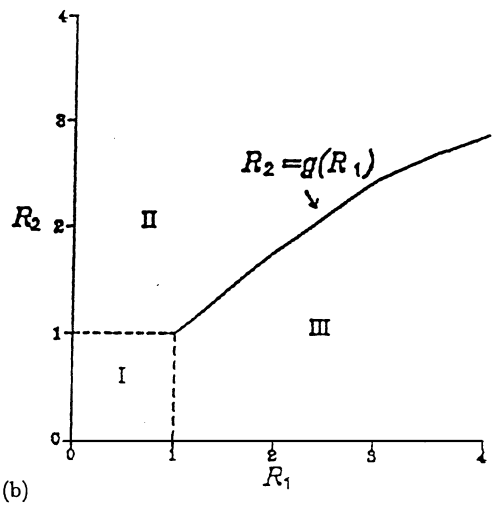
2 A two-strain TB model

The increasing recent number of outbreaks of active TB signal the creation of new opportunities for the development of resistant strains. In this section we modify our earlier model (System (1.1)) to take into account the possible appearance of resistant strains due to the deficient compliance with treatment schedules. We add two additional classes L_2 (latent) and J (infectious)



(a)

Fig. 2a. A bifurcation diagram for the system (2.5) in the case $q = 0$. There are four regions I, II, III, and IV in the parameter space $(\mathscr{R}_1, \mathscr{R}_2)$. In the region I E_1 is a global attractor and other equilibria are unstable when they exist. In regions II and IV E_3 does not exist and E_2 and E_4 are l.a.s., respectively. In the region III E_3 exists and is l.a.s.



(b)

Fig. 2b. A bifurcation diagram for the system (2.5) in the case $q > 0$. There are three regions I, II and III in the parameter space $(\mathscr{R}_1, \mathscr{R}_2)$ (E_4 does not exist.), and they correspond to stabilities of E_1, E_2, and E_3, respectively

representing the developmental stages of resistant strains. Since it is very hard to cure a patient with resistant TB we ignore the treatment of the resistant strain. Furthermore, we assume that J individuals can infect S, L_1, and T individuals. From the disease transmission diagram (see Fig. 2b) we can write the following system of ordinary differential equations:

$$\frac{d}{dt} S = \Lambda - \beta c S \frac{I}{N} - \beta^* c S \frac{J}{N} - \mu S$$

$$\frac{d}{dt} L_1 = \beta c S \frac{I}{N} - (\mu + k) L_1 - r_1 L_1 + p r_2 I + \beta' c T \frac{I}{N} - \beta^* c L_1 \frac{J}{N}$$

$$\frac{d}{dt}I = kL_1 - (\mu + d)I - r_2 I$$

$$\frac{d}{dt}T = r_1 L_1 + (1 - p - q)r_2 I - \beta' cT\frac{I}{N} - \beta^* cT\frac{J}{N} - \mu T$$

$$\frac{d}{dt}L_2 = qr_2 I - (\mu + k')L_2 + \beta^* c(S + L_1 + T)\frac{J}{N}$$

$$\frac{d}{dt}J = k'L_2 - (\mu + d')J$$

$$N = S + L_1 + I + T + L_2 + J\,, \tag{2.1}$$

where β^* is the probability that treated individuals become infected by one resistant-TB infectious individual per contact per unit of time; d' and k' have similar meanings as d and k; $p + q$ is the proportion of those treated infectious individuals who did not complete their treatment. The proportion p modifies the rate that departs from the latent class; $qr_2 I$ gives the rate at which individuals develop resistant-TB because they did not complete the treatment of active TB. Therefore $p \geqq 0$, $q \geqq 0$ and $p + q \leqq 1$.

For System (2.1) the first octant in the state space is positively invariant. By adding the equations in (2.1) we get the equation for $\frac{d}{dt}N$:

$$\frac{d}{dt}N = \Lambda - \mu N - dI - d'J\,. \tag{2.2}$$

Since $\frac{d}{dt}N(t) < 0$ for $N > \Lambda/\mu$, all solutions of (2.1) with nonnegative initial data approach, enter, or stay inside the subset Ω defined by $0 \leqq S + L_1 + I + T + L_2 + J \leqq \Lambda/\mu$. Hence, without loss of generality, we can only consider solutions of (2.1) on Ω. Using Lemma 1 found in the Appendix we get the inequality

$$N_\infty \geqq \frac{\Lambda}{\mu + d + d'}\,.$$

Since the right hand side of (2.1) is continuously differentiable there exists a unique solution on a maximal forward time interval for any nonnegative initial data, and hence the initial value problem (2.1) with initial data in Ω is well posed.

The basic reproductive numbers for the two-strain model are given by

$$\mathscr{R}_1 = \left(\frac{\beta c + pr_2}{\mu + d + r_2}\right)\left(\frac{k}{\mu + k + r_1}\right)$$

and

$$\mathscr{R}_2 = \left(\frac{\beta^* c}{\mu + d'}\right)\left(\frac{k'}{\mu + k'}\right),$$

respectively. We can interpret $\mathscr{R}_1$ and $\mathscr{R}_2$ as the average numbers of secondary infectious cases produced by an ordinary TB strain and one resistant-TB strain infectious individual during his or her *effective* infectious period, respectively.

If we let

$$\mathscr{R}_0 = \max\{\mathscr{R}_1, \mathscr{R}_2\} ,$$

then $\mathscr{R}_0 = 1$ gives a threshold condition. The disease will die out if $\mathscr{R}_0 < 1$ while the disease may become endemic if $\mathscr{R}_0 > 1$.

Next we consider the case when the infection probabilities per contact for the treated class is the same as that of the susceptible class, i.e., $\beta' = \beta$. Then System (2.1) together with (2.2) is equivalent to the following system:

$$\begin{aligned}
\frac{d}{dt}N &= \Lambda - \mu N - dI - d'J ,\\
\frac{d}{dt}L_1 &= \beta c(N - L_1 - I - L_2 - J)\frac{I}{N}\\
&\quad - (\mu + k)L_1 - r_1 L_1 + pr_2 I - \beta^* c L_1 \frac{J}{N} ,\\
\frac{d}{dt}I &= kL_1 - (\mu + d)I - r_2 I , \qquad (2.3)\\
\frac{d}{dt}L_2 &= qr_2 I - (\mu + k')L_2 + \beta^* c(N - I - L_2 - J)\frac{J}{N} ,\\
\frac{d}{dt}J &= k'L_2 - (\mu + d')J .
\end{aligned}$$

To simplify future expressions we introduce the following notation:

$$R_{1a} = \frac{\beta c}{\mu + k + r_1}, \qquad R_{1b} = \frac{k}{\mu + d + r_2} ,$$

When $q \neq 0$, the system (2.3) has three equilibria $E_i, i = 1, 2, 3$. (In this case there is no boundary equilibrium, that is, an equilibrium where only the first strain is present.) E_1 is the disease-free equilibrium

$$E_1 = \left(\frac{\Lambda}{\mu}, 0, 0, 0, 0\right) .$$

E_2 describes the case where only the second strain is present, that is,

$$E_2 = \left(N_2, 0, 0, (\mu + d')m\left(1 - \frac{1}{\mathscr{R}_2}\right)N_2,\ k'm\left(1 - \frac{1}{\mathscr{R}_2}\right)N_2\right),$$

where

$$N_2 = \frac{\Lambda}{\mu + d'k'm(1 - \frac{1}{\mathscr{R}_2})}, \qquad m = \frac{1}{\mu + d' + k'} .$$

E_2 exists only when $\mathscr{R}_2 > 1$ (for any $\mathscr{R}_1 > 0$). E_3 describes the case where both strains are present:

$$E_3 = (N^*, L_1^*, I^*, L_2^*, J^*) .$$

The expression of E_3 and the region where E_3 exists are described explicitly in the Appendix.

The case $q = 0$ gives the boundary equilibrium E_4 (if $\mathscr{R}_1 > 1$):

$$E_4 = \left(N_4, \frac{e(\mathscr{R}_1 - 1)}{R_{1b}} N_4, e(\mathscr{R}_1 - 1)N_4, 0, 0 \right),$$

where

$$N_4 = \frac{\Lambda}{\mu + de(\mathscr{R}_1 - 1)}, \qquad e = \frac{1}{R_{1a}(1 + R_{1b})}. \tag{2.4}$$

To conduct an analytical analysis of asymptotical behaviors of the equilibrium points we assume that there is no disease-induced death rate, i.e., $d = d' = 0$. For $d > 0$ and $d' > 0$ our numerical simulations support similar results (see Sect. 3). Hence we have that

$$\frac{d}{dt} N = \Lambda = \mu N ,$$

and, consequently, $N(t) \to \Lambda/\mu$ as $t \to \infty$. Without loss of generality (see Thieme 1992, 1994) we assume that our population has reached its limiting value, i.e.,

$$N \equiv \Lambda/\mu \equiv W + L_1 + I + L_2 + J .$$

By introducing the fractions

$$x_1 = \frac{L_1}{N}, \qquad x_2 = \frac{I}{N}, \qquad y_1 = \frac{L_2}{N}, \qquad y_2 = \frac{J}{N},$$

and eliminating the equation for $\frac{d}{dt} N$ we obtain from (2.3) the equivalent limiting system

$$\begin{aligned}
\frac{d}{dt} x_1 &= \beta c(1 - x_1 - x_2 - y_1 - y_2)x_2 - (\mu + k + r_1)x_1 + pr_2 x_2 - \beta^* c x_1 y_2 \\
\frac{d}{dt} x_2 &= k x_1 - (\mu + r_2)x_2 \\
\frac{d}{dt} y_1 &= q r_2 x_2 - (\mu + k')y_1 + \beta^* c(1 - x_2 - y_1 - y_2)y_2 \\
\frac{d}{dt} y_2 &= k' y_1 - \mu y_2 .
\end{aligned} \tag{2.5}$$

Obviously

$$0 \leqq x_1 + x_2 + y_1 + y_2 \leqq 1 \tag{2.6}$$

for all time $t \geqq 0$. With this notation, we are able to establish the following result:

Theorem 2. *Assume that* $q = 0$, $d = d' = 0$. *Then*

(a) *The disease-free equilibrium* E_1 *of System* (2.5) *is g.a.s. if* $\mathscr{R}_0 < 1$, *i.e., if* $\mathscr{R}_1 < 1$ *and* $\mathscr{R}_2 < 1$.

(b) *If* $\mathcal{R}_1 > 1$, *then there exists a threshold curve given by the function* $f(\mathcal{R}_1)$ *such that the boundary equilibrium* E_4 *of* (2.5) *is l.a.s. if* $\mathcal{R}_2 < f(\mathcal{R}_1)$ *and unstable if* $\mathcal{R}_2 > f(\mathcal{R}_1)$. *Moreover,* $f(\mathcal{R}_1) > 1$ *for all* $\mathcal{R}_1 > 1$ *and* $f(1) = 1$.

(c) *If* $\mathcal{R}_2 > 1$, *then there exists a second threshold curve given by the function* $g(\mathcal{R}_1)$ *such that the boundary equilibrium* E_2 *of* (2.5) *is l.a.s. if* $\mathcal{R}_1 < 1$ *or if* $\mathcal{R}_1 > 1$ *and* $\mathcal{R}_2 > g(\mathcal{R}_1)$. E_2 *is unstable if* $\mathcal{R}_1 > 1$ *and* $\mathcal{R}_2 < g(\mathcal{R}_1)$. *Moreover* $g(\mathcal{R}_2) > f(\mathcal{R}_1) > 1$ *for all* $\mathcal{R}_1 > 1$ *and* $g(1) = 1$.

(d) *The equilibrium* E_3 *of* (2.5) *exists if* $\mathcal{R}_1 > 1$ *and* $f(\mathcal{R}_1) < \mathcal{R}_2 < g(\mathcal{R}_1)$. *At the quasi-steady state* $\frac{d}{dt}x_2 = \frac{d}{dt}y_2 = 0$, *i.e., when* $x_2 = R_{1b}x_1$ *and* $y_2 = R_{2a}y_1$, *the corresponding positive equilibrium is l.a.s. when it exists.*

The proof of Theorem 2 is given in the Appendix. Our numerical simulations suggest that E_3 is l.a.s. not only at the quasi-steady state (when the conditions in (d) are satisfied) but possibly in general. Theorem 2 states that using $\mathcal{R}_1$ and $\mathcal{R}_2$ as parameters the existences as well as the stabilities of all possible equilibria of (2.5) can be completely determined by threshold conditions $\mathcal{R}_0 = 1$, $\mathcal{R}_2 = f(\mathcal{R}_1)$, and $\mathcal{R}_2 = g(\mathcal{R}_1)$ under the assumptions of the theorem. Figure 2a gives a bifurcation diagram for the case $q = 0$.

Remarks: 1. We see, from the proof of Theorem 2, that the functions $f(\mathcal{R}_1)$ and $g(\mathcal{R}_1)$ can be determined by fixing all parameter values related to the first strain except β. Hence $\mathcal{R}_1$ can be varied by varying β. Furthermore, since $\mathcal{R}_2$ is a monotone increasing function of β^*, the bifurcation diagram can be drawn using β and β^* instead of $\mathcal{R}_1$ and $\mathcal{R}_2$.

2. Our numerical simulations suggest that non-trivial equilibria $E_i (i = 2, 3, 4)$ are g.a.s. whenever they are l.a.s. when the quasi-steady state assumption is dropped.

We now consider the case $q > 0$. In this case System (2.5) has only three equilibrium E_1, E_2 and E_3. We have established the following result:

Theorem 3. *Assume that* $q > 0$, $d = d' = 0$. *Then*

(i) *The disease-free equilibrium* E_1 *of* (2.5) *is g.a.s. if* $\mathcal{R}_0 < 1$, *i.e., if* $\mathcal{R}_1 < 1$ *and* $\mathcal{R}_2 < 1$.

(ii) *If* $\mathcal{R}_2 > 1$, *then the boundary equilibrium* E_2 *of* (2.5) *is l.a.s. if* $\mathcal{R}_1 < 1$ *or if* $\mathcal{R}_1 > 1$ *and* $\mathcal{R}_2 > g(\mathcal{R}_1)$. ($g$ *is the function given in Theorem* 2(c)). E_2 *is unstable if* $\mathcal{R}_1 > 1$ *and* $\mathcal{R}_2 < g(\mathcal{R}_1)$.

(iii) *The equilibrium* E_3 *of* (2.5) *exists iff* $\mathcal{R}_1 > 1$ *and* $\mathcal{R}_2 < g(\mathcal{R}_1)$. (*In this case* E_2 *is unstable.*) *When it exists,* E_3 *is l.a.s. at the quasi-steady state* $\frac{d}{dt}x_2 = \frac{d}{dt}y_2 = 0$, i.e., *when* $x_2 = R_{1b}x_1$ *and* $y_2 = R_{2a}y_1$.

The proof of Theorem 3 can be found in the Appendix. Figure 2b gives a bifurcation diagram for the case $q > 0$. For the case when $d > 0$, $d' > 0$ our numerical simulations and bifurcation diagram also support similar results (see Fig. 5) when the quasi-steady state approximation assumption is dropped.

3 Numerical results

In this section we study the system (2.3) numerically to support our analytical results as well as to provide evidences that our results are likely to hold in more general situations. First we "extend" the result of Theorem 2(d) numerically. Our simulations support the stability of the interior equilibrium E_3 for the system (2.5) not only at the quasi-steady state but in general (see Fig. 3). Figure 3 presents some phase portraits for the system (2.5) which show that (for parameters in a clearly defined region (see Fig. 2)) the corresponding "l.a.s." equilibria $E_i (i = 2, 3, 4)$ attract all solutions with positive initial data (see Fig. 3). Similar simulations have been carried out when $q > 0$ to support the same conclusion (see Theorem 3) that the interior equilibrium E_3 is asymptotically stable whenever it exists not only at the quasi-steady state. The non-trivial equilibrium switches stability as the parameters change as specified in the bifurcation diagram (see Fig 2b and Fig. 4). We also "extend" the results of Theorem 3 numerically to the case when $d > 0$ and $d' > 0$. This is done by establishing explicit functional relationship between β and β' and by showing numerically that this function plays a role similar to that of the function g in Theorem 3 (see Fig. 5).

For the construction of Fig. 3 we have selected for illustration purposes the following parameter values: $\mu = 0.0143$ ($1/\mu = 70$ years), $\beta = 13$, $c = 1$, $k = 1$, $q = 0$, $p = 0.5$, $r_1 = 1$, $r_2 = 2$, $k' = 1$, $\Lambda = 35$, $d = d' = 0$. This choice of parameter values gives $\mathscr{R}_1 = 3.45$. Using the formula for $f(\mathscr{R}_1)$ and $g(\mathscr{R}_1)$ (see the Appendix) we get

$$(\mathscr{R}_1, \mathscr{R}_2) \in \text{III} \quad \text{iff } 1.34 < \mathscr{R}_2 < 4.13 ,$$

$$(\mathscr{R}_1, \mathscr{R}_2) \in \text{II} \quad \text{iff } \mathscr{R}_2 > 4.13 ,$$

$$(\mathscr{R}_1, \mathscr{R}_2) \in \text{IV} \quad \text{iff } \mathscr{R}_2 < 1.34 .$$

The value of $\mathscr{R}_2$ for Fig. 3a is chosen to be 2. Our simulations show that E_3 attracts all solutions with positive initial data. Values of $\mathscr{R}_2$ for Fig. 3b and Fig. 3c are 4.5 and 1.2, and our simulations support the global stabilities of E_2 and E_4, respectively.

Figure 4 is for the case when $q > 0$, and $d = d' = 0$. The parameter values used in Fig. 4 are $k = 0.5$, $k' = 1$, $\mu = 0.0143$, $q = 0.01$, $p = 0.4$, $r_1 = 2$, $r_2 = 1$, $d = 0$, $d' = 0$, $\Lambda = 500$. $\bar{\beta}$ is chosen to be 13 as it corresponds to $\bar{\mathscr{R}}_1 = 2.627$ and $\bar{\mathscr{R}}_2 = 2.7364$. This last selection implies that E_2 is l.a.s. if $\mathscr{R}_2 < 2.7364$ and it also implies that E_3 exists and is l.a.s. if $\mathscr{R}_2 > 2.7364$. In Fig. 4, the values for $\mathscr{R}_2$ are chosen to be (a) $\mathscr{R}_2 = 0.9$, (b) $\mathscr{R}_2 = 1.5$, (c) $\mathscr{R}_2 = 2$, and (d) $\mathscr{R}_2 = 3$.

Our analytic results for the stabilities of equilibria E_2 and E_3 (see Theorem 3) hold only for $d = 0$ and $d' = 0$. Since the death rate d' from resistant-TB may be high, one would like to know if similar results hold when $d > 0$ and $d' > 0$. Our numerical simulations suggest that this is the case. We first find from (2.3) a necessary condition under which the interior equilibrium E_3 exists

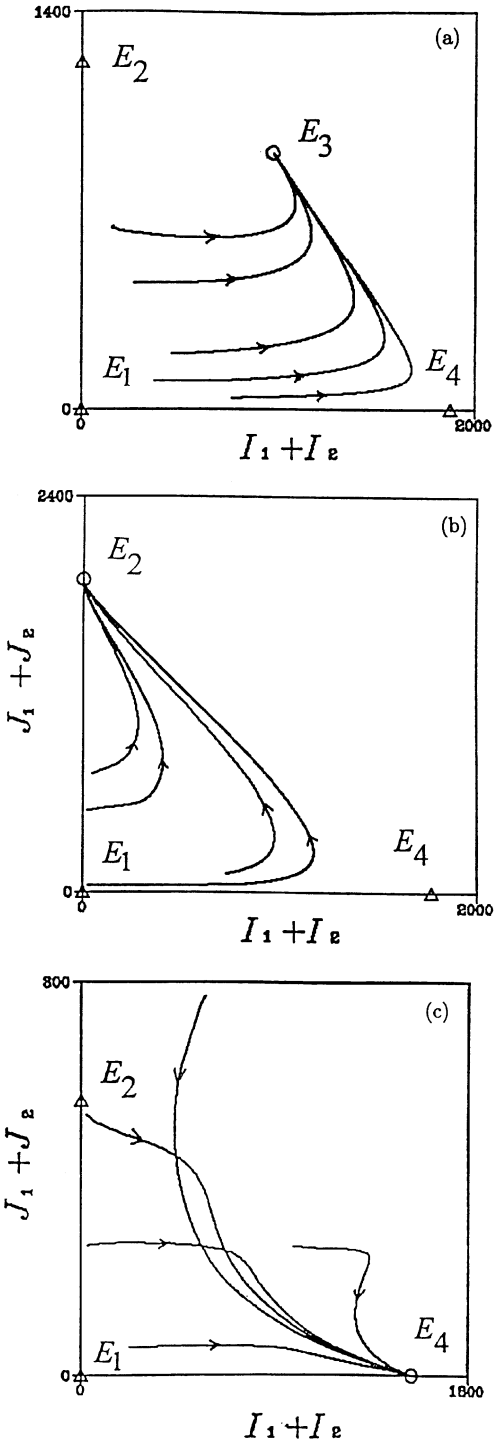


Fig. 3a–c. Phase portraits of solutions to (2.5). The parameter values for all three graphs are chosen to be: $\mu = 0.143$, $\beta = 13$, $c = 1$, $k = 1$, $q = 0$, $p = 0.5$, $r_1 = 1$, $r_2 = 2$, $k' = 1$, $\Lambda = 35$, $d = d' = 0$. This choice of parameters gives a fixed value $\mathscr{R}_1 = 3.45$. In **a** $\mathscr{R}_2 = 2$ and hence $(\mathscr{R}_1, \mathscr{R}_2) \in$ III. In **b** $\mathscr{R}_2 = 4.5$ and hence $(\mathscr{R}_1, \mathscr{R}_2) \in$ II. In **c** $\mathscr{R}_2 = 1.2$ and hence $(\mathscr{R}_1, \mathscr{R}_2) \in$ IV. Circle "○" indicates a stable equilibrium, and triangle "Δ" indicates an unstable equilibrium. (Graphs are given by Phsplan)

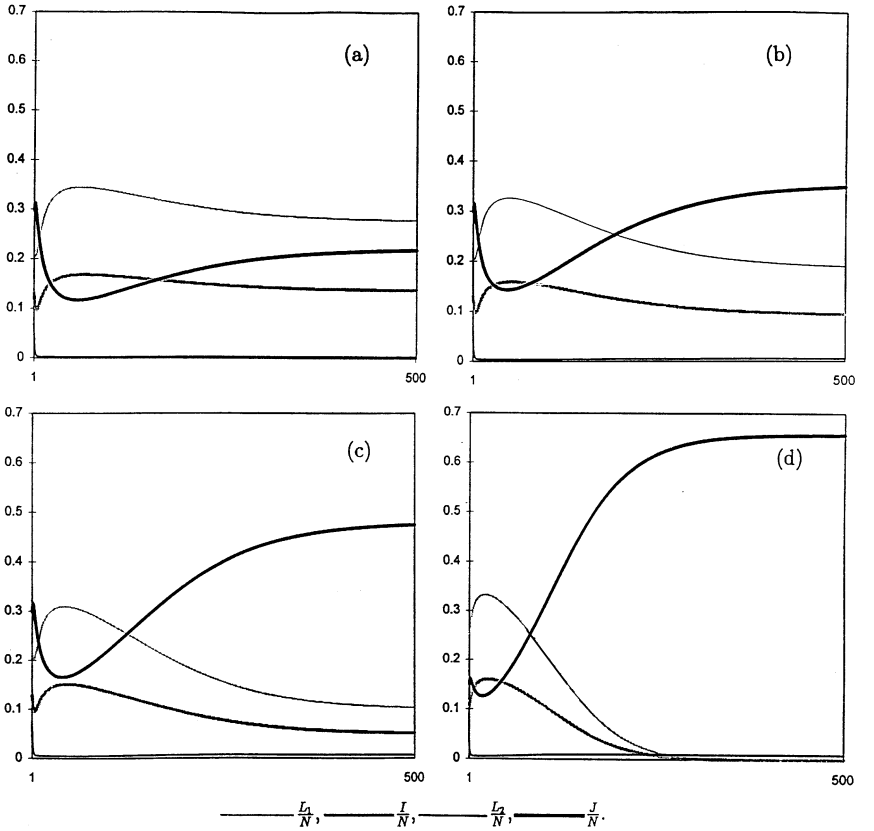


Fig. 4a–d. Plots of fractions of infected and infectious population versus time in the case $q > 0$, $d = d' = 0$. Parameters for all graphs are chosen to be: $\mu = 0.143$, $\beta = 13$, $c = 1$, $k = 0.5$, $q = 0.01$, $p = 0.4$, $r_1 = 2$, $r_2 = 1$, $k' = 1$, $\Lambda = 500$, $d = d' = 0$. In **a** $\mathscr{R}_2 = 0.9$ and hence $(\mathscr{R}_1, \mathscr{R}_2) \in$ III. In **b** $\mathscr{R}_2 = 1.5$ and $(\mathscr{R}_1, \mathscr{R}_2) \in$ III. In **c** $\mathscr{R}_2 = 2$ and $(\mathscr{R}_1, \mathscr{R}_2) \in$ III. In **d** $\mathscr{R}_2 = 3$ and $(\mathscr{R}_1, \mathscr{R}_2) \in$ IV

(see (25) in the Appendix). Using β and β^* as parameters we can establish a functional relationship between β and β^*, that is, $\beta^* = h(\beta)$ where

$$h(\beta) = \frac{1}{2c}(D_2 + \sqrt{D_2^2 + 4D_1\beta})$$

with

$$D_1 = cR_{1b}\left(1 + \frac{1}{R_{2a}}\right)\frac{\mu + k'}{R_{2a}},$$

and

$$D_2 = \frac{\mu + k'}{R_{2a}} + (pr_2R_{1b} - (\mu + k + r_1))\left(1 + \frac{1}{R_{2a}}\right).$$

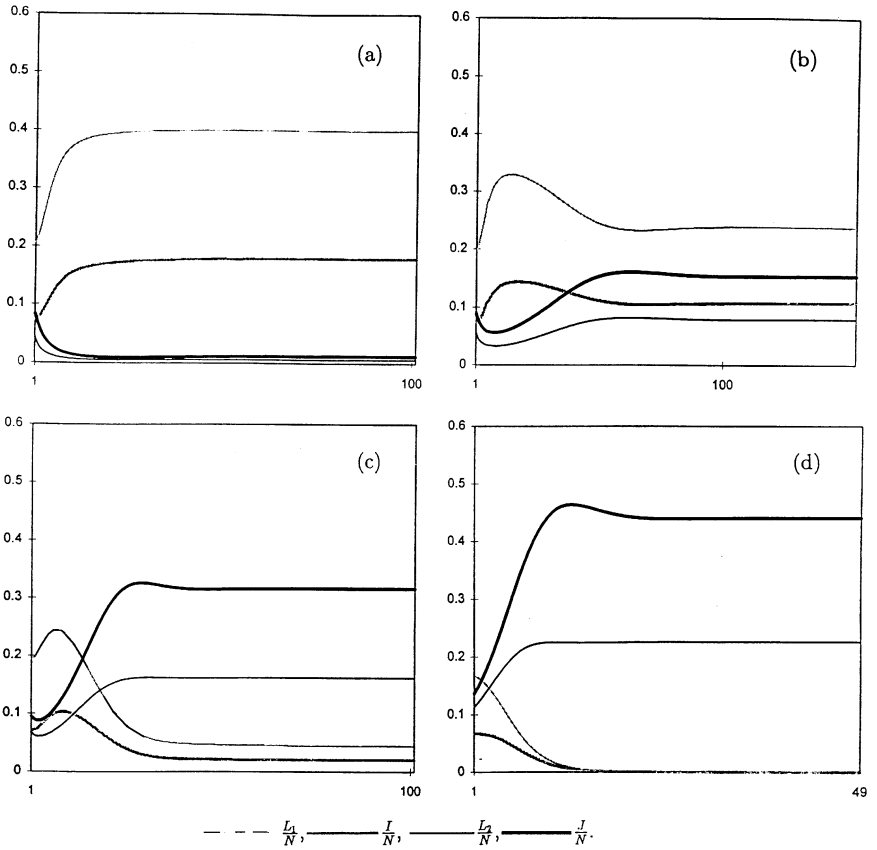


Fig. 5a–d. Plots of fractions of infected and infectious population versus time in the case $q > 0$, $d > 0, d' > 0$. Parameters for all graphs are chosen to be: $\mu = 0.143$, $\beta = 13$, $c = 1$, $k = 0.5$, $q = 0.01$, $p = 0.4$, $r_1 = 2$, $r_2 = 1$, $k' = 1$, $\Lambda = 500$, $d = 0.1$, $d' = 0.5$. In **a** $\mathscr{R}_2 = 0.9$ and hence $(\mathscr{R}_1, \mathscr{R}_2) \in$ III. In **b** $\mathscr{R}_2 = 1.5$ and $(\mathscr{R}_1, \mathscr{R}_2) \in$ III. In **c** $\mathscr{R}_2 = 2$ and $(\mathscr{R}_1, \mathscr{R}_2) \in$ III. In **d** $\mathscr{R}_2 = 3$ and $(\mathscr{R}_1, \mathscr{R}_2) \in$ IV

Then we have that $u^* = a - bv^* > 0$ (a necessary condition for E_3 to exist, see (11) and (14) in Appendix) iff

$$\beta^* < h(\beta), \mathscr{R}_1 > 1 \ . \tag{3.1}$$

By analogy with the proof for the case $d = d' = 0$ (see the proof of Theorem 3 in the Appendix) we would guess that E_3 exists and is l.a.s. if (3.1) holds and E_2 is l.a.s. if $\beta^* > h(\beta)$. Our guess is clearly supported by Fig. 5.

The parameter values used in Fig. 5 are the same as those used in Fig. 4 except that $d = 0.1$ and $d' = 0.5$. We choose $\bar{\beta}$ to be 13 (corresponding to $\bar{\mathscr{R}}_1 = 2.39$) and therefore $\bar{\beta}^* = h(\bar{\beta}) = 1.116$ (corresponding to $\bar{\mathscr{R}}_2 = 2.139$). We conclude that E_2 is l.a.s. if $\beta^* < 1.116$ (or $\mathscr{R}_2 < 2.139$) and E_3 exists and is l.a.s. if $\beta^* > 1.116$ (or $\mathscr{R}_2 > 2.139$). In Fig. 5 the values for β^* are chosen to be

(a) $\beta^* = 0.4695(\mathscr{R}_2 = 0.9)$, (b) $\beta^* = 0.78(\mathscr{R}_2 = 1.5)$, (c) $\beta^* = 1.04(\mathscr{R}_2 = 2)$, and (d) $\beta^* = 1.565(\mathscr{R}_2 = 3)$.

4 Discussion

In this paper we introduced a basic model to study the dynamics of resistant TB. First we analyzed a one-strain TB model in order to understand its transmission dynamics in the absence of resistance. We proceeded to analyze a two-strain model for TB and resistant-TB with the purpose of determining the role that lack of drug treatment compliance by TB patients plays on the maintenance of antibiotic resistant strains. To make the role of antibiotic resistance transparent, we first studied a special version of our two-strain model with two competing strains of TB: the typical strain plus a resistant strain that was not the result of antibiotic resistance. In this last situation, we found that co-existence is possible but rare while later we noticed that co-existence is almost certain when the second strain is the result of antibiotic resistance. In our two-strain model there is a superinfection-like term $\beta * cL_1J/N$. Is this necessary to obtain the co-existence result since it is well known that superinfection can cause co-existence (see Levin and Pimentel (1981) and Nowak and May (1994))? The answer is no. In fact is can be shown that in the absence of the superinfection-like term co-existence is still almost the rule when the second strain is the result of antibiotic resistance. Our mathematical analysis was based on quasi-steady state approximations. This kind of assumption usually can be justified analytically by using small parameter methods. But all parameters in Model (2.1) are of about the same order of magnitude. For example, the latent period $1/k$ for TB can range from 1 year to more than 20 years, the treatment period $1/r$ can range from 1/2 to 2 years, and the life expectancy $1/\mu$ is about 70 years. Therefore we can not really use small parameter method to support the assumption of quasistationary state. Our results under the quasi-steady state assumption were confirmed with the help of numerical simulations and the construction of bifurcation diagrams that support the plausibility of our hypotheses.

Our results show that coexistence of naturally resistant strains is limited. Deterministic models such as the one we have used here in fact suggest that competitive exclusion is the preferred outcome and not coevolution. Our results do not differ significantly from those developed by Levin and Pimentel (1991) in which it was found that coexistence of two strains of maxoma were only possible within a window of opportunity in parameter space. The recent work on super-infection by May and Nowak (1994) and Nowak and May (1994) suggest that super-infection would enhance coexistence and presumably coevolution. However, their models assume that the population size of the competing hosts are constant. If this assumption is removed then again the size of the region of coexistence shrinks and may even become a disconnected set (see Castillo-Chavez and Velasco-Hernández (1996); Mena-Lorca, Velasco-Hernández and Castillo-Chavez (1995)). In fact, we strongly believe

that the key mechanism for the support of species diversity is driven by spatial heterogeneity (see Durrett and Levin (1994); Bolker and Pacala (1995)).

In this article, we have just looked at the basic transmission dynamics of TB on a homogeneously mixing population (the null model). Our results are clear: such populations cannot support pathogen diversity, that is, competitive exclusion seems to be the preferred outcome. Furthermore, we found that antibiotic resistant (just as pesticide resistance) enhances – (one may even say promotes) – coexistence. This is not surprising but reminds us of the challenges facing public health officials. Resistant TB will remain a serious threat to our communities as long as many members of our society do not have regular access to medical care. The cost of treating individuals with active TB is over $20,000 while the price of infection with resistant strains of TB is very often death. Obviously more work needs to be done with models that incorporate more realism. We outline our progress on three fronts in the rest of this conclusion.

A natural criticism of our basic model is that it did not take into account long and variable periods of latency – a key feature of TB. We have in fact looked at the effects of long and variable periods of latency (rather than the exponentially distributed delays used in this article) and we have found that their addition makes no difference in the qualitative dynamics of TB (see Castillo-Chavez and Feng (1996a)). However, we chose not to incorporate this analysis in this article as its emphasis is on the study of resistant TB. The incorporation of distributed delays would have obscured the objective of this article while making the mathematics ugly. The coupling of a two-strain model with different distributed delays for their latent periods is at this point under investigation. However, we are not hopeful that we will be able to fully analyze such a complex model. Nevertheless, we plan to publish some preliminary results in this direction in the near future.

A person infected with TB may develop active TB in a variety of ways. One possibility is that such a person may develop active TB as a consequence of exogenous reinfection (i.e., acquiring a new infection from another infectious individual; Smith 1993). We have begun to look at the role that exogenous reinfection has on the transmission dynamics of common strains of TB (see Castillo-Chavez and Feng (1996a)). Our preliminary results seem to support our hypothesis that exogenous reinfection has a drastic effect on the qualitative dynamics of TB. More explicitly, the incorporation of exogenous reinfection into the basic TB model of Sect. 1 allows for the possibility of a subcritical bifurcation. That is, not only an endemic equilibrium may occur at the critical value of the reproductive number $\mathscr{R}_0 = 1$ but our system can have multiple endemic equilibria for $\mathscr{R}_0 < 1$. This type of behavior has been observed in recent epidemiological models but in the context of sexually-transmitted diseases (see Hadeler and Castillo-Chavez 1995, (1995)). Our analysis is almost complete and we plan to publish these results elsewhere in the near future.

Finally, it is clear that mixing plays a key role in TB transmission. We are particularly interested in looking at the effects of age-dependent contact rates on TB dynamics. The formulation of models with age-dependent contact

rates, even under the assumption of proportionate mixing, leads to hyperbolic systems of partial differential equations that are difficult to analyze. Nevertheless, we have managed to obtain some preliminary results and we have used them to study optimal vaccination policies for the BCG vaccine on age-structured populations (see Castillo-Chavez and Feng (1996b)).

Acknowledgements. This research was partially supported by NSF grant DEB-925370 (Presidential Faculty Fellowship Award) to Carlos Castillo-Chavez and by the U.S. Army Research Office through the Mathematical Science Institute of Cornell University (contract DAAL03-91-0027).

Appendix

In this Appendix we provide the details of the proofs of Theorems 1, 2 and 3 as well as the statements of needed preliminary results.

For a bounded real-valued function f on $[0, \infty)$ we define

$$f_\infty = \liminf_{t\to\infty} f(t), \qquad f^\infty = \limsup_{t\to\infty} f(t) .$$

Lemma 1 [Thieme (1993)]. *Let $f:[0, \infty)\to \boldsymbol{R}$ be bounded and twice differentiable with bounded second derivative. Let $t_n \to \infty$ and $f(t_n)$ converge to f^∞ or f_∞ for $n\to\infty$. Then*

$$f'(t_n)\to 0, \qquad n\to\infty .$$

Proof of Theorem 1: Let $\mathscr{R}_0 < 1$. Choose a sequence $t_n\to\infty$ such that

$$I(t_n)\to I^\infty, \qquad \frac{d}{dt}I(t_n)\to 0 .$$

Using the equation for $\frac{d}{dt}I$ in (1.1) and Lemma 1 we have

$$I^\infty \leqq \frac{k}{\mu + r_2 + d} L^\infty . \tag{1}$$

Similarly, choosing a sequence $s_n\to\infty$ such that

$$L(s_n)\to L^\infty, \qquad \frac{d}{dt}L(s_n)\to 0 ,$$

and using the equation for $\frac{d}{dt}L$ in (1.1) we get

$$\begin{aligned} 0 &\leqq \beta c I^\infty - (\mu + k + r_1) L^\infty \\ &\leqq (\mu + k + r_1)(\mathscr{R}_0 - 1) L^\infty . \end{aligned}$$

This implies that $L^\infty \leqq 0$ (since $\mathscr{R}_0 < 1$). But since $L_\infty \geqq 0$, we have that $L^\infty = L_\infty = 0$, and

$$L(t)\to 0, \qquad t\to\infty .$$

By (1) we also have that

$$I(t)\to 0, \qquad t\to\infty .$$

Adding equations in (1.1) gives

$$\frac{d}{dt}N = \Lambda - \mu N - dI \, .$$

Using Lemma 1 we have

$$N_\infty \geqq \frac{1}{\mu}(\Lambda - dI^\infty) = \frac{\Lambda}{\mu} \, .$$

Note that $\frac{d}{dt}N(t) < 0$ for $N > \Lambda/\mu$. Hence, without loss of generality, we can consider only solutions of (1.1) with $N(t) \leqq \Lambda/\mu$ (it is easy to check that solutions of (1.1) are nonnegative for nonnegative initial data). Then $N^\infty \leqq \frac{\Lambda}{\mu}$. It follows that

$$N_\infty = N^\infty = \frac{\Lambda}{\mu} \, .$$

Hence E^0 is g.a.s. when $\mathscr{R}_0 < 1$.

Let $\mathscr{R}_0 > 1$.

Note that under the assumption $\beta' = \beta$ (1.1) is equivalent to the following system:

$$\begin{aligned} \frac{d}{dt}N &= \Lambda - \mu N - dI \\ \frac{d}{dt}L &= \beta c(N - L - I)\frac{I}{N} - (\mu + k + r_1)L \\ \frac{d}{dt}I &= kL - (\mu + d + r_2)I \, . \end{aligned} \tag{2}$$

The unique endemic equilibrium of (2) is given by $E^* = (N^*, L^*, I^*)$, where

$$N^* = \frac{\alpha \mathscr{R}_0 \Lambda}{dk(\mathscr{R}_0 - 1) + \mu\alpha\mathscr{R}_0}$$

$$L^* = \frac{\mu + d + r_2}{k} I^*$$

$$I^* = \frac{k(\mathscr{R}_0 - 1)}{\alpha \mathscr{R}_0} N^* \, ,$$

and

$$\alpha = \mu + d + r_2 + k \, .$$

Noticing that

$$\frac{(N^* - L^* - I^*)}{N^*} = \frac{1}{\mathscr{R}_0}$$

we can write the Jacobian of (2) J at E^* as

$$J = \begin{pmatrix} -\mu & 0 & -d \\ a(\mathscr{R}_0 - 1) & -(a\mathscr{R}_0 + \mu + r_1 + k) & \frac{\beta c}{\mathscr{R}_0} - a\mathscr{R}_0 \\ 0 & k & -(\mu + d + r_2) \end{pmatrix},$$

where

$$a = \frac{\beta c}{\mathscr{R}_0} \frac{I^*}{N^*} .$$

The characteristic equation is

$$\lambda^3 + A\lambda^2 + B\lambda + C = 0 ,$$

where

$$A = a\mathscr{R}_0 + 3\mu + k + r_1 + r_2 + d ,$$

$$B = a\mathscr{R}_0(2\mu + k + r_2 + d) + \mu(2\mu + k + r_1 + r_2 + d) ,$$

$$C = \mu a\mathscr{R}_0(\mu + k + r_2 + d) + kad(\mathscr{R}_0 - 1) .$$

It is clear that $A, C > 0$. Since it can be easily checked that

$$AB > C ,$$

then the Routh-Hurwitz stability conditions are satisfied. It follows that E^* is l.a.s..

Proof of Theorem 2: (a) Let $\mathscr{R}_1 < 1$, $\mathscr{R}_2 < 1$. By Lemma 1 and the $\frac{d}{dt}x_2$, $\frac{d}{dt}y_2$ equations in (2.5) we have

$$\begin{aligned} x_2^\infty &\leqq R_{1b}x_1^\infty , \qquad x_{2\infty} \geqq R_{1b}x_{1\infty} \\ y_2^\infty &\leqq R_{2a}y_1^\infty , \qquad y_{2\infty} \geqq R_{2a}y_{1\infty} \end{aligned} \tag{3}$$

Using the $\frac{d}{dt}y_1$ equation in (2.5) and choosing $t_n \to \infty$ such that

$$y_1(t_n) \to y_1^\infty , \qquad \frac{d}{dt}y_1(t_n) \to 0, \qquad t \to \infty ,$$

we get

$$0 \leqq -(\mu + k')y_1^\infty + \beta^* c(1 - x_2 - y_1 - y_2)^\infty y_2^\infty .$$

Using (3) and (2.6) we get

$$0 \leqq y_1^\infty(\mathscr{R}_2 - 1) ,$$

and it is shown that $y_1^\infty \leqq 0$ since $\mathscr{R}_2 < 1$. As $y_1^\infty \geqq 0$, we have that $y_1^\infty = 0$. The inequalities in (3) also imply that $y_2^\infty = 0$. Similarly, using (2.6), the equation for $\frac{d}{dt}x_1$ in (2.5), and the inequalities in (3) we conclude that

$$0 \leqq x_2^\infty(\mathscr{R}_1 - 1) .$$

Since $\mathscr{R}_1 < 1$, we have that $x_2^\infty \leqq 0$. It follows that $x_1^\infty = x_2^\infty = 0$. Hence

$$\lim_{t\to\infty} x_1(t) = \lim_{t\to\infty} x_2(t) = \lim_{t\to\infty} y_1(t) = \lim_{t\to\infty} y_2(t) = 0 .$$

(b) Let $\mathscr{R}_1 > 1$. Denote the corresponding equilibrium E_4 for (2.5) by $(\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2)$. Recall that

$$R_{1a} = \frac{\beta c}{\mu + k + r_1}, \qquad R_{1b} = \frac{k}{\mu + d + r_2},$$

Then

$$(\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2) = \left(\frac{e(\mathscr{R}_1 - 1)}{R_{1b}}\right), \qquad e(\mathscr{R}_1 - 1), 0, 0),$$

where e is given by (2.4). The eigenvalues of the Jacobian of (2.5) at E_4 are given by the following two equations (keeping in mind that $q = 0$):

$$\begin{aligned} &\lambda^2 + (\beta c \bar{x}_2 + 2\mu + k + r_1 + r_2)\lambda + \beta c \bar{x}_2(\mu + r_2 + k) = 0, \\ &\lambda^2 + (2\mu + k')\lambda + \mu(\mu + k') + k'\beta^* c(\bar{x}_2 - 1) = 0. \end{aligned} \tag{4}$$

The first equation in (4) always has two eigenvalues with negative real parts. Both roots of the second equation in (4) have negative real parts if and only if the constant term is positive which is equivalent to

$$\mathscr{R}_2 < \frac{1}{1 + \frac{(1 - \mathscr{R}_1)}{R_{1a}(1 + R_{1b})}}.$$

Let

$$f(\mathscr{R}_1) = \frac{1}{1 + \frac{(1 - \mathscr{R}_1)}{R_{1a}(1 + R_{1b})}}, \tag{5}$$

it follows that E_4 is l.a.s. if $\mathscr{R}_2 < f(\mathscr{R}_1)$. If $\mathscr{R}_2 > f(\mathscr{R}_1)$, then the second equation in (4) has one root with positive real part, and hence E_4 is unstable. It is easy to show that $f(\mathscr{R}_1) > 1$ for all $\mathscr{R}_1 > 1$ and $f(1) = 1$.

(c) Let $\mathscr{R}_2 > 1$. Note that the corresponding equilibrium E_2 now is

$$(\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2) = \left(0, 0, \left(1 - \frac{1}{\mathscr{R}_2}\right)\frac{\mu}{\mu + k'}, \left(1 - \frac{1}{\mathscr{R}_2}\right)\frac{k'}{\mu + k'}\right).$$

The eigenvalues w of the Jacobian of (2.5) at the correponding equilibrium E_2 are given by the following two equations:

$$\begin{aligned} &w^2 + (\mu + k + r_1 + \mu\mathscr{R}_2 + r_2)w + \alpha = 0, \\ &w^2 + (2\mu + k' + \mu(\mathscr{R}_2 - 1))\, w + \mu(\mu + k')(\mathscr{R}_2 - 1) = 0, \end{aligned} \tag{6}$$

where

$$\alpha = (\mu + k + r_1)(\mu + r_2)(1 - \mathscr{R}_1) + \mu(\mu + r_2)(\mathscr{R}_2 - 1) + k\beta c\left(1 - \frac{1}{\mathscr{R}_2}\right).$$

To simplify expressions we introduce the following two notations

$$R_{2a} = \frac{k'}{\mu + d'}, \qquad q_{1a} = \frac{\mu}{\mu + k + r_1}.$$

In the second equation of (6), since the coefficients are all positive there are always two roots with negative real parts. In the first equation of (6), both roots have negative real parts iff $\alpha > 0$ which is true if $\mathscr{R}_1 < 1$ or if $\mathscr{R}_1 > 1$ and

$$q_{1a}\mathscr{R}_2^2 + (1 - \mathscr{R}_1 - q_{1a} + R_{1a}R_{1b})\mathscr{R}_2 - R_{1a}R_{1b} > 0 \,. \tag{7}$$

As a quadratic function of $\mathscr{R}_2$, the left hand side of (7) has exactly one positive root which we denote by $R_2^+ = g(\mathscr{R}_1)$. Then $\alpha > 0$ iff $\mathscr{R}_1 < 1$ or

$$\mathscr{R}_1 > 1 \quad \text{and} \quad \mathscr{R}_2 > g(\mathscr{R}_1) \,, \tag{8}$$

where

$$g(\mathscr{R}_1) = \frac{1}{2q_{1a}}(\mathscr{R}_1 - 1 + q_{1a} - R_{1a}R_{1b} + \sqrt{(\mathscr{R}_1 - 1 + q_{1a} - R_{1a}R_{1b})^2 + 4q_{1a}R_{1a}R_{1b}}) \,. \tag{9}$$

To check that $g(\mathscr{R}_1) > 1$, let $F(z)$ be the function defined by the left hand side of (7) (as a function of $\mathscr{R}_2$). Then $F(R_2^+) = 0$, and $z < R_2^+$ iff $F(z) < F(R_2^+)$ or $F(z) < 0$. If $\mathscr{R}_1 > 1$, then it is easy to see that $F(1) < 0$. Hence $1 < R_2^+$, i.e., $g(\mathscr{R}_1) > 1$. Then we conclude that E_2 is l.a.s. if $\mathscr{R}_1 < 1$ or if (8) holds.

If $\mathscr{R}_1 > 1$ and $\mathscr{R}_2 < g(\mathscr{R}_1)$, then $\alpha < 0$ and the first equation in (6) has one root with positive real part, and hence E_2 is unstable.

Note that

$$\begin{aligned} F(f(\mathscr{R}_1)) &= R_{1a}(\mathscr{R}_1 - 1)(q_{1a} + q_{1a}R_{1b} + \mathscr{R}_1 - 1 - R_{1a} - R_{1a}R_{1b}) \\ &\leqq \frac{R_{1a}(1 - \mathscr{R}_1)}{(\mu + k + r1)(\mu + r2)}(kr_2(1 + p) + (r_1 + \beta c)(\mu + r_2)) \,, \\ &\leqq 0 \end{aligned}$$

for all $\mathscr{R}_1 > 1$. If follows that $g(\mathscr{R}_1) > f(\mathscr{R}_1)$ for all $\mathscr{R}_1 > 1$. It is easy to see that $g(1) = 1$.

(d) To solve for E_3 (for any $q \geqq 0$), let

$$u = \frac{I}{N}, \qquad v = \frac{J}{N} \,. \tag{10}$$

From the third and the fifth equations in (2.3) we have

$$L_1 = \frac{\mu + d + r_2}{k} I, \qquad L_2 = \frac{\mu + d'}{k'} J.$$

Substituting into the second equation in (2.3) and using (10) we get

$$R_{1a}(1 + R_{1b})u + \left(R_{1a}R_{1b}\left(1 + \frac{1}{R_{2a}}\right) + \frac{\beta^*}{\beta} R_{1a}\right)v + 1 - \mathscr{R}_1 = 0 \,.$$

This gives

$$u = a - bv \,, \tag{11}$$

where

$$a = \frac{\mathscr{R}_1 - 1}{R_{1a}(1 + R_{1b})}, \qquad b = \frac{R_{1b}(1 + \frac{1}{R_{2a}}) + \frac{\beta^*}{\beta}}{(1 + R_{1b})} \,. \tag{12}$$

Substituting into the fourth equation in (2.3) we get

$$Av^2 + Bv + C = 0 \,, \tag{13}$$

where

$$A = 1 + \frac{1}{R_{2a}} - b \,,$$

$$B = \frac{1}{\mathscr{R}_2} + \frac{bqr_2}{\beta^* c} + a - 1 \,,$$

$$C = -\frac{aqr_2}{\beta^* c} \,.$$

The positive real roots v^* of (13), if there are any, are given by

$$v^* = \frac{1}{2A}(-B \pm \sqrt{B^2 - 4AC}) \,. \tag{14}$$

E_3 can be expressed (for any $q \geqq 0$) as

$$N^* = \frac{\Lambda}{\mu + du^* + d'v^*} \,,$$

$$L_1^* = \frac{\mu + d + r_2}{k} u^* N^* \,,$$

$$I^* = u^* N^* \,, \tag{15}$$

$$L_2^* = \frac{\mu + d'}{k'} v^* N^* \,,$$

$$J^* = v^* N^* \,,$$

where v^*, u^* are given by (14) and (11).

We next look at the feasibility of E_3 in the case $q = 0$. E_3 is feasible iff $u^* > 0$ and $v^* > 0$. Note that $q = 0$ implies $C = 0$. Hence

$$v^* = -\frac{B}{A} \,.$$

The equivalent condition for $v^* > 0$ is that A and B have opposite signs which turns out to be if the following condition is satisfied:

$$f(\mathscr{R}_1) < \mathscr{R}_2 < \frac{\beta c}{\mu} \,. \tag{16}$$

Here we have used the fact that $f(\mathscr{R}_1) < \frac{\beta c}{\mu}$. Expression (16) implies that $A > 0$ and $B < 0$. By (11) the equivalent condition for $u^* > 0$ is that $a/b > v^*$ which is satisfied when the inequality (7) or (8) change direction, i.e., when

$$\mathscr{R}_2 < g(\mathscr{R}_1) \,. \tag{17}$$

It can be checked that $g(\mathscr{R}_1) = R_2^+ < \frac{\beta c}{\mu}$. Then (16) and (17) imply that $f(\mathscr{R}_1) < \mathscr{R}_2 < g(\mathscr{R}_1)$, which also implies by part (c) that $\mathscr{R}_1 > 1$. It follows that E_3 exists iff

$$\mathscr{R}_1 > 1, \qquad f(\mathscr{R}_1) < \mathscr{R}_2 < g(\mathscr{R}_1)\,. \tag{18}$$

At the quasi-steady state, $x_2 = R_{1b}x_1, y_2 = R_{2a}y_1$, System (2.5) reduces to a two dimensional system:

$$\begin{aligned}\frac{d}{dt}x_1 &= (\mu + k + r_1)(\mathscr{R}_1 - 1)x_1 - \beta c(R_{1b} + R_{1b}^2)x_1^2 \\ &\quad - (\beta c R_{1b}(1 + R_{2a}) + \beta^* c R_{2a})x_1 y_1 \\ \frac{d}{dt}y_1 &= (\mu + k')\mathscr{R}_2\left(\left(1 - \frac{1}{\mathscr{R}_2}\right)y_1 - R_{1b}x_1 y_1 - (1 + R_{2a})y_1^2\right).\end{aligned} \tag{19}$$

The positive equilibrium of (19) corresponding to E_3 is

$$x_1 = \frac{u^*}{R_{1b}}, \qquad y_1 = \frac{v^*}{R_{2a}}\,. \tag{20}$$

One can show that the Jacobian of (19) at (20) has two eigenvalues with negative real part(s) iff

$$\mathscr{R}_2 < \frac{\beta c}{\mu}\,. \tag{21}$$

Notice that $\mathscr{R}_2 < R_2^+ < \frac{\beta c}{\mu}$. The local stability follows.

This finishes the proof.

Remark. Functions $f(\mathscr{R}_1)$ and $g(\mathscr{R}_1)$ can be written in different forms depending on which parameter(s) we want to vary. For example if we want β to be a varying parameter with other parameters fixed, we can write f and g as the following (noticing that $\mathscr{R}_1$ is a function of β):

$$f(\mathscr{R}_1(\beta)) = \frac{1}{1 + (1 - \mathscr{R}_2(\beta))((1 + R_{1b})(\frac{\mathscr{R}_1(\beta)}{R_{1b}}) - \frac{pr_2}{\mu + k + r_1}))^{-1}},$$

$$g(\mathscr{R}_1(\beta)) = \frac{1}{2q_{1a}}(-l + \sqrt{l^2 + 4q_{1a}(\mathscr{R}_1(\beta) - n)})\,,$$

where

$$l = \frac{\mu(k + r_1) + r_1 r_2 + (1 - p)kr_2}{(\mu + k + r_1)(\mu + r_2)}\,,$$

$$n = \frac{kpr_2}{(\mu + k + r_1)(\mu + r_2)}\,.$$

Proof of Theorem 3: (i) is true since the proof of part (a) in Theorem 2 is valid for any $q \geqq 0$.

Notice that q does not appear in both the expression of E_2 and the Jacobian of (2.5) at E_2. Also q is not involved in the proof of theorem 2(c). (ii) follows immediately.

For part (iii), recall that v^* is the solution of (13). As $C < 0$ there exists a unique positive real root v^* iff $A > 0$, which (after some algebra) is equivalent to

$$\mathscr{R}_2 < \frac{\beta c}{\mu} . \tag{22}$$

As $bv^* > 0$, a necessary condition for $u^* > 0$ is that $a > 0$, i.e., that

$$\mathscr{R}_1 > 1 . \tag{23}$$

Let $G(z) = Az^2 + Bz + C$ (a function given by the left hand side of (13)), then $G(v^*) = 0$. Since G is a parabola with $A > 0$, and since v^* is the only positive root, $u^* > 0$ (or $a/b > v^*$) iff

$$G\left(\frac{a}{b}\right) > G(v^*) = 0 .$$

This is equivalent to

$$\left(1 + \frac{1}{R_{2a}}\right)\frac{a}{b} + \frac{1}{\mathscr{R}_2} - 1 > 0 . \tag{24}$$

After some algebra we can see that (24) is equivalent to

$$q_{1a}\mathscr{R}_2^2 + (1 - \mathscr{R}_1 - q_{1a} + R_{1a}R_{1b})\mathscr{R}_2 - R_{1a}R_{1b} < 0 . \tag{25}$$

Note that (25) has just the opposite direction as that of the inequality (7) which implies that

$$\mathscr{R}_2 < g(\mathscr{R}_1) . \tag{26}$$

Hence $u^* > 0$ iff (23) and (26) hold. Also noticing that $g(\mathscr{R}_1) = R_2^+ < \frac{\beta c}{\mu}$, we see that (26) implies (22), i.e., $u^* > 0$ implies that $v^* > 0$ exists and is unique. Hence E_3 exists iff $\mathscr{R}_1 > 1$ and $\mathscr{R}_2 < g(\mathscr{R}_1)$. The proof of the local asymptotical stability is similar to the one in Theorem 2 (d). This finishes the proof.

References

Anderson, R. M.: Population dynamics of infectious disease. London New York: Chapman and Hall 1982

Anderson, R. M., May, R. M.: Population biology of infectious diseases. Springer-Verlag Berlin Heidelberg New York: 1982

Anderson, R. M., May, R. M.: Infectious Disease of Humans. Oxford New York Tokyo: Oxford University Press 1991

Ayvazian, L. F.: History of tuberculosis. Reichman, L. B.; Hershfield, E. S. (eds) Tuberculosis: A Comprehensive International Approach. pp. 1–20 New York 1993

Bloom, B. R.: Tuberculosis: Pathogenesis, Protection, and Control. ASM Press. Washington, DC, 1994

Bolker, B., Pacala, S. W.: Understanding the ecological implications of spatial pattern formation using ensemble models. To appear. American Naturalist (1995)

Castillo-Chavez, C., Feng, Z.: Mathematical models for the disease dynamics of tuberculosis. Submitted to the Proceedings of the Fourth International Conference on Mathematical Population Dynamics (1996a)

Castillo-Chavez, C., Feng, Z.: Optimal vaccination strategies for TB in age-structure populations. Submitted Journal of Mathematical Biology (1996b)
Castillo-Chavez, C., Hethcote, H. W., Andreason, V. Levin, S. A., Liu, W.: Cross-immunity in the dynamics of homogeneous and heterogeneous populations. In: Gross, L., Hallam, T. G., Levin, S. A. (eds) Mathematical ecology. Proceedings, Autumn Course Research Seminars, Trieste 1986, pp. 303–316. Singapore: World Scientific Publ. Co. 1988
Castillo-Chavez, C., Hethcote, H. W., Andreason, V. Levin, S. A., Liu, W.: Epidemiological models with age structure, proportionate mixing, and cross-immunity. J. Math. Biol. **27**, 233–258 (1989)
Castillo-Chavez, C., Thieme, R. H.: Asymptotically autonomous epidemic models. In: Proc. Third Intern. Conf. on Mathematical Population Dynamics (O. Arino and M. Kimmel, eds), pp. 33–50 (1995)
Castillo-Chavez, C., Velasco-Hernández, J.: On the relationship between tight coevolution and superinfection. Manuscript (1996)
Dietz, K.: Epidemiologic interference of virus populations. J. Math. Biol. **8**, 291–300 (1979)
Dietz, K., Schenzle, D.: Proportionate mixing models for age-dependent infection transmission. J. Math. Biol. **22**, 117–120 (1985)
Durrett, R., Levin, S.: The importance of Being discrete (and spatial). Theoretical Population Biol. **46**, 363–394 (1994)
Feng, Z., Thieme, R. H.: Multi-annual outbreaks of childhood disease revisited: the impact of isolation. **128**, 93–130 Math. Biosci. (1995)
Feng, Z.: A mathematical model for the dynamics of childhood disease under the impact of isolation. Ph. D. Thesis, Arizona State University (1994)
Gao, L. Q., Hethcote, H. W.: Disease transmission models with density-dependent demographics. J. Math. Biol. **22**, 117–120 (1985)
Greenhalgh, D.: Some results for an SEIR epidemic model with density dependence in the death rate. IMA J. Math. Appl. Med. Biol. **9**, 67–106 (1992)
Hadeler, K. P., Castillo-Chavez, C.: A core group model for disease transmission. Math. Biosci. **128**, 41–56 (1995)
Hethcote, H. W.: Qualitative analysis for communicable disease models. Math. Biosci. **28**, 335–356 (1976)
Hethcote, H. W., Stech, H. W., van den Driessche, P.: Periodicity and stability in epidemic models: a survey. In: Busenberg, S., Cooke, K. L. (eds.) Differential equations and applications in ecology, epidemics and population problems, pp. 65–82. New York: Academic Press 1981
Hethcote, H. W., Van Ark, J. W.: Epidemiological models for heterogenous populations: proportionate mixing, parameter estimation and immunization programs. Math. Biosci., **84**, 85–118 (1987)
Hopewell, P. C.: Overview of clinical tuberculosis. Bloom B. R. (ed.) Tuberculosis: Pathogenesis, protection, and control, pp. 25–46 (1994)
Kent, J. H.: The epidemiology of multi-drug resistant tuberculosis in the United States. Medical Clinics of North America. **77**, 1–6 (1991)
Kochi, A.: The Global Tuberculosis Situation and the New Control Strategy of the World Health Organization. Tubercle **72**, 1–6 (1991)
Kolata, G.: First documented case of TB passed on airliner is reported by the U.S.. New York Times. March 3, 1995
May, R. M., Nowak, M. A.: Superinfection, Metapopulation dynamics, and the evolution of diversity. J. Theor. Biol. (1994)
Mena-Lorca, J., Velasco-Hernández, J., Castillo-Chavez, C.: Superinfection, virulence and density-dependent mortality in an epidemic model. BU-1299 Biometrics Unit, Cornell University. (1995)
Miller, Bess.: Preventive Therapy for Tuberculosis. Medical Clinics of North America. **77**, 1263–1275 (1993)
MMWR: Morbidity and mortality weekly report. **44** (8), 137–154 (1995)
Nowak, M. A., May, R. M.: Superinfection and the evolution of parasite virulence. Proc. R. Soc. Lond. **B255**, 81–87 (1994)
Press release WHO/89: TB Epidemic Outpacing New Drug Development. 17 Nov. 1994

Pugliese, A.: An SEI epidemic model with varying population size. In: Busenberg, S. and Martelli, M., (eds.) Differential equation models in biology, epidemiology and ecology. Lecture notes in biomathematics. **92**, Springer-Verlag, Berlin, Heidelberg, New York, 121–138, (1991)

Reichman, L. B., Hershfield, E. S. (editors): Tuberculosis: A Comprehensive International Approach. Lung Biology in Health and Disease, Volume **66**, Marcel Dekker, Inc. New York, Basel, Hong Kong (1993)

Salyers, A. A., Whitt, D. D.: Bacterial Pathogenesis. Washington, D.C. 1994

Schenzle, D.: An age-structured model of pre- and post-vaccination measles transmission. IMA J. Math. Appl. Med. Biol. **1**, 169–191 (1984)

Smith, P. G., Moss, A. R.: Epidemiology of tuberculosis. Bloom B. R. (ed.) Tuberculosis: Pathogenesis, protection, and control, pp. 47–59 (1994)

Styblo, K.: Selected Papers: Epidemiology of Tuberculosis. Royal Netherlands Tuberculosis Association. **24**, 55–62 (1991)

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

Thieme, R. H.: Convergence results and a Poincaré-Bendixson tricho-tomy for asymptotical autonomous differential equations. J. Math. Biol. **30**, 755–763 (1992)

Thieme, R. H.: Asymptotical autonomous differential equations in the plane. Rocky Mountain J. Math. **24**, 351–380 (1994)