Modeling Tuberculosis Contact Tracing and the Role of Social Cluster

Xiaxia Xu[1], Baojun Song[2]

[1] Department of Applied Mathematics, Xi'an Jiaotong University, Xi'an, China

[2] Department of Mathematical Sciences, Montclair State University, Montclair, NJ

Abstract

We propose a discrete model to investigate the role of contact tracing program reducing the newcases and prevalence of tuberculosis. We observe that the tuberculosis contact tracing program has no effect on the basic reproduction number R_0 but the size of social cluster has. On the other hand a contact tracing program can speed up the process of TB elimination. We compute the partial rank correlation coefficient (PRCC), based on Latin hypercube sampling (LHS), to evaluate the effect of input parameters on the magnitude of the newcases and prevalence. The most influential parameters are identified and ranked.

keywords: tuberculosis, discrete model, TB contact tracing, control strategy

1 Introduction

Tuberculosis (TB) is an infectious bacterial disease caused by Mycobacterium tuberculosis, which most commonly affects the lungs. The disease can also attack any part of the body such as the kidney, spine, and brain. If not treated or treat improperly, TB disease can be fatal.

The disease is transmitted from person to person via droplets from the throat and lungs of people with the active respiratory disease[16, 20]. In most people who breathe in TB bacteria and become infected, the body is able to fight the bacteria to stop them from growing at the beginning, we call this latent TB period. People with latent TB infection do not feel sick, do not have any symptoms, and cannot spread the disease. However, active TB individuals are able to spread the disease to others [2, 20]. It is well known that TB can not be spread by shaking hands, sharing food or drink, touching bed linens or toilet seats, sharing toothbrushes or kissing [20].

As an ancient disease, TB remains one of the major causes of disability and death worldwide. In 2010, an estimated 8.8 million people fell ill with TB. Of the 1.4 million deaths, 95 percent occurred in developing countries[17]. Figure 1 shows the new cases of active TB per year in the U.S.A from 1953 to 2011, data taken from [18, 19], which continues to exhibit a downward trend. The annual case of TB has been declining steadily but raise sightly in the 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[2].

There are some global research communities committed to finding new ways to better understand, diagnose, treat, and ultimately prevent TB. Contact tracing with subsequent treatment of individuals latently infected with Mycobacterium tuberculosis is a cornerstone of tuberculosis control in the United States and other low burden countries. In epidemiology, contact tracing is defined as the identification and diagnosis of people who may have come into contact with an infected person. This is the method that has been used to control endemic contagious diseases for decades. A disease investigation begins when an individual is identified as having a communicable disease. An investigator interviews the patient, family members, physicians, nurses, and anyone else who may have knowledge of the primary patient's contacts, anyone who might have been exposed, and anyone who might have been the source of the disease. Then the contacts are screened to see if they have or have ever had the disease.

It is well known that smallpox is one of the few infectious diseases that was conquered by human beings. However, the final elimination of smallpox was not attributed to the smallpox vaccine, but to contact tracing. Smallpox could be controlled only because the sores and scars prevented infected persons from escaping detection. Fellow villagers and tribesmen were encouraged in various ways to identify infected persons. When a person with smallpox was identified, he or she was quarantined, and all the persons in the surrounding community or village were vaccinated. In this way smallpox was eventually reduced to isolated outbreaks and then eradicated.

A disease that is spread by respiratory contact, such as tuberculosis, may require

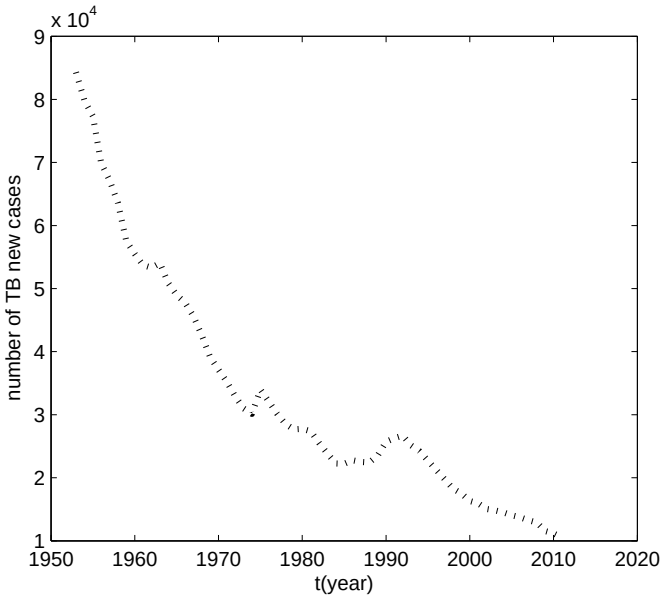

Figure 1: *Tuberculosis new cases in USA from 1953 to 2011.*

screening tens to hundreds of persons. Can TB contact tracing eventually eradicate this long-standing infectious disease, as in the case of smallpox? There are at least two epidemiological factors that distinguish smallpox and TB. First, latent TB is asymptomatic. The other is that latent TB has a much longer latent period than smallpox. We suspect there is a long way to go for the elimination of TB by using contact tracing (if even possible). This research will quantify the TB contact tracing program, then use mathematical models to test this hypothesis, and ultimately to compare this strategy with other ones.

2 Discrete-time TB model with contact tracing

One efficient way to exercise TB contact tracing is to look into the social cluster where a great deal of intimate contacts takes place. To treat latent TB, we begin with tracing the contacts within the social cluster after identifying an active TB. A social cluster of an individual may include the co-workers, family members and acquaintances.

We develop a social-cluster model in a discrete-time scale via extending the previous work in [1, 5] in order to quantify the practice of TB contact tracing. Social clusters are divided into two types: TB active clusters and TB inactive clusters. Each TB active cluster contains one active TB, while TB inactive cluster does not. Let S_1 and E_1 be the respective total number of susceptible individuals and latent TB in the TB active cluster, and let S_2 and E_2 be the respective total number of susceptible individuals and latent TB in the TB inactive clusters; Let I be the total number of active TB. $N_1 = S_1 + E_1$ is the total number of people in TB active clusters, while $N_2 = S_2 + E_2$ is the total number of people in TB inactive clusters (as shown in Table 1). We assume that each social cluster includes exactly the same number of n individuals on average. When a latent TB individual becomes infectious, her/his social cluster moves to TB active cluster.

Let p be the survival rate for populations in the absence of the disease. Latently-infected survivals of a TB inactive cluster, are assumed to develop active TB at the rate kpE_2. Each new infectious individual 'activates' a new cluster and in the process 'moves' $(n-1)kpE_2$ individuals from TB inactive clusters into TB active cluster ones. It is assumed that a fraction $\frac{S_2}{N_2}$ of these individuals goes to the S_1 class per time step while the last

Table 1: *Model variables. Each TB active cluster contains one active TB.*

Variables	Description
S_1	total number of susceptible individuals in the TB active cluster
E_1	total number of latent-TB individuals in the TB active cluster
I	active TB
S_2	total number of susceptible individuals in the TB inactive cluster
E_2	total number of latent-TB individuals in the TB inactive cluster
$N_1 = S_1 + E_1$	total number of people in TB active cluster
$N_2 = S_2 + E_2$	total number of people in TB inactive cluster
$N = N_1 + N_2 + I$	total number of population

fraction $\frac{E_2}{N_2}$ goes to E_1 class per time step. The number of people who do not change their status of social cluster in the S_2 subpopulation is counted by

$$pS_2(t) - (n-1)kpE_2(t)\frac{S_2(t)}{N_2(t)} = p(S_2(t) + (k+1-kn)E_2(t))\frac{S_2(t)}{N_2(t)}.$$

Similarly, the number of people who do not change their status of social cluster in the E_2 subpopulation is

$$(1-k)pE_2(t) - (n-1)kpE_2(t)\frac{E_2(t)}{N_2(t)} = p((1-k)N_2(t) - (n-1)kE_2(t))\frac{E_2(t)}{N_2(t)}.$$

Furthermore, since infectious individuals recover or die (at rate $\gamma I(t)$) then the rate at which TB active clusters become inactive is $\gamma I(t)$. Bookkeeping requires that recovery (or cluster dissolution) must be accompanied by the returned of cluster members to the

TB inactive cluster population. It is assumed that $(n-1)\gamma pI(t)\frac{S_1(t)}{N_1(t)}$ individuals are returned to S_2 and $(n-1)\gamma pI(t)\frac{E_1(t)}{N_1(t)}$ are returned to E_2 respectively. The identity $N_1(t) = (n-1)I(t)$ reduces the last expression for the flow from $S_1(t)$ to $S_2(t)$ into $\gamma pS_1(t)$ and that from E_1 to E_2 into $\gamma pE_2(t)$. The (assumed) low prevalence of active TB implies that $N_1 << N_2$. We use this observation to neglect births into the TB active cluster population. TB progression in the E_1 class is extremely unlikely because of the short average life of TB active clusters. The above considerations lead to the following basic discrete time model:

$$\begin{cases} S_1(t+1) = (1-\beta-\gamma)pS_1(t) + (n-1)kpE_2(t)\dfrac{S_2(t)}{N_2(t)}, \\ E_1(t+1) = \beta pS_1(t) + (1-\gamma)pE_1(t) + (n-1)kpE_2(t)\dfrac{E_2(t)}{N_2(t)}, \\ I(t+1) = kpE_2(t) + (1-\gamma)I(t), \\ S_2(t+1) = \Lambda + \gamma pS_1(t) + p(S_2(t) + (k+1-kn)E_2(t))\dfrac{S_2(t)}{N_2(t)}, \\ E_2(t+1) = \gamma pE_1(t) + p((1-k)N_2(t) - (n-1)kE_2(t))\dfrac{E_2(t)}{N_2(t)}. \end{cases} \tag{1}$$

It can be checked that the evolution of the total population ($N(t) = S_1(t) + E_1(t) + I(t) + S_2(t) + E_2(t)$) is governed by

$$N(t+1) = \Lambda + pN(t) + (1-\gamma-p)I(t).$$

Recall that $1-\gamma$ is the survival rate of the active-TB individuals, we have that $1-\gamma \leq p$. If there is no treatment and no the TB-induced death, then $1-\gamma = p$, that means the survival proportion of the infectious is the same as no-infectious. Then $N(t+1) = \Lambda + pN(t)$ has a closed and simplest form in population dynamics. We note $\beta + \gamma < 1$ and $nk < 1$

guarantee all terms in model (1) are positive. Therefore, all solutions will remain positive.

TB contact tracing is applied when a TB inactive cluster change into TB active cluster ones. Of $(n-1)kpE_2(t)\frac{E_2(t)}{N_2(t)}$ newly found latent TB, a proportion q of them are treated and the remaining is left untreated (of course, when q=1, we treat all newly identified latent TB). The parameter q contains much more tractable information about treating latent TB, such as the adherence of the antibiotics administration, capability of identifying all members of a social cluster, and vulnerability of members in social clusters. Therefore, TB contact tracing campaign treat $q(n-1)kpE_2(t)\frac{E_2(t)}{N_2(t)}$ of latent TB cases in active social cluster per unit time. Model (1) equipped with TB contact tracing now has form:

$$\begin{cases} S_1(t+1) = (1-\beta-\gamma)pS_1(t) + (n-1)kpE_2(t)\dfrac{S_2(t)}{N_2(t)} + q(n-1)kpE_2(t)\dfrac{E_2(t)}{N_2(t)}, \\ E_1(t+1) = \beta pS_1(t) + (1-\gamma)pE_1(t) + (1-q)(n-1)kpE_2(t)\dfrac{E_2(t)}{N_2(t)}, \\ I(t+1) = kpE_2(t) + (1-\gamma)I(t), \\ S_2(t+1) = \Lambda + \gamma pS_1(t) + p(S_2(t) + (k+1-kn)E_2(t))\dfrac{S_2(t)}{N_2(t)}, \\ E_2(t+1) = \gamma pE_1(t) + p((1-k)N_2(t) - (n-1)kE_2(t))\dfrac{E_2(t)}{N_2(t)}. \end{cases} \tag{2}$$

The size of social cluster n and parameter q are the key players in analyzing model (2). We will focus on simulating model (2) based on our analytic work on model (1). The impact of cluster size n and treatment parameter q to the TB contact tracing campaign should be quantified from this model.

3 Model analysis

3.1 The basic reproduction number and stability

To calculate the basic reproduction number of the model, the next generation operate approach can't be applied [6]. We compute the reproduction number by considering the stability of disease-free equilibrium.

There always exists the disease-free equilibrium $e_0 = (0, 0, 0, \frac{\Lambda}{1-p}, 0)$ for model (2). Next we study the stability of e_0. The Jacobi matrix evaluated at e_0 is

$$J|_{e_0} = \begin{bmatrix} (1-\beta-\gamma)p & 0 & 0 & 0 & (n-1)kp \\ \beta p & (1-\gamma)p & 0 & 0 & 0 \\ 0 & 0 & 1-\gamma & 0 & kp \\ \gamma p & 0 & 0 & p & -(n-1)kp \\ 0 & \gamma p & 0 & 0 & (1-k)p \end{bmatrix}. \tag{3}$$

The characteristic equation is

$$[\lambda-(1-\gamma)](\lambda-p)\left\{[\lambda-(1-\beta-\gamma)p][\lambda-(1-\gamma)p][\lambda-(1-k)p]-(n-1)k\gamma\beta p^3\right\}=0. \tag{4}$$

Clearly, we have two eigenvalues $\lambda_1 = 1-\gamma$ and $\lambda_2 = p$, both of them satisfy $|\lambda_i| < 1, i = 1, 2$. The other eigenvalues satisfy the equation

$$F(\lambda) = [\lambda-(1-\beta-\gamma)p][\lambda-(1-\gamma)p][\lambda-(1-k)p]-(n-1)k\gamma\beta p^3 = 0. \tag{5}$$

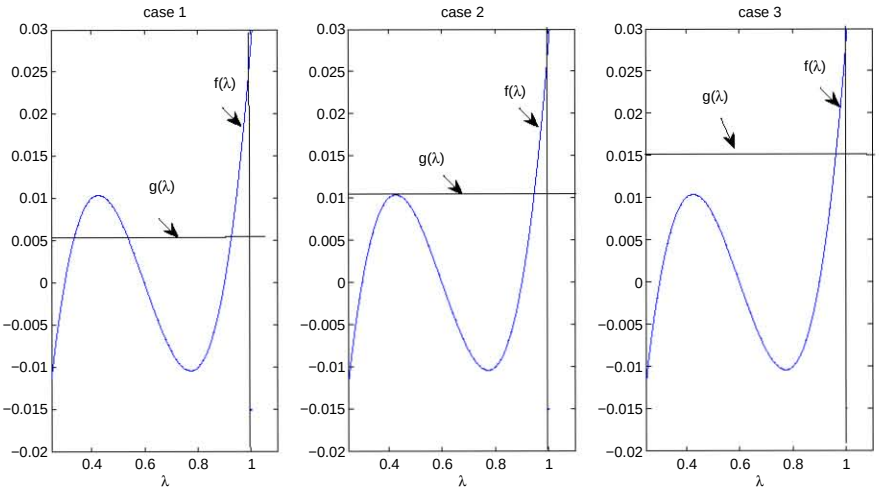

Figure 2: Three cases for eigenvalues. Case 1: three different real eigenvalues; Case 2: a real double root happens; Case 3: a pair of imaginary roots happens.

Let

$$f(\lambda) = [\lambda - (1-\beta-\gamma)p][\lambda - (1-\gamma)p][\lambda - (1-k)p],$$

and

$$g(\lambda) = (n-1)k\gamma\beta p^3.$$

We observe that $f(\lambda) = 0$ has three roots $\tilde{\lambda}_1 = (1-\beta-\gamma)p$, $\tilde{\lambda}_2 = (1-\gamma)p$ and $\tilde{\lambda}_3 = (1-k)p$, and $|\tilde{\lambda}_i| < 1$, $i = 1, 2, 3$ satisfied. The eigenvalues are the point where the two equations $f(\lambda)$ and $g(\lambda)$ intersect. Three cases may happen, see Figure 2 for details. From Figure 2, when $g(\lambda) = (n-1)k\gamma\beta p^3 < f(1)$ is satisfy, the absolute values of $\lambda_i, i = 3, 4, 5$ are less than unity. That is

$$(n-1)k\gamma\beta p^3 < [1-(1-\beta-\gamma)p][1-(1-\gamma)p][1-(1-k)p],$$

which can be rewritten as

$$\frac{\beta p(n-1)}{1-(1-\beta-\gamma)p} \cdot \frac{\gamma p}{1-(1-\gamma)p} \cdot \frac{kp}{1-(1-k)p} < 1.$$

Therefore we define

$$R_0 = \left(\frac{\beta p(n-1)}{1-(1-\beta-\gamma)p}\right) \cdot \left(\frac{\gamma p}{1-(1-\gamma)p}\right) \cdot \left(\frac{kp}{1-(1-k)p}\right).$$

Each term in R_0 has clear epidemiological interpretation. The first term $\frac{\beta p(n-1)}{1-(1-\beta-\gamma)}$ represents the number of infections produced by one infectious individual [5] as follows:

When one inactive latent TB progresses to active TB, all the rest $n-1$ members in his/her social cluster move to the active cluster. Thus, the dynamic satisfies

$$S_1(t+1) = (1-\beta-\gamma)pS_1(t),$$

$$E_1(t+1) = \beta p S_1(t).$$

with the initial conditions $S_1(0) = n-1, E_1(0) = 0$. Thus,

$$S_1(t) = (n-1)[(1-\beta-\gamma)p]^t,$$

The number of infections produced by one active TB during his/her lifetime is

$$\sum_{i=1}^{\infty} E_1(i) = \beta p \sum_{i=0}^{\infty} S_1(i) = \lim_{i\to\infty} \frac{(n-1)\beta p(1-(1-\beta-\gamma)^i p^i)}{1-(1-\beta-\gamma)p} = \frac{(n-1)\beta p}{1-(1-\beta-\gamma)p}.$$

The second term $\frac{\gamma p}{1-(1-\gamma)p}$ stands for the fraction of people who progress from active latent TB to inactive latent TB. The last fraction of R_0 $\frac{kp}{1-(1-k)p}$ represents the fraction of infected people who develop active-TB during their lifetime.

Following the above analysis, we have the theorem.

Theorem: The disease-free equilibrium e_0 of model (2) is locally asymptotically stable if

$R_0 < 1$ and unstable if $R_0 > 1$.

In this section we considered the stability of the disease-free equilibrium, and got the expression of the basic reproduction number R_0. From the expression, we note that the contact tracing rate q has no effect on R_0. This means that we can not eliminate the disease if $R_0 > 1$.

3.2 The endemic equilibrium

If there is no treatment and no the TB-induced death, then $1 - \gamma = p$, that means the survival proportion of the infectious is the same as no-infectious. Then

$$N(t+1) = \Lambda + pN(t), \tag{6}$$

$N^* = \Lambda/(1-p)$ is a global attractor for (6). Using limiting equations $\lim_{t\to\infty} N(t) = N^*$, we reduce the five dimensional system (2) into four dimensional one

$$\begin{cases} E_1(t+1) = \beta p(N^* - E_1(t) - I(t) - S_2(t) - E_2(t)) + (1-\gamma)pE_1(t) + (n-1)kpE_2(t)\dfrac{E_2(t)}{N_2(t)}, \\ I(t+1) = kpE_2(t) + (1-\gamma)I(t), \\ S_2(t+1) = \Lambda + \gamma p(N^* - E_1(t) - I(t) - S_2(t) - E_2(t)) + p(S_2(t) + (k+1-kn)E_2(t))\dfrac{S_2(t)}{N_2(t)}, \\ E_2(t+1) = \gamma pE_1(t) + p((1-k)N_2(t) - (n-1)kE_2(t))\dfrac{E_2(t)}{N_2(t)}. \end{cases}$$

First of all, we consider about the epidemic equilibrium while $q = 0$ as follows:

$$\begin{cases} E_1 + E_2 = \beta p(N^* - E_1 - I - S_2 - E_2) + pE_1 + (1-k)pE_2, & (7) \\ I = kpE_2 + (1-\gamma)I, & (8) \\ S_2 = \Lambda + \gamma p(N^* - E_1 - I - S_2 - E_2) + pS_2 - (n-1)kp\dfrac{E_2 S_2}{N_2}, & (9) \\ E_2 = \gamma pE_1 + p(1-k)E_2 - (n-1)kp\dfrac{E_2^2}{N_2}. & (10) \end{cases}$$

From (8), we have

$$E_2 = \frac{\gamma}{kp} I. \quad (11)$$

From (7) we obtain

$$(1 - p + \beta p)E_1 + [1 - (1-k)p + \beta p]E_2 = \beta p(N^* - S_2 - I).$$

Rewrite it we get

$$S_2 = (N^* - I) - \frac{1 - p + \beta p}{\beta p} E_1 - \frac{1 - (1-k)p + \beta p}{\beta p} E_2, \quad (12)$$

$$N_2 = (N^* - I) - \frac{1 - p + \beta p}{\beta p} E_1 - \frac{1 - (1-k)p}{\beta p} E_2. \quad (13)$$

Substituting (13) into (10),

$$\begin{aligned} (n-1)kpE_2^2 =& [\gamma pE_1 - (1 - p(1-k))E_2] \left((N^* - I) - \frac{1 - p + \beta p}{\beta p} E_1 - \frac{1 - (1-k)p}{\beta p} E_2 \right) \\ =& -\gamma p \frac{1 - p + \beta p}{\beta p} E_1^2 + \left(\gamma p(N^* - I) + \frac{1 - p + \beta p - \gamma p}{\beta p} (1 - (1-k)p) E_2 \right) E_1 \\ &+ \frac{(1 - p(1-k))^2}{\beta p} E_2^2 - (N^* - I)(1 - p(1-k))E_2. \quad (14) \end{aligned}$$

Substituting (11) into (14), we get an equation of variables E_1 and I,

$$\begin{aligned} 0 =& \gamma p \frac{1 - p + \beta p}{\beta p} E_1^2 - \left(\gamma p(N^* - I) + \frac{1 - p + \beta p - \gamma p}{\beta p} (1 - (1-k)p) \frac{\gamma}{kp} I \right) E_1 \\ &+ \left((n-1)kp - \frac{(1 - p(1-k))^2}{\beta p} \right) (\frac{\gamma}{kp} I)^2 + (N^* - I)(1 - p(1-k)) \frac{\gamma}{kp} I. \quad (15) \end{aligned}$$

From (9) we have

$$(1-p+\gamma p)S_2 + (n-1)kpE_2\frac{S_2}{N_2} = \Lambda + \gamma p(N^* - I - E_1 - E_2). \quad (16)$$

After submit (11),(12) and (13) into (16), we obtain

$$A_1(I)E_1^2 + A_2(I)E_1 + A_3(I) = 0, \quad (17)$$

where

$$\begin{aligned}
A_1(I) =& [(1-p)^2 + p(\beta+\gamma)(1-p)](1-p+\beta p),\\
A_2(I) =& \{(1-p+\beta p)\left[2(1-(1-k)p) + \beta p(1-(n-1)kp-\gamma p)\right] - \beta p\gamma p(1-(1-k)p)\}\frac{\gamma I}{kp}\\
&+ [\beta p\gamma p(1-p+2\beta p) - 2\beta p(1-p+\beta p)](N^*-I) + \beta p(1-p+\beta p)\Lambda,\\
A_3(I) =& (1-p+\gamma p)\left(\left[(1-(1-k)p)\frac{\gamma I}{kp} - \beta p(N^*-I)\right]^2 + \beta p(1-(1-k)p)(\frac{\gamma I}{kp})^2\right)\\
&- \beta^2p^2(1-p+\gamma p)\left((N^*-I)\frac{\gamma I}{kp}\right)\\
&+ k\beta^2p^3(n-1)(N^*-I)\frac{\gamma I}{kp} - (n-1)k\beta p^2(1-(1-k-\beta)p)(\frac{\gamma I}{kp})^2\\
&- \gamma\beta p^2(1-(1-k)p)(\frac{\gamma I}{kp})^2 + \gamma\beta p^2(1-(1-k-\beta)p)(N^*-I)\frac{\gamma I}{kp}\\
&+ \beta p\Lambda(1-(1-k)p)\frac{\gamma I}{kp} - \beta^2p^2(N^*-I)(\Lambda+\gamma p(N^*-I)).
\end{aligned}$$

By calculation of the two equations of (15) and (17) we can consider the endemic equilibrium of System (2) in the absence of contact tracing.

4 Numerical simulation

In this section, we will estimate the parameter values using least square method to let the simulation fit the annual reported new cases in Figure 1 as well as we can. Based on

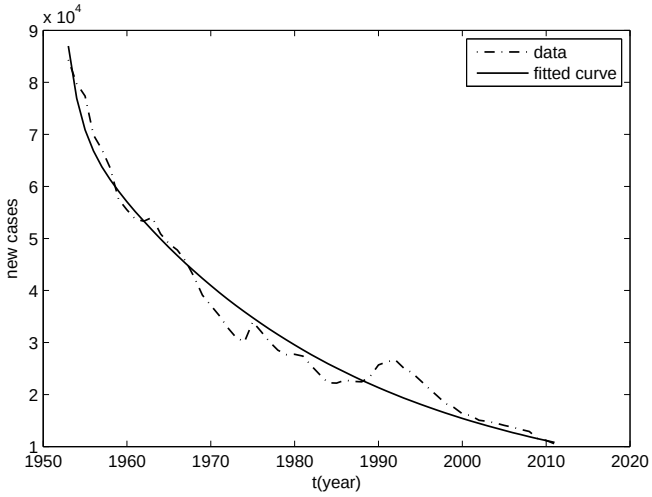

Figure 3: *The best-fit solution obtained by fitting $\beta p S_1(t)$ in model (2) to the reported number of TB new cases in Figure 1.*

the estimation, the impacts of social cluster and the contact tracing on the dynamics of TB can be studied. Finally, we will compute the partial rank correlation coefficient to evaluate the effect of input parameters on the magnitude of TB new cases and prevalence.

4.1 Estimation of parameter values

We estimate the parameter values by using the Least Square Method (LSM) to make sure the simulation curves based on the model fit the reported TB new cases in U.S.A. as well as possible. That is, let

$$h(\Theta) = \Sigma_{t=1953}^{2011} \|\beta p S_1(t) - RN(t)\|^2$$

where Θ is the vector composed of fitting parameters $\beta, \gamma, p, n, k, q, \Lambda$. $RN(t)$ is the number of reported TB cases at the year t as show in Figure 1, $\beta p S_1(t)$ is the solution of TB new

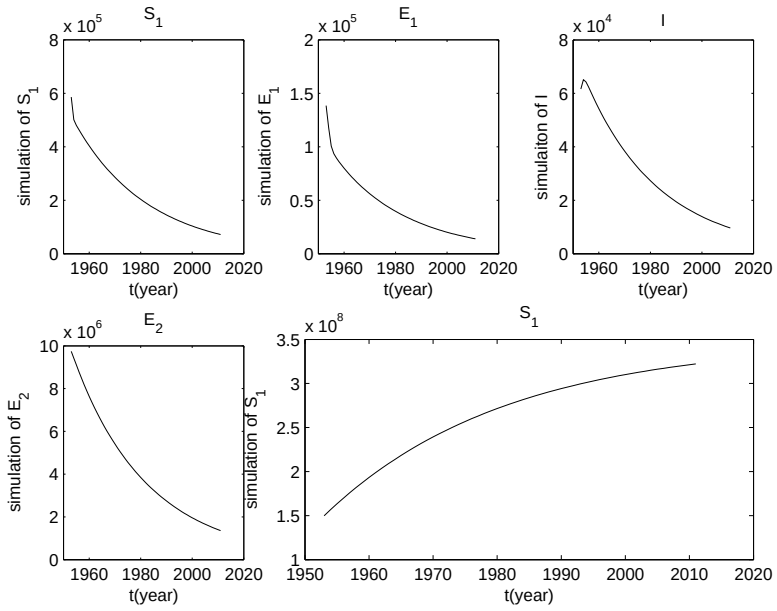

Figure 4: *Behavior of model (2) based on the parameter estimation.*

cases in system (2). Then the values for the parameters are estimated such that $h(\Theta)$ achieves its minimum. The estimated parameters are shown in Table 2 and the best fitted curve is shown in Figure 3. The simulations for other variables of model (2) are shown in Figure 4. From the website [15] we know the total population of U.S.A. is about $3.12 * 10^8$, the total number of population in our simulation fits the reported number well. From Figure 4, we also know that the disease in U.S.A. is decreasing.

4.2 Social cluster size impact on R_0

The basic reproductive number R_0 is the expected number of secondary infectious cases generated by an average infected individual during the infective period in an entirely

Table 2: *Model parameters and the estimated values ($R_0 = 0.1750$).*

Parameter s	Description	Estimation
β	effective transmission rate	0.15
γ	removed rate of infectious individual	0.7750
p	survival rate for susceptible and latent TB	0.9649
n	average size of social cluster	10
k	progression rate from inactive latent TB to active TB	0.0055
q	effective treatment rate of active latent TB	0.9
Λ	recruitment	$1.22 * 10^7$

susceptible population. This quantity determines the potential for an infectious agent to start an outbreak, and the extention of transmission in the absence of control measures. From the above estimation, we know the basic reproductive number is estimated to be $R_0 = 0.1750$ for the TB transmission in U.S.A, which means the disease is going to extinction. However, we are still interested in the variance in R_0 caused by the varied control parameters, to try to decrease the secondary infections and thus the new infection each year.

Figure 5 shows the contour plot of R_0 versus the removed rate out of active TB γ and the size of social clusters n. It indicates that the basic reproduction number decreases as the number of social clusters n declines. For a fixed social clusters number n, the basic reproduction number R_0 will firstly increase and then decrease as the remove rate γ varies

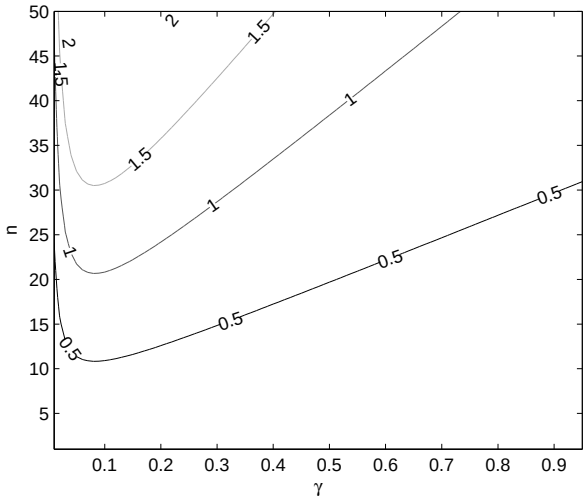

Figure 5: *Contour plot of R_0 vs. removed rate of active TB γ and the size of social clusters n.*

from 0 to 1. This shows that increasing the treatment rate of active TB, which leads to an increase in γ, is not always helpful to decrease the new infections, it can sometimes exacerbate the disease transmission. From this figure we can also give the critical values for the remove rate out of active TB γ for various number of social cluster n to keep the basic reproduction number below unity. For example, when the number of social cluster n is fixed at 30, to keep $R_0 < 1$, the remove rate γ should be less than 0.02 or bigger than 0.328.

4.3 Contact tracing impact on TB dynamic

From the expression of the basic reproduction number R_0, we know that the contact tracing has no effect on R_0, which means that we can not eliminate the disease by implementing

the contact tracing strategy when the disease is an epidemic originally. To investigate how the contact tracing affects the disease dynamics, we consider the numerical numbers of the latent active TB according to different levels of the contact tracing.

Figure 6 shows the number of active latent TB (E_1) with q varying from the baseline value to its 90% and 110%. Figure 6(a) indicates that the disease declines faster with an increased q. If we define a critical value for E_1 as the elimination value of the disease (for example, while the number of E_1 is less than 1, then the disease is eliminated), then the simulation with a larger q can reach the elimination level earlier than the simulation with a smaller q, we omit plots here. That means the strategy with a stronger contact tracing (q is larger) can eliminate the disease earlier than others. In other words, the contact tracing can speed up the process of TB elimination. Figure 6(b) shows that the steady state size of the active latent TB decreases greatly when the contact tracing p increases. This indicates that the contact tracing is an effective control measure even though it can not help to eliminate the disease.

In a summary, by implementing the strategy of contact tracing, we can speed up the process of TB elimination, or we can effectively control the total infection numbers.

4.4 Sensitivity and uncertainty analysis

In this section, an uncertainty and sensitivity analysis, based on the Latin Hypercube Sampling (LHS) scheme ([8, 9, 11, 14]), are carried out to examine which parameter is sensitive to the disease spread or which measure is most effective to lower the TB

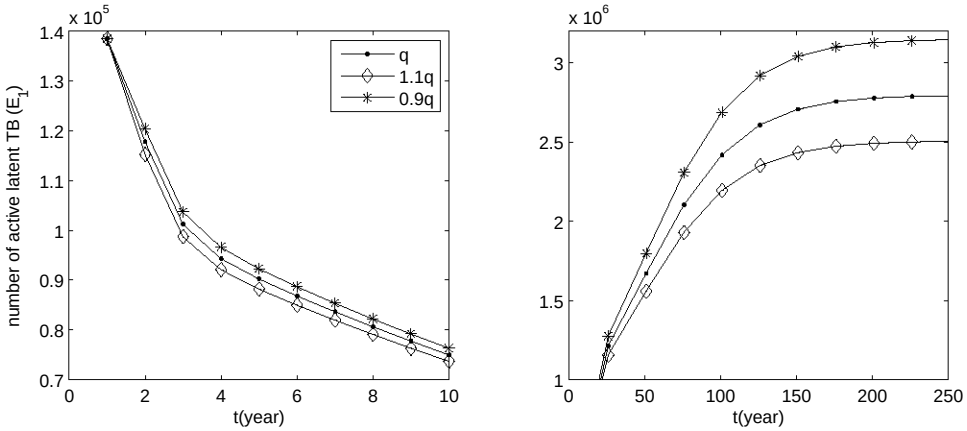

Figure 6: *Contact tracing affects on the disease while* $n = 10$ *(a) and* $n = 100$ *(b). All other parameters are as shown in Table 2.*

incidence and prevalence. The LHS scheme is an extremely efficient scheme that enables the exploration of the entire parameter space of the model with a minimum number of computer simulations.

4.4.1 Preparation for sensitivity and uncertainty analysis

In the absence of available data on the distribution of parameters β, γ, n, q, we use normal distribution as in [14]. We assume the distribution of parameters p and k are a triangular shape. Let the values of parameters in Table 2 be the mean value of parameters, or be the peaking values for the triangular distributed parameters, respectively. Secondly we select appropriate variances or minimum and maximum values for the distribution. The details are as follows. The nature death rate equals to 1/60 year^{-1} in [5], 0.0189 year^{-1} in [12] and 0.01 year^{-1} in [3], respectively. Thus the survival rate equals to 0.9833, 0.9811

and 0.99 correspondingly. Here we let the survival rate p be triangular peaking at 0.9649, which is also the minimum value for the simulations. The progression rate from latent inactive TB to active TB k is beween 0.00256 and 0.0016 in [2] and [1]. The progression rate k has a triangular distribution peaking at 0.0055 and the minimum value is 0.0016. See details in Table 3.

The LHS-matrix is then built by assembling 2000 samples from each parameter's distribution. Each row of the LHS-matrix is a combination of parameter values. So, 2000 solutions are then simulated [11, 14].

4.4.2 Results of sensitivity analysis

A sensitivity analysis, carried out by estimating the partial rank correlation coefficients (PRCC) for each input parameter and each outcome variable, can identify which parameters are important in contributing to the variability of outcomes ([7, 8, 10, 11]). The sigh of PRCC indicates the qualitative relationship between each input variable and each output variable, that is, the positive value of the PRCC for the majority of the variable implies that when the value of the input variable increases, the future number of the outcomes increase. The magnitude indicates the contribution of input parameter to the prediction of the value of outcome variable.

To know whether the contribution of any input parameter to the outcomes vary over an entire time interval during model dynamics, PRCC values are calculated for multiple time points and plotted versus time. This allows us to assess whether the contribution of

one parameter is obvious over an entire time interval during the progression of the model dynamics [14]. Figures 7(a) and 7(b) show PRCC values for new cases and prevalence plotted over 10 years, respectively. There are two PRCC values (the progression rate k and the survival rate p) for both new cases and prevalence are significantly different from zero. This may conclude that new cases and prevalence are quite sensitive to k and p. We can observe from Figure 7(a) that the PRCC values of the survival rate p varies much while the rest of others vary little over time. Comparing the two figures we know the PRCC for prevalence varies much more than that for new cases, especially at the very beginning of the simulation.

Figure 8 focuses on the PRCC values between the input parameters and the two outcomes (new cases and prevalence) 10 years later (at the year of 2021). The PRCC values together with the p-values are calculated as shown in Table 3. If p-value is small, say no more than 0.05, then the correlation between the input parameter and the outcomes is statistically significant at the level 0.05. The correlations between biological behavioral transmission parameters p, n, k and the two outcomes are statistically significant at the level 0.0001 ($p - NC < 0.0001$), and that between γ and the two outcomes are significant at the level 0.01. From Table 3, the correlation between q and the outcomes are not significant ($p - NC > 0.1$).

The PRCC value between the progression rate from inactive latent TB to active TB k and the new cases (equals to 0.9314) is larger than the ones between other parameters and new cases, followed by the PRCC between the survival rate p and new cases (equals

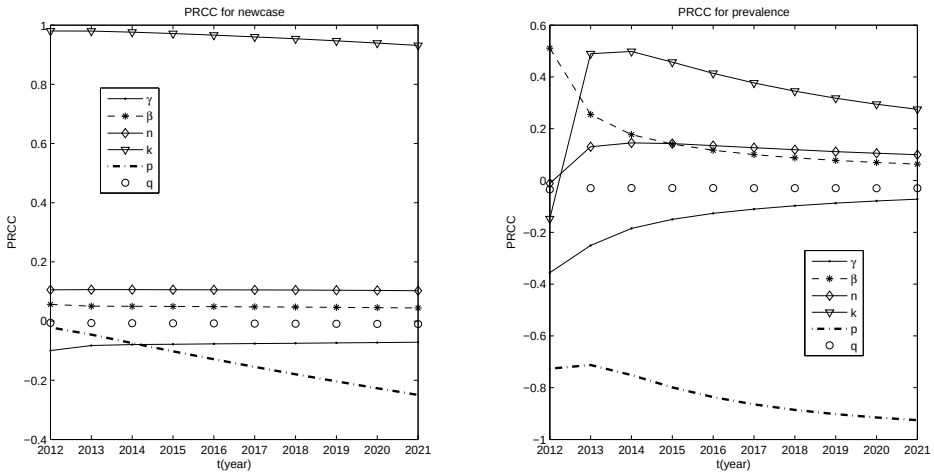

Figure 7: *PRCC of the parameters for new cases (left) and prevalence (right) over 10 years. All the parameters came from LHS.*

to -0.2495). This implies that the progression rate k contributes the most to new cases. Thus an increase in k will greatly increase new cases, and an increase in p will lead to an obvious decrease in new cases. From Figure 8(b), the PRCC value between the survival rate p and the prevalence (equals to -0.9253) is largest among all the PRCCs, then followed by PRCC between the progression rate k and prevalence (equals to 0.2748). So that the prevalence is sensitive to p and k. An increase in p and a decrease in k can both lead to a obvious decline in prevalence. From Figure 8 and Table 3, decreasing the cluster size n can also effectively reduce new infections and prevalence (the PRCCs between n and new cases equals to 0.1027, between n and prevalence equals to 0.1027).

Table 3: *PRCC between each parameter and the two outcomes: new cases (NC) and prevalence (Prev) 10 years later (at the year of 2021). We denote p-NC and p-Prev as zeros while they are smaller than 0.0001. If p-NC and p-Prev are small, say no more than 0.05, then the correlation between the input parameter and the outcome variables are significant.*

Para	Min	Max	Function shape	PRCC (NC)	p-NC	PRCC (Prev)	p-Prev
γ	0.7750*0.95	0.7750*1.05	normal (mean 0.7750)	-0.0712	0.0014	-0.0720	0.0013
β	0.15*0.95	0.15*1.05	normal (mean 0.15)	0.0444	0.0473	0.0633	0.0046
n	10*0.95	10*1.05	normal (mean 10)	0.1027	0	0.0997	0
k	0.0016	0.0055	triangular (peak at 0.0055)	0.9314	0	0.2748	0
p	0.9649	0.99	triangular (peak at 0.9649)	-0.2495	0	-0.9253	0
q	0.9*0.95	0.9*1.05	normal (mean 0.9)	-0.0100	0.6544	-0.0300	0.1792

4.4.3 Results of uncertainty analysis

A Latin Hypercube Sampling (LHS) uncertainty analysis is performed to explore the variability of the outcomes due to the uncertainty in estimating the input parameters [10, 11]. This uncertainty analysis technique enables the degree of prediction imprecision to be quantified. In this subsection, we explore the uncertainty in estimating the values of the input six variables on the prediction precision of the two outcomes: the new TB cases and the prevalence of TB at the year of 2021 in U.S.A. The frequency distributions for the two outcome variables are derived from the results of the uncertainty analysis. The frequency distribution of new cases is skewed slightly to the right (the skewness equals to -0.3249) and the frequency distribution of prevalence is skewed to the left (the skewness equals to 0.4418) (see Figure 9). The descriptive statistics for the two distributions are shown in Table 4. The maximum (equals to 9729 for new cases, equals to 0.3116% for prevalence) and minimum (equals to 2626 for new cases, equals to 0.2943% for prevalence) of the two distributions reflect the likely ranges of possible outcomes, rather than the absolute upper ($\Lambda/(1-p)$) and lower (0) bounds of the system; These show that at the year 2021, at least 2626 new TB infections will be produced and the prevalence will be lager than 0.2943%. The 90% confidence interval for new cases is 3707 to 8384, this prediction imprecision is due to the uncertainty in estimating the values of the input parameters. The 90% confidence interval for prevalence is 0.2969% to 0.3078%, which indicates about 3 persons living with active TB per 1000 persons in the year of 2021. The prediction precision of the prevalence is fairly high because of the narrow distribution interval of the

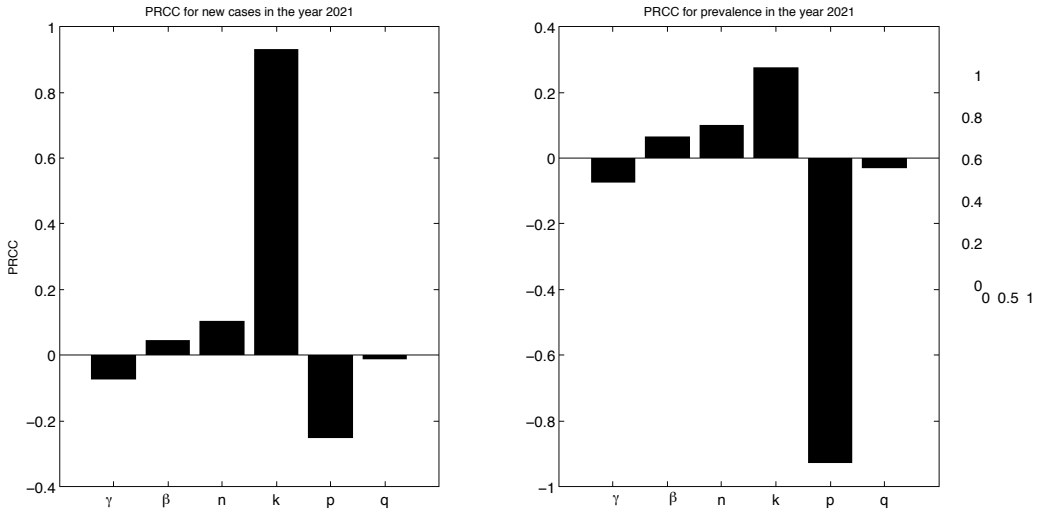

Figure 8: *PRCC for new cases (left) and prevalence (right) when t=2021. All the parameters came from LHS.*

prevalence.

5 Conclusion and discussion

We studied a discrete model considering the TB contact tracing and social cluster. From the expression of R_0, we know that the basic reproduction number has no relationship with effective treatment rate of active latent TB q. Therefore the contact tracing cannot help to eliminate the active TB when $R_0 > 1$. However, the contact tracing can speed up the process of TB elimination when $R_0 < 1$ and reduce the epidemic size when $R_0 > 1$ (see Figure 6). By implementing contact tracing, both the new infections and the prevalence of the disease can be reduced from the results of sensitivity analysis. In fact,

Table 4: *Descriptive statistics from the uncertainty analysis*

Statistic	new cases	prevalence(%)
Minimum	2626	0.2943
Maximum	9729	0.3116
Mean	6305	0.3017
Median	6515	0.3013
Variance	$2.09 * 10^6$	$1.12 * 10^{-5}$
Standard	$1.45 * 10^3$	0.0033
5th percentile	3707	0.2969
95th percentile	8384	0.3078
Skewness	-0.3249	0.4418

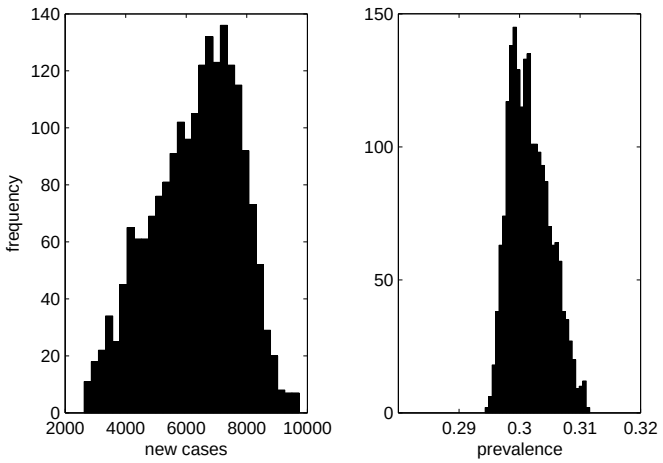

Figure 9: *The frequency distribution of new cases and prevalence of U.S.A, after 10 years, produced by 2000 numerical simulations in the LHS/PRCC sensitivity analysis.*

comparing with the parameters, the contributions of contact tracing to both the new infections and the prevalence of the disease vary slightly during the whole progression of the disease transmission. Thus, the effect of contact tracing on disease transmission depends less on time. The contribution of the social cluster size to the two outcomes, new cases and prevalence, also depends less on time; However, the correlations between them are statistically significant as shown in the sensitivity analysis. The sensitivity analysis also indicates that decreasing the social cluster size can effectively reduce the new cases and prevalence. The contour plot of R_0 (Figure 5) illustrates that TB infection can be reduced by cutting down the communication between the active TB and the surrounding community (R_0 decrease as the social cluster size decrease).

Additionally, from the sensitivity analysis, new infections are very sensitive to the removed rate from active latent TB to inactive latent TB and to the survival rate. The result is consistent with the sensitivity results in [13], where they got the same conclusion that the disease transmission is quite sensitive to the progression rate from latent TB to active TB. They also got the result that their model is quite sensitive to the transmission probability. However, in our TB model, the transmission probability contributes not so much to the disease.

Acknowledgements

We would like to thank Dr. Carlos Castillo-Chavez, Executive Director of the Mathematical and Theoretical Biology Institute (MTBI), for giving us the opportunity to participate in

this research program. We would also like to thank Co-Executive Summer Directors Dr. Erika T. Camacho and Dr. Stephen Wirkus for their efforts in planning and executing the day to day activities of MTBI. We also want to give special thanks to José Manuel Vega-Guzmán for valuable discussion and comments. This research was conducted in MTBI at the Mathematical, Computational and Modeling Sciences Center (MCMSC) at Arizona State University (ASU). This project has been partially supported by grants from the National Science Foundation (NSF - Grant DMPS-0838705), the National Security Agency (NSA - Grant H98230-11-1-0211), the Office of the President of ASU, and the Office of the Provost of ASU.

References

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

[2] C. Castillo-Chavez, B. Song. Dynamical models of tuberculosis and their applications, *Math. Bios. Eng.* 1(2004)361-404.

[3] C. P. Bhunu, W. Garira, Z. Mukandavire, M. Zimba. Tuberculosis transmission model with chemoprophylaxis and treatment, *Bull. Math. Bio.* 70(2008)1163–1191.

[4] H. Cao, Y. Zhou, B. Song. Complex dynamics of discrete SEIS model with simple demography, *Disc. Dyn. Nat. Soc.* (2011)doi:10. 1155/2011/653937.

[5] J. P. Aparicio, A. F. Capurro, C. Castillo-Chavez. Transmission and Dynamics of Tuberculosis on Generalized Households, *J. Theor. Biol.* 206(2000)327–341.

[6] L. J. S. Allen, P. van den Driessche. The basic reproduction number in some discrete-time epidemic models, *J. Diff. Equa. Appl.* 14(2008)1127–1147.

[7] L. Rong, Z. Feng, A. S. Perelson. Emergence of HIV-1 Drug Resistance During Antiretroviral Treatment, *Bull. Math.Biol.* 69(2007)2027–2060.

[8] M. A. Sanchez, S. M. Brower. Uncertainty and Sensitivity Analysis of the Basic Reproductive Rate, *Amer. J. Epid.* 145(1997)1127–1137.

[9] S. Marino, I. B. Hogue, C. J. Ray, D. E. Kirschner. A methodology for performing global uncertainty and sensitivity analysis in systems biology, *J. Theor. Biol.* 254(2008)178–196.

[10] S. M. Blower, D. Hartel, H. Dowlatabadi, R. M. Anderson and R. M. May. Drugs, Sex and HIV: A Mathematical Model for New York City. *Phil. Trans. R. Soc. Lond. B* 321(1991)171–187.

[11] S. M. Blower, H. Dowlatabadi. Sensitivity and Uncertainty Analysis of Complex Model of Disease Transmission: An HIV Model, as an example, *Inter. Stat. Review.* 62(1994)229–243.

[12] S. Whang, S. Choi, E. Jung. A dynamic model for tuberculosis transmission and optimal treatment strategies in South Korea, *J. Theor. Bio.* 279(2011)120–132.

[13] T. Cohen, C. Colijn, B. Finklea and M. Murry. Exogenous re-infection and the dynamics of tuberculosis epidemics: local effects in a network model of transmission, *J. R. Soc. Interface.* 4(2007)523–531.

[14] X. Wang, Y. Xiao, J. Wang, X. Lu. A mathematical model of effects of environmental contamination and presence of volunteers on hospital infections in China, *J. Theor. Bio.* 293(2012)161–173.

[15] http://geography.about.com/od/obtainpopulationdata/a/uspopulation.htm, Accessed July 21, 2012.

[16] http://www.who.int/topics/tuberculosis/en/ Accessed July 16, 2012.

[17] http://www.niaid.nih.gov/topics/tuberculosis/Pages/default.aspx, Accessed July 15, 2012.

[18] http://www.cdc.gov/tb/statistics/default.htm, Accessed July 21, 2012.

[19] http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6111a2.htm, Accessed July 21, 2012.

[20] http://www.cdc.gov/tb/topic/basics/default.htm, Accessed July 15, 2012.