**Spora: A Journal of Biomathematics**

Volume 5 | Issue 1 Article 2

2019

# Combating Tuberculosis: Using Time-Dependent Sensitivity Analysis to Develop Strategies for Treatment and Prevention

Kendall B. Clark
*Northwestern University*, kendallclark2023@u.northwestern.edu

Mayleen Cortez
*California State University, Channel Islands*, mayleen.cortez136@myci.csuci.edu

Cristian Hernandez
*Regis University*, 3hcristian@gmail.com

Beth E. Thomas
*St. Mary's College of Maryland*, bethomas@smcm.edu

Allison L. Lewis
*Lafayette College*, lewisall@lafayette.edu

# Combating Tuberculosis: Using Time-Dependent Sensitivity Analysis to Develop Strategies for Treatment and Prevention

**Cover Page Footnote**
This research was supported by grant number DSM-1560301 from the National Science Foundation, and was conducted during the Research Experience for Undergraduates Emerging Scholars Program at St. Mary's College of Maryland in the summer of 2018.

Intercollegiate Biomathematics Alliance

SPORA
A Journal of Biomathematics

ISSN 2473-3067 (Print)
ISSN 2473-5493 (Online)

# Combating Tuberculosis: Using Time-Dependent Sensitivity Analysis to Develop Strategies for Treatment and Prevention

Kendall B. Clark[1], Mayleen Cortez[2], Cristian Hernandez[3], Beth E. Thomas[4], Allison L. Lewis[5,*]

*Correspondence:
Dr. Allison L. Lewis
Dept. of Mathematics
Lafayette College
730 High St.
Easton, PA 18042, USA
lewisall@lafayette.edu

**Abstract**

Although many organizations throughout the world have worked tirelessly to control tuberculosis (TB) epidemics, no country has yet been able to eradicate the disease completely. We present two compartmental models representing the spread of a TB epidemic through a population. The first is a general TB model; the second is an adaptation for regions in which HIV is prevalent, accounting for the effects of TB/HIV co-infection. Using active subspaces, we conduct time-dependent sensitivity analysis on both models to explore the significance of certain parameters with respect to the spread of TB. We use the results of this sensitivity analysis to determine the most effective strategies for treatment and prevention throughout the epidemic.

**Keywords:** time-dependent sensitivity analysis, active subspaces, tuberculosis, SEIR, compartmental modeling

## 1 Introduction

Tuberculosis (TB) is a potentially fatal, airborne infection caused by the bacterium *Mycobacterium tuberculosis*. TB typically affects the respiratory system (pulmonary TB), though it may also damage other organ systems (extra-pulmonary TB) [25]. The progression of this infection varies with each case; some individuals become infectious immediately after encountering the bacterium, while others enter an exposed—or latent—state. Those who become immediately infectious (and thus contagious) are classified as having *fast-track* TB. These individuals may be treated via a therapeutic drug regimen. Those who enter the exposed state have *slow-track* TB, and may develop active infections later on in one of two ways: *endogenous reactivation* or *exogenous reinfection*. Endogenous reactivation occurs when a latently infected individual becomes infectious after the TB bacteria inside them is activated. Exogenous reinfection occurs when a latently infected individual becomes infectious after they are externally infected by a new TB bacterium. This is sometimes referred to as a "mixed infection" [18]. However, 90 to 95 percent of latent TB patients remain asymptomatic and noninfectious for life [3]. Exposed individuals may be treated with a regimen of chemoprophylaxis.

Tuberculosis has been kept under relative control in the western world since the beginning of the twentieth century. In 1921, the Bacillus Calmette-Guérin (BCG) vaccine was developed by French medical researchers, Albert Calmette and Camille Guérin. Ideally administered during infancy, the BCG vaccine remains a standard medical practice, although it has only fifty percent efficacy [12].

Despite numerous medical advances, TB has made a recent resurgence in the developing world, specifically in the regions of Sub-Saharan Africa, Eastern Europe, and Asia. New drug resistant strains of *M. tuberculosis*, in addition to the simultaneous rapid spread of human immunodeficiency virus (HIV), prompted the declaration of a public health emergency in 2005. According to the Africa Regional Committee of the World Health Organization, the TB incidence rate has doubled overall in the past fifteen years. Moreover, in areas where HIV is especially prevalent, the rate has tripled and in regions that have been severely affected by both TB and HIV, the rate has quadrupled [26]. This suggests that being HIV+ significantly increases one's susceptibility to tuberculosis, accelerating the spread of an epidemic. This occurs because tuberculosis is an opportunistic infection, meaning that active cases are more likely to develop or increase in severity when the body's immune system is in a weakened state. Accordingly, we choose to focus both on general TB epidemics as well as those that occur in HIV-prevalent regions.

A number of previous studies have been conducted to model the proliferation of tuberculosis in the developing world. Ozcaglar et. al [18] provide an extensive overview

[1]Northwestern University, Evanston, IL, [2]California State University Channel Islands, Camarillo, CA, [3]Regis University, Denver, CO, [5]St. Mary's College of Maryland, St. Mary's City, MD, [4]Department of Mathematics, Lafayette College, Easton, PA

of the work done with regard to the mathematical modeling of TB epidemics. Expanding upon the work described in [18], our investigation not only encompasses the analysis of compartmental models for the spread of TB and its co-infection with HIV, but also introduces the use of time-dependent sensitivity analysis to pinpoint the exact times in an epidemic for which specific treatment or prevention strategies are most effective.

Our goal is to identify which parameters have the most substantial impact upon the size of the exposed and infectious populations, in order to determine any vulnerabilities that could be exploited to drive the epidemic towards eventual eradication. We begin in Section 2 by constructing two Susceptible - Exposed - Infectious - Recovered (SEIR) compartmental models for the spread of a tuberculosis epidemic through a population. The first is a general model for the spread of tuberculosis, while the second accounts for the effects of HIV prevalence within the population. In Section 3, we describe our method for conducting time-dependent sensitivity analysis using active subspace techniques. In Section 4, we apply these techniques to our two compartmental models to determine the most effective strategies for combating tuberculosis at different stages of the epidemic.

# 2 Model Formulation

In this study, we utilize compartmental models to describe the spread of tuberculosis through a population. The goal of a compartmental model is to show how materials flow from one mutually exclusive state to another. Changes in states are often modeled with differential equations, taking into account what enters and leaves each state over time.

Historically, compartmental models play an important role in the mathematical modeling of infectious diseases. In the early twentieth century, William Kermack and Anderson McKendrick introduced a compartmental model called the Susceptible - Infectious - Recovered (SIR) model [10]. SIR models, often referred to as Kermack-McKendrick models, represent the possible states in which an individual can exist during an epidemic.

Since SIR models have mutually exclusive states, members of a population can only exist in one state at a time. However, individuals may move from one compartment to another as their infection status changes due to the epidemic. Basic SIR models account for two types of transitions: the transition from susceptible (S) to infectious (I) and from infectious (I) to recovered (R). The rate at which someone moves from S to I is called the rate of infection. This rate represents individuals moving from S to I after coming into contact with the disease. The rate at which an infected individual moves into the recovered population is referred to as the rate of recovery.

In this study, we use a modified version of the original SIR model: the Susceptible - Exposed - Infectious - Recovered (SEIR) model. The addition of the Exposed state (E) takes into account the proportion of the population with latent TB. As previously stated, when an individual has latent TB, they have been exposed to the *M. tuberculosis* bacteria but are asymptomatic and not contagious.

Here, we introduce two SEIR models. One consists of the aforementioned states and does not account for HIV/TB co-infection. Throughout this paper we will refer to it as our general TB model. The other model is an extension of our general SEIR model and accounts for HIV prevalence within a population. We will refer to this model as our HIV/TB co-infection model.

## 2.1 General Tuberculosis Model

The general TB model is illustrated in Figure 1. Each individual resides in either the susceptible, exposed, infectious, or recovered population at any given time. The pathways and rates of transfer between states are described by the parameters in Table 1. The general TB model includes the following assumptions:

1. Everyone is born susceptible to tuberculosis.
2. Relapse from the recovered state means re-contracting active TB, i.e. a relapsed individual goes straight from recovered to infectious.
3. Drug treatment for both latent and active tuberculosis is available, but treatment is not administered to the entire E and I populations.
4. There are no drug-resistant TB strains present.
5. Vaccines are not 100% effective and may decline in efficacy over time.

The following system of ordinary differential equations defines the instantaneous rates of change of the populations in each compartment. Parameter units are chosen to ensure consistency in the ODE system; all four equations have units of individuals/time.

$$\frac{dS}{dt} = \beta - \lambda SI - \delta S - \nu S \tag{1}$$

$$\frac{dE}{dt} = p\lambda SI - (\omega + \sigma_c + \delta)E \tag{2}$$

$$\frac{dI}{dt} = (1-p)\lambda SI + \omega E - (\delta + \tau + \gamma + \sigma_t)I + \theta R \tag{3}$$

$$\frac{dR}{dt} = \nu S + \sigma_c E + (\sigma_t + \gamma)I - (\delta + \theta)R \tag{4}$$

Equation (1) illustrates the change in the susceptible population with respect to time. As per assumption (1)

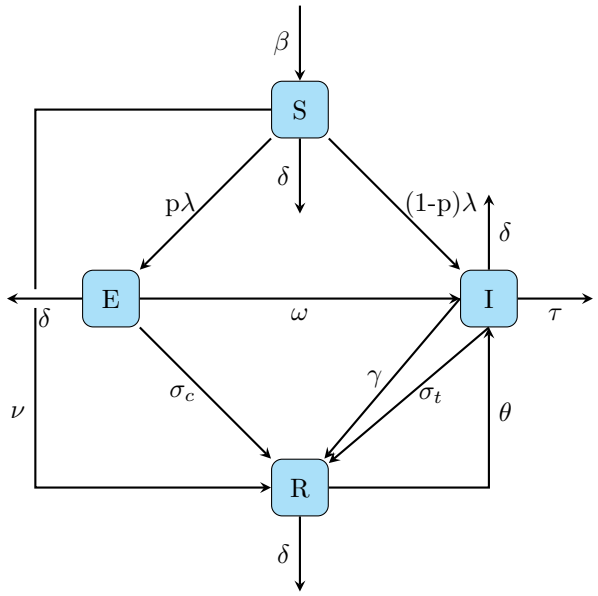


Figure 1: General Tuberculosis SEIR Model.

Table 1: Parameters for General TB Model.

| Parameter | Meaning | Units |
|---|---|---|
| $\beta$ | birth rate | $\frac{\text{individuals}}{\text{time}}$ |
| $\delta$ | natural death rate | $\frac{1}{\text{time}}$ |
| $\nu$ | rate of vaccination | $\frac{1}{\text{time}}$ |
| $\lambda$ | rate of infection | $\frac{1}{\text{individuals}\times\text{time}}$ |
| $\omega$ | rate of deterioration | $\frac{1}{\text{time}}$ |
| $\tau$ | death rate due to TB | $\frac{1}{\text{time}}$ |
| $\gamma$ | rate of recovery | $\frac{1}{\text{time}}$ |
| $\sigma_c$ | latent TB treatment | $\frac{1}{\text{time}}$ |
| $\sigma_t$ | active TB treatment | $\frac{1}{\text{time}}$ |
| $\theta$ | rate of reinfection | $\frac{1}{\text{time}}$ |
| $p$ | slow TB proportion | — |

above, all individuals born into the population, represented by $\beta$, enter the susceptible population. Those individuals who are vaccinated leave the susceptible population, transitioning directly to the recovered population at a rate of $\nu$. We also take into account those who leave the susceptible population after coming into contact with infectious individuals and contracting the disease. A fixed percentage, $p$, of these individuals will develop slow-track TB, while the remainder will move directly to the infectious state. Natural death occurs at the same rate in each of the four populations and is represented by the $\delta$ terms in Equations (1)–(4).

Equation (2) includes the proportion, $p$, of those individuals from the susceptible population that develop slow TB and enter the exposed population. Individuals may leave the exposed population due to deterioration to the active TB infection, natural death, or treatment via chemoprophylaxis, represented by $\omega$, $\delta$, and $\sigma_c$, respectively.

Equation (3) displays an influx of individuals coming from both the susceptible population (via fast TB) and the exposed population (via deterioration of latent TB into the active infection). This equation also shows individuals who are naturally recovering from active TB or recovering via therapeutic treatment and thus leaving the infectious population, represented by rates $\gamma$ and $\sigma_t$, respectively. Here we have two "death" terms; $\tau$ represents people dying due to TB, while $\delta$ indicates death due to natural causes, as in the other three states. Finally, some individuals from the recovered population relapse and re-enter the infectious population at a rate of $\theta$.

Equation (4) models how individuals from the susceptible, exposed and infectious populations move into the recovered population due to vaccination, treatment and natural recovery. Leaving the recovered population are those individuals who are naturally dying and relapsing, as previously discussed. Because TB vaccines are not 100% effective and may decline in efficacy over time [14], we include vaccinated individuals in the recovered class, which allows for them to re-enter the E-I-R cycle via the relapse mechanism.

Figure 2 illustrates the population dynamics of a tuberculosis epidemic over a span of 300 years. This graph is depicted using our established ODE system with estimated parameter values listed in Table 2. We use an initial population of 1000 people, 990 of which are susceptible and 10 exposed. These initial conditions are comparable to those used in the epidemiological model presented in [18]. We choose these values to ensure that our model dynamics illustrate an *epidemic* (defined by the CDC as a greater incidence of disease than what is expected in a region [19]) as opposed to an endemic prevalence of TB. Because we allow for individuals to be born into—and die out of—the population, our system is not

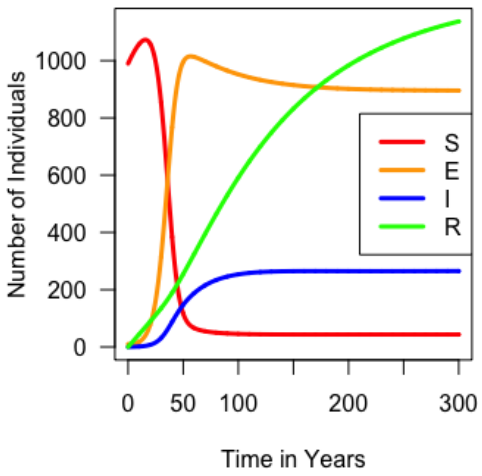


Figure 2: SEIR dynamics for a population with nominal parameter values as listed in Table 2, with initial values $S = 990$, $E = 10$, $I = 0$, and $R = 0$.

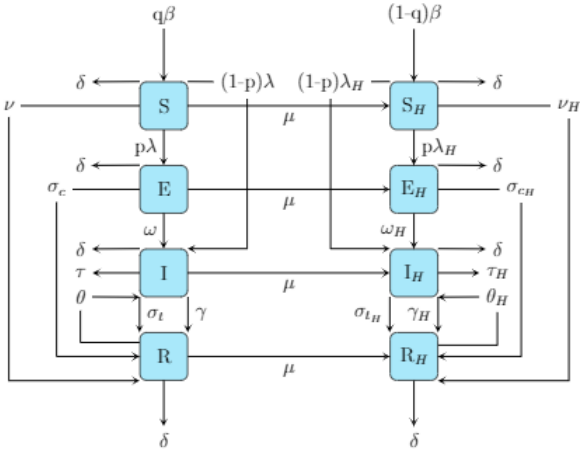


Figure 3: HIV/TB Co-infection Model.

closed. Thus, it is possible for the total population to exceed 1000 at later times in the epidemic. The susceptible population initially thrives then drastically decreases until year 50 when it begins to level off. The exposed population acts in opposition to this, increasing rapidly until year 50. The infectious population also increases and begins leveling off around year 50, however it increases at a lesser rate than the exposed population. The recovered population—which consists of vaccinated individuals as well as those whose TB has been successfully treated—steadily increases in this time frame. In the eventual steady-state, we see that approximately 35% of the population has contracted latent TB, while roughly 10% have active TB. These percentages are in alignment with data from high-TB regions in Africa [27]. All remaining individuals are still susceptible or have moved into the recovered category via either successful treatment of their TB or vaccination.

## 2.2 HIV/TB Co-infection Model

We now present a model that more accurately represents the spread of tuberculosis in regions where HIV is prevalent. This co-infection model has eight compartments, seen in Figure 3, which separate S, E, I, and R into HIV+ and HIV− populations. We also introduce two new parameters. The parameter $\mu$ represents the HIV incidence rate: the rate at which people in each compartment contract HIV and move into the respective HIV+ population. The parameter $q$ represents infants who are born HIV− as a proportion of the birth rate, $\beta$; thus, $(1-q)\beta$ accounts for HIV+ births.

The remainder of the parameters have equivalent interpretations to their counterparts in the general TB model. Parameters subscripted with an $H$, as shown in Table 2, have altered values to account for the effects of a weakened immune system. One additional assumption is added: the HIV incidence rate, $\mu$, remains the same for all four HIV− populations.

As before, the rate of change of each population is represented by a system of ordinary differential equations. Each equation has similar interpretations to the equations for the general TB model. In the system, the term $(I + I_H)$ represents the total number of infectious individuals, whether in the HIV+ or HIV− populations. Note that an infectious person in the HIV− population can infect a susceptible individual in their own population or in the HIV+ population. Similarly, an infectious individual in the HIV+ population can infect a person in their own population or in the HIV− population.

$$\frac{dS}{dt} = q\beta - \lambda(I + I_H)S - (\nu + \delta + \mu)S$$

$$\frac{dS_H}{dt} = (1-q)\beta - \lambda_H(I + I_H)S_H - (\nu_H + \delta)S_H + \mu S$$

$$\frac{dE}{dt} = p\lambda(I + I_H)S - (\delta + \sigma_c + \omega + \mu)E$$

$$\frac{dE_H}{dt} = p\lambda_H(I + I_H)S_H + \mu E - (\delta + \sigma_{c_H} + \omega_H)E_H$$

$$\frac{dI}{dt} = (1-p)\lambda(I + I_H)S + \omega E - (\gamma + \delta + \tau + \sigma_t + \mu)I + \theta R$$

$$\frac{dI_H}{dt} = (1-p)\lambda_H(I + I_H)S_H + \omega_H E_H + \mu I + \theta_H R_H - (\gamma_H + \tau_H + \sigma_{t_H} + \delta)I_H$$

$$\frac{dR}{dt} = (\sigma_t + \gamma)I + \sigma_c E + \nu S - (\delta + \theta + \mu)R$$

$$\frac{dR_H}{dt} = (\sigma_{t_H} + \gamma_H)I_H + \sigma_{c_H} E_H + \nu_H S_H - (\delta + \theta_H - \mu)R_H$$

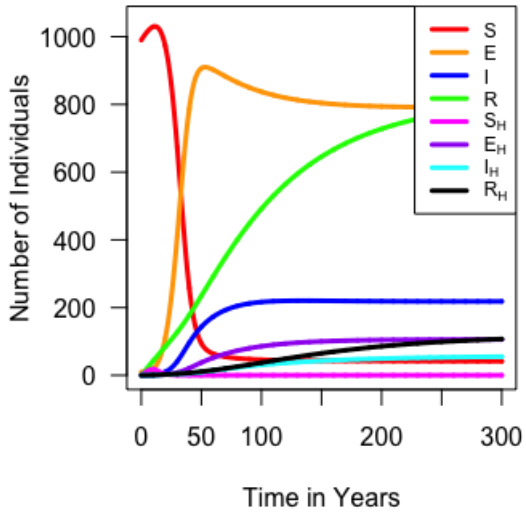


Figure 4: HIV/TB Co-infection Population Dynamics.

Figure 4 illustrates the dynamics of the eight populations in the HIV/TB co-infection model. At the steady state, the total percentage of HIV+ individuals is 12.86%. For comparison, HIV prevalence among adults aged 15–49 in Nigeria, South Africa, and Swaziland is 3.2%, 18.1%, and 27.4%, respectively [27].

# 3 Sensitivity Analysis Using Active Subspaces

Our objective is to identify which parameters, or combinations of parameters, contribute most to the fluctuations in exposed and infected individuals, to identify vulnerabilities in the system that would allow us to drive that number toward zero with appropriate treatment and prevention strategies. To perform this sensitivity analysis, we use techniques derived from the construction of active subspaces.

## 3.1 Active Subspace Construction

Active subspaces can be used in sensitivity analysis to identify key directions in the input space that are most influential in changing a quantity of interest, $f$. An active subspace identifies a set of orthogonal directions—or vectors—in the input space, each of which is a set of weights that define a linear combination of the original parameters. The first vector in the active subspace represents the direction along which $f$ changes the most; relative importance of the direction decreases as you move along the columns of the active subspace. Those directions which are deemed unimportant, or noninfluential to the quantity of interest, are discarded. In this way, active subspaces are a powerful tool for dimension reduction; here, we will use them as a means of identifying influential parameters or sets of parameters for our sensitivity analysis.

Constantine [5] outlines the process for identifying an active subspace via Monte Carlo estimation of the eigendecomposition of the uncentered covariance matrix. He goes on to state that this is equivalent to analyzing the singular value decomposition (SVD) of a matrix of gradient vectors, collected at a set of $M$ Monte Carlo parameter samples. We utilize this second method, and outline it briefly in Algorithm 1.

There are a number of methods for choosing where to partition $S$. Some rely on visual gaps in the singular value spectrum [5, 13]; others choose the dimension in order to satisfy a user-defined error tolerance [7, 23]. Depending on the desired use for the active subspace, a balance must be struck between minimizing the dimension to reduce computational needs and maximizing the accuracy of the estimate. Here we utilize the "gap-based" method, choosing the dimension of the active subspace to reflect an observed gap in the singular values. The partition of $S$ determines the dimension of the active subspace; that is, the number of singular values that are deemed significant determines the number of directions that are included in the active subspace.

## 3.2 Obtaining Sensitivity Metrics from Active Subspaces

Our primary use for the active subspace in this investigation is to construct a global sensitivity metric, which—like similar procedures such as Morris Screening [17] or the construction of Partial Rank Correlation Coefficients (PRCC) [15]—allows us to rank parameters in terms of their overall influence on the quantity of interest. However, the advantage to using active subspaces as an intermediary step to constructing the sensitivity ranking is that information regarding linear combinations of significant parameters may be obtained; that is, a parameter may not be flagged as significant on its own, but may be extremely significant when observed in combination with another parameter. In high-dimensional models where significant dimension reduction must occur before analysis can proceed, this information can be utilized to redefine the quantity of interest over key directions of influence, reducing the dimension of the model while still minimizing the information loss from discarded parameters [5].

After identifying the dimension, $n$, of the active subspace, we can compute *activity scores* for each of the $m$ parameters, which weight each parameter's contributions from the important directions in the input space accord-

**Algorithm 1** Estimating the Active Subspace

(1) Draw $M$ parameter samples $\mathbf{x}_j = [x_1, \ldots, x_m]$, for $j = 1, \ldots, M$ and $m$ the number of parameters, independently from a specified density function. Here, we use

$$\mathcal{U}^m(\mathbf{x}^* - 0.2\mathbf{x}^*, \mathbf{x}^* + 0.2\mathbf{x}^*),$$

where $\mathbf{x}^*$ represents the vector of nominal parameter values. That is, we draw samples from a uniform hypercube where parameter values may vary up to 20% from their defined nominal values.

(2) For each $\mathbf{x}_j$, compute the gradient of $f$,

$$\nabla_{\mathbf{x}} f_j = \nabla_{\mathbf{x}} f(\mathbf{x}_j),$$

where $f(\mathbf{x})\colon \mathbb{R}^m \to \mathbb{R}$ represents the quantity of interest. In cases where the gradient is not easily computed via an analytic solution, estimate the gradient using finite differences,

$$\nabla_{\mathbf{x}} f_j \approx \frac{f(\mathbf{x} + h\mathbf{e}_j) - f(\mathbf{x})}{h},$$

where $h$ is a small step size and $\mathbf{e}_j$ represents the standard basis vector with a one in the $j$th spot and zeros elsewhere.

(3) Compile the gradient vectors into a matrix:

$$G = \frac{1}{\sqrt{M}}[\nabla_{\mathbf{x}} f_1 \ldots \nabla_{\mathbf{x}} f_M].$$

(4) Compute the SVD: $G = USV^T$. This decomposes the matrix $G$ into its singular values (contained in $S$) and matrices containing the left and right singular vectors ($U$ and $V$, respectively) [16].

(5) Partition $S$ according to significance of the singular values,

$$S = \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix},$$

where

$$S_1 = \text{diag}(s_1, \ldots, s_n),$$
$$S_2 = \text{diag}(s_{n+1}, \ldots, s_m).$$

(6) Partition $U = [U_1 \quad U_2]$, such that $U_1$ contains the first $n$ singular vectors and $U_2$ contains the remainder. The vectors in $U_1$ comprise the *active* subspace.

ing to the significance of the corresponding singular value. For parameter $i$ and active subspace dimension $n$, this takes the form of

$$\alpha_i(n) = \sum_{j=1}^{n} {s_j}^2 {u_{i,j}}^2,$$

where $u_{i,j}$ represents the $(i,j)$th component of the active subspace matrix $U_1$. For more details on the use of activity scores as a sensitivity metric, see [6].

In a time-dependent scenario, we must account for the fact that parameter sensitivities may change as a function of time. We now have a quantity of interest of the form $f(t;\mathbf{x})\colon \mathbb{R}^{m+1} \to \mathbb{R}$, where the function depends both on time and on the value of the parameters assigned for each run.

To handle the time-dependence, we take snapshots of the active subspace at time steps $[t_1, t_2, \ldots, t_N]$. This method has been used previously to construct response surfaces (low-dimensional alternatives to a complex model) for an HIV model in [11]. In our investigation, we use these snapshots to compute an activity score for each parameter at each time step, and observe how parameter sensitivities change with respect to time.

# 4 Analysis and Discussion

We now apply the methods of Section 3 to the two compartmental models constructed in Section 2 and analyze the results. All numerical solutions and figures used in the analysis were produced using the `ode` function with default method `LSODA`—which switches automatically between stiff and non-stiff systems—from the `deSolve` package in R [21].

## 4.1 Setup

To begin, we define a scalar quantity of interest for our model, which we take to be the sum of our exposed ($E$) and infectious ($I$) individuals as a proportion of the total population ($N$) at a particular time $t$. Thus, the quantity of interest at time $t$ is the proportion of the population that is affected by tuberculosis in either the latent or active form:

$$f(t) = \frac{E(t) + I(t)}{N(t)}.$$

We construct an active subspace for every two year period over a duration of 300 years, beginning with $t = 1$ up to $t = 301$. To construct our active subspace at each time $t$, we sample the parameters from a multivariate uniform distribution, where parameter values range from $\pm 20\%$ of the nominal values listed in Table 2. Note that

those parameters listed with asterisks in Table 2 are fixed, leaving six varying parameters in the general TB model and 13 in the HIV/TB co-infection model. In each case, we fix those parameters over which medical professionals would have no control, to focus our analysis on those that may have some flexibility in the field. Wherever possible, we use nominal values found in the literature. In cases where values are unavailable, we carefully choose our nominal values to yield the expected model dynamics and satisfy our biological understanding. For instance, we would expect that HIV+ individuals would recover from TB less frequently than their HIV− counterparts due to their potentially weakened immune systems; thus, we choose $\gamma_H < \gamma$. We note that additional work would need to be done to verify these estimates before any conclusions drawn here could be used in a clinical setting. However, as we intend for this investigation to serve as a proof of concept for a procedural methodology rather than to provide concrete recommendations to clinicians in the field, we proceed with these estimates.

The gradient vectors at each of the sample sets are combined into a gradient matrix $G$, from which an active subspace is constructed—see Algorithm 1. To determine the dimension of the active subspace at each time step, we investigate the relative magnitude of the singular values, plotted in Figure 5 for the general TB model. Note that the singular values are normalized so that $\sigma_1 = 1$ in every case, making comparison easier. We use a cutoff value of 0.01 when choosing the dimension; any singular value that attains a value greater than this at any point over the 300 years is deemed significant, and the corresponding singular vector is included in the active subspace for all time steps. For both the general TB and HIV/TB co-infection models, we determined the dimension of the active subspace to be 3. For each of the 151 active subspaces, we compute activity scores with $n = 3$ for each of the varying parameters, to identify which are most influential in changing the value of the quantity of interest. We note that with such a small parameter space, dimension reduction is not strictly necessary here as computation time is not a concern. However, contributions from the final four columns of the $U$ matrix are minimal, and their effects upon the activity scores can be considered negligible.

## 4.2 Analysis of General TB Model

The activity scores of the general TB model are graphed over time for each of the six varying parameters in Figure 6. We observe that in the beginning of the epidemic, $\lambda$, the rate of infection, is by far the most significant parameter. Therefore, we would recommend that resources be allocated with a focus on decreasing contact between the infectious and susceptible populations.

Table 2: Nominal parameter values for both the general TB and co-infection models. Asterisks indicate fixed values. Those parameters that are not fixed are sampled from uniform distributions centered at their listed values.

| Parameter | Nominal, Source |
|---|---|
| ($\beta$) birth rate | 21.3*, [8] |
| ($q$) proportion born HIV− | 0.987*, [22] |
| ($\delta$) natural death rate | 0.00861*, [9] |
| ($\nu, \nu_H$) rate of vaccination | 0.00445, 0.002, [26] |
| ($\lambda, \lambda_H$) rate of infection | 0.0018, 0.08, Est. |
| ($p$) slow-track TB proportion | 0.95*, [2] |
| ($\omega, \omega_H$) rate of deterioration | 0.0084, 0.01, [24, 20] |
| ($\tau, \tau_H$) death rate due to TB | 0.00079*, 0.0375*, [24] |
| ($\gamma, \gamma_H$) rate of recovery | 0.02*, 0.0001*, Est. |
| ($\sigma_c, \sigma_{c_H}$) latent TB treatment | 0.005, 0.0001, Est. |
| ($\sigma_t, \sigma_{t_H}$) active TB treatment | 0.005, 0.0001, Est., [1] |
| ($\theta, \theta_H$) rate of reinfection | 0.0005, 0.009, [2, 4] |
| ($\mu$) HIV incidence rate | 0.0025, [24] |

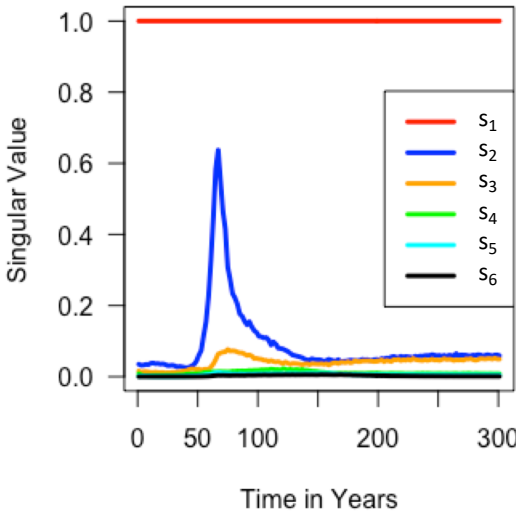


Figure 5: Normalized singular values for the general TB model.

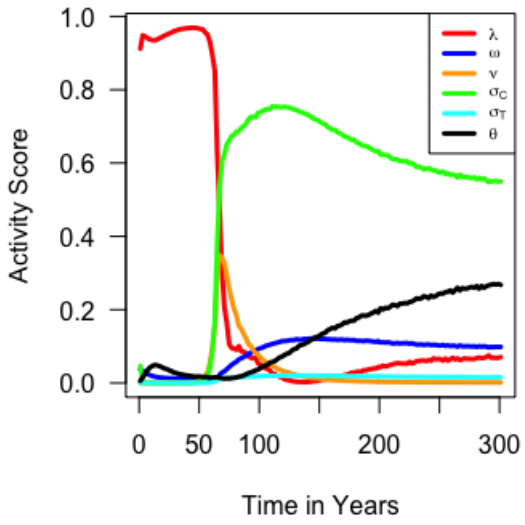


Figure 6: Activity Scores for General TB Model

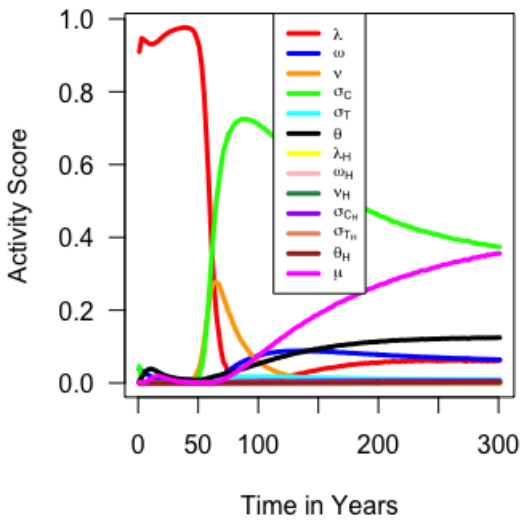


Figure 7: Activity Scores for HIV/TB Co-infection.

This could take the form of quarantine for infectious individuals and/or encouraging susceptible individuals to get vaccinated, thereby moving them (at least temporarily) to the recovered population.

Around year 50 of the epidemic, the significance of $\sigma_c$ and $\nu$, the rates of latent treatment and vaccination, increases. This indicates that treatment of exposed individuals and vaccination of susceptible individuals both play a large role in driving down the total number of people affected by TB during this time frame. Allocating resources to these two venues would be the most effective form of controlling the epidemic at this stage. Note that this could include any of the following: improving vaccination efficacy, increasing the number of vaccinations given, improving current treatments for slow TB, increasing the number of individuals receiving treatment, and monitoring those being treated to make sure they complete the full treatment regimen.

Towards the end of the epidemic, $\sigma_c$ remains highly significant. The significance of $\theta$ and $\omega$ increase as well. The rise in significance of the relapse rate, $\theta$, coupled with the decrease in significance of $\nu$, indicates that our focus should shift from vaccinating the susceptible population to monitoring recovered individuals at the risk of relapse. Since the rate of deterioration, $\omega$, is the rate at which latently infected individuals become infectious, we would still want to keep a large focus on latent TB treatment as well as therapeutic treatment for those that do progress to active infections.

### 4.3 Analysis of HIV/TB Co-infection Model

Figure 7 shows the activity scores of the HIV/TB co-infection model. For the most part, we observe similar behavior as in the general TB case. During the beginning of the epidemic, the most significant parameter is $\lambda$, the rate of infection, indicating a need for quarantine of infectious individuals and increased vaccination rates for the susceptible population to reduce the number of individuals at risk of infection. As time progresses, $\sigma_c$ and $\nu$ once again increase in significance, indicating the efficacy of latent treatment and vaccinations in controlling TB epidemics. As in the general TB model, we see an increase in significance of the relapse rate, $\theta$ towards the end of the epidemic.

One notable difference in the HIV/TB co-infection model is the significance of $\mu$, the HIV contraction rate. Beginning around year 75, the activity scores for $\mu$ undergo a steady increase through to the end of the observed time period, eventually becoming just as significant as the treatment rate of latent TB, $\sigma_c$. This suggests that one of the most effective ways of driving down the number of individuals affected by TB would actually be to focus on controlling the spread of HIV, as the TB infection takes advantage of weakened immune systems such as those exhibited by HIV+ individuals. Spreading awareness of HIV and emphasizing safe-sex methods could be vital to controlling the spread of both HIV and TB in these regions.

## 5 Conclusion

Tuberculosis is an airborne disease, therefore it is treated and prevented through strategies such as quar-

antine, vaccinations, and drug treatment. These medical approaches have been used for decades, quelling the spread of TB for brief periods at a time, though the infection often returns with a vengeance. As of the twenty-first century, the prevalence of HIV as well as drug resistant strains of TB have incited a modern day epidemic in many developing regions of the world. Given the recent increase in the incidence of tuberculosis on a global scale, we conducted this research with the goal of accurately modeling tuberculosis spread in varying environments and aiding in the design of a protocol for determining effective methods of prevention and treatment.

Through use of time-dependent sensitivity analysis using techniques derived from active subspaces, we were able to pinpoint those parameters in our proposed models that are most influential in driving down the proportion of the population that is affected by TB at any given time. Among those parameters that were deemed most significant were the rate of infection (toward the beginning of the epidemic), the latent treatment and vaccination rates (toward the middle of the epidemic), and the relapse and deterioration rates (as the 300 year observed time period draws to an end). In addition, we discovered that in regions where HIV is prevalent, controlling the spread of HIV is vital to the eventual eradication of TB.

We stress that all conclusions drawn from this model assume the validity of the parameter values listed in Table 2. As data is not available for each of the four states over a full epidemic period for model fitting, we have used parameter values from the literature when available and chosen values to ensure expected behavior of the state variables when not. Additional model validation and verification would be required before the conclusions drawn here could be utilized at the clinical level. As such, we present this investigation as a procedural methodology, and claim that for a properly validated model, our procedure can illuminate not only which strategies of treatment and prevention may be most effective, but perhaps even more importantly, *when* each strategy would be most effective. This can help us to achieve optimal allocation of scarce resources for treatment and prevention as we fight for the eventual eradication of TB on a global scale.

## Acknowledgements

This research was supported by grant number DSM-1560301 from the National Science Foundation, and was conducted during the Research Experience for Undergraduates Emerging Scholars Program at St. Mary's College of Maryland in the summer of 2018.

## References

[1] Bekker, L., & Wood, R. (2010). The Changing Natural History of Tuberculosis and HIV Coinfection in an Urban Area of Hyperendemicity. *Clinical Infectious Diseases, 50(S3).* doi:10.1086/651493

[2] Blower, S. M., Hopewell, P. C., Mclean, A. R., Moss, A. R., Porco, T. C., Sanchez, M. A., & Small, P. M. (1995) The intrinsic transmission dynamics of tuberculosis epidemics. *Nature Medicine, 1*, 815–821. doi:10.1038/nm0895-815.

[3] Center for Disease Control. (2014). Tuberculosis (TB). Retrieved July 2018 from `https://www.cdc.gov/tb/publications/factsheets/general/ltbiandactivetb.htm`

[4] Chaisson, R. E., & Churchyard, G. J. (2010). Recurrent Tuberculosis: Relapse, Reinfection, and HIV. *The Journal of Infectious Diseases, 201(5),* 653–655. doi:10.1086/650531.

[5] Constantine, P. G. (2015). *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies*, SIAM Spotlight.

[6] Constantine, P. G. & Diaz, P. (2017). Global sensitivity metrics from active subspaces. *Reliability Engineering & System Safety, 162*, 1–13. doi:10.1016/j.ress.2017.01.013.

[7] Halko, N., Martinsson, P. G., & Tropp, J.A. (2011). Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review, 53(2)*, 217–288.

[8] IndexMundi. (2015) South Africa Birth Rate. Retrieved July 2018 from `https://www.indexmundi.com/south_africa/birth_rate.html`

[9] IndexMundi. (2018) South Africa Death Rate. Retrieved July 2018 from `https://www.indexmundi.com/south_africa/death_rate.html`

[10] Kermack, W. O. & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London, Series A, 115(772)*, 700–721.

[11] Loudon, T. & Pankavich, S. (2017). Mathematical analysis and dynamic active subspaces for a long term model of HIV. *Mathematical Biosciences and Engineering, 14(3)*, 709–733.

[12] Luca, S. & Mihaescu, T. (2013). History of BCG Vaccine. *Maedica: A Journal of Clinical Medicine, 8(1)*, 53–58.

[13] Luo, W. & Li, B. (2013). Combining eigenvalues and variation of eigenvectors for order determination. *Biometrika, 99(1)*, 1–12.

[14] MacIntyre, C. R. (2007). New Developments in BCG Vaccine: Implications for Tuberculosis Control. *Epidemiology and Infection, 135(2),* 177–180.

[15] Marino, S., Hogue, I. B., Ray, C. J., & Kirschner, D. E. (2009). A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of Theoretical Biology, 254(2009)*, 178–196.

[16] Meyer, C. D. (2000). Matrix Analysis and Applied Linear Algebra. *SIAM.*

[17] Morris, M. D. (1991). Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics, 33(2)*, 161–174.

[18] Ozcaglar, C., Shabbeer, A., Vandenberg, S. L., Yener, B., & Bennett, K. P. (2012). Epidemiological models of Mycobacterium tuberculosis complex infections. *Math Biosciences, 236(2)*, 77–96. doi: 10.1016/j.mbs.2012.02.003.

[19] Principles of Epidemiology—Lesson 1 – Section 11. (n.d.). Retrieved from `https://www.cdc.gov/csels/dsepd/ss1978/lesson1/section11.html`

[20] Shea, K. M., Kammerer, J. S., Winston, C. A., Navin, T. R., & Horsburgh, C. R., Jr. (2014). Estimated rate of reactivation of latent tuberculosis infection in the United States, overall and by population subgroup. *American Journal of Epidemiology, 179(2),* 216–225. doi:10.1093/aje/kwt246.

[21] Soetaert, K., Petzoldt, T., & Setzer, R. W. (2010). Package deSolve: Solving Initial Value Differential Equations in R. `http://desolve.r-forge.r-project.org/`

[22] The South African National AIDS Council. (2017) Let Our Actions Count: South Africa's National Strategic Plan for HIV, TB and STIs 2017–2022. *SANAC.* Retrieved July 2018 from `https://sanac.org.za//wp-content/uploads/2017/06/NSP_FullDocument_FINAL-1.pdf`

[23] Stoyanov, M. & Webster, C. G. (2015). A gradient-based sampling approach for dimension reduction of partial differential equations with stochastic coefficients. *International Journal for Uncertainty Quantification, 5(1)*, 49–72.

[24] TBfacts.org. (2017) TB Statistics South Africa – National, incidence, provincial. *Global Health Education.* Retrieved July 2018 from `https://www.tbfacts.org/tb-statistics-south-africa/`

[25] Vynnycky, E. & White, R. (2010). *An Introduction to Infectious Disease Modeling.* New York: Oxford University Press.

[26] World Health Organization. (2013). TB Emergency Declaration. Retrieved July 2018 from `http://www.who.int/tb/features_archive/tb_emergency_declaration/en`

[27] World Health Organization. (2017). Global Health Observatory Data. Retrieved July 2018 from `http://www.who.int/gho/en/`