

Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

D.P. MOUALEU, S. RÖBLITZ, R. EHRIG, P. DEUFLHARD

Parameter Identification for a Tuberculosis Model in Cameroon

Herausgegeben vom
Konrad-Zuse-Zentrum für Informationstechnik Berlin
Takustraße 7
D-14195 Berlin-Dahlem

Telefon: 030-84185-0
Telefax: 030-84185-125

e-mail: bibliothek@zib.de
URL: <http://www.zib.de>

ZIB-Report (Print) ISSN 1438-0064
ZIB-Report (Internet) ISSN 2192-7782

Parameter Identification for a Tuberculosis Model in Cameroon

D.P. Moualeu, S. Röblitz, R. Ehrig, P. Deuffhard

Zuse Institute Berlin, Germany

Abstract

A deterministic model of tuberculosis in sub-Saharan Africa in general and Cameroon in particular including lack of access to the treatment and weak diagnose capacity is designed and analyzed with respect to its transmission dynamics. The model includes both frequency- and density-dependent transmissions. It is shown that the model is mathematically well-posed and epidemiologically reasonable. Solutions are non-negative and bounded whenever the initial values are non-negative. A sensitivity analysis of model parameters is performed and most sensitive parameters of the model are identified using a state-of-the-art Gauss-Newton Method. In particular, parameters representing the proportion of individuals having access to medical facilities have a large impact on the dynamics of the disease. It has been shown that an increase of these parameter values over the time can significantly reduce the disease burden in the population within the next 15 years.

1 Introduction

Tuberculosis (TB) is a preventable and curable disease caused by *Mycobacterium tuberculosis* (Mtb) that most often affects the lungs. To date, TB claims the second largest number of victims due to a single infectious agent right after Human Immunodeficiency Virus and Acquired Immune Deficiency Syndrome (HIV/AIDS) [WHO12a]. Mostly, young adults are affected by tuberculosis, in their most productive years. In 2010, 8.8 million people worldwide were infected. Over 95% of TB deaths occur in 22 low- and middle-income countries mostly located in Sub-Saharan Africa and in South-East-Asia [WHO12a]. Sub-Saharan Africa carries the greatest proportion of new cases per population with over 260 cases per 100 000 population in 2011 [WHO12a]. Tuberculosis is an ancient and complex infectious disease on which a large number of theoretical studies have been carried out. Mtb's infection can remain latent, become active, or it can progress from latent TB to active TB either by endogenous re-activation and/or exogenous re-infection. Active TB is most of the time acquired through co-infection of Mtb with other diseases (diabetes, HIV/AIDS) or some substance abuse such as alcohol and tobacco. The mathematical analysis of biomedical and disease transmission models can significantly contribute to the understanding of the mechanisms of those processes and to the design of potential therapies [AM91, Thi03].

Diseases (including heart disease and cancer) cause orders of magnitude more deaths in the world than anything else, even wars and famines [Mur02, RH04]. The earliest mathematical models describing the TB dynamics have been built in the 1960s by the statistician H. T. Waaler, chief of the Norwegian TB control services [WGA62].

The models focused on the prediction and control strategies using simulation approaches. Later on, Waaler [Waa68], Revelle and coworkers [RLF67] and Ferebee [Fer67] developed many other mathematical models with the same aim, Waaler considered an exponential population dynamics in the absence of TB [WGA62]. Waaler's first linear model did not describe the mechanics of TB transmission. He introduced a new model of 160 linear equations in [Waa68], keeping the same structure, but including BCG vaccinated and different recovered classes for 20 different age classes. Using the model of Brogger and Waaler as a template, Revelle firstly introduced nonlinear systems of ordinary differential equations (ODEs) that model TB dynamics [RLF67]. Revelle developed an optimization model and studied the minimal cost strategy against TB.

Blower and colleagues [BSH96, VF97, BD02] discussed the persistence condition of tuberculosis inside the population and determined the basic reproduction ratio \mathcal{R}_0 (the average number of new infectious cases caused by a single infectious case in a fully susceptible population over the course of the entire infectious period). A sensitivity analysis of \mathcal{R}_0 has been performed by several authors [BSH96]. However, the sensitivity analysis of parameters on \mathcal{R}_0 does not really illustrate the impact of these parameters on the global trajectory of the system, especially in the presence of backward bifurcation. Blower and colleagues found in their model that $1 < \mathcal{R}_0 < 9$, and the most important parameters are the infection rate, the probability of fast progression, the re-activation rate, and the TB related death rate. Chavez and colleagues [CCF97] developed a mathematical analysis of a TB model without fast progression. Thereafter, most publications include sophisticated mathematical theories to study the dynamics of tuberculosis, such as center manifold theory and Lyapunov functions [FCCC00, MSK03, MBE11, TDB09].

The challenge of TB control in developing countries is due to the increase of TB incidence by a high level of undiagnosed infectious population and lost sight population with respect to diagnosed infectious cases. Undiagnosed infectious population means people who have not yet been to a hospital for diagnosis or have not been detected, but have a pulmonary TB [BTT⁺11, RKL⁺07] when lost sight population are people who have been diagnosed as having active TB, begun their treatment and quitted before the end. Compared to existing results [CCF97, BGMZ08, MSK03, BOP⁺08, ACCC02, CM99, MBE11, OK07, FCCC00] and references therein, our work differs from these studies in that our model, in addition to undiagnosed infectious and lost sight population, also considers the aspects of exogenous re-infections, disease relapse as well as primary active TB cases, natural recovery and traditional medicine or self-medication (practiced in Sub-Saharan Africa). Also, it is recognized that undiagnosed population, lost sight population and exogenous re-infections are important components of TB transmission in Sub-Saharan Africa. A new mathematical model will be designed and qualitatively analyze to gain insights into the transmission dynamics and control of TB in a population in developing countries. The infective class is divided into three subgroups with different properties: i) diagnosed infectious population, ii) undiagnosed infectious population and iii) lost sight population. According to the National Committee of Fight against TB of Cameroon (NCFT) [oFAT01], about 8% of diagnosed infectious that begin their therapy treatment never return to the hospital for the rest of sputum examinations and treatment, and then become lost sight. This class of TB epidemiological models can be extended to many classes of infective individuals and data for many other African countries.

The quite high rate (5 to 17 %) of lost sight individuals among the Cameroonian population raises a number of concerns. Indeed, what is happening with the undiagnosed cases of active TB and lost sight? How do these people affect the dynamics of

TB in Sub-Saharan Africa? What are the conditions for the diagnosed rate, i.e. the proportion of the diagnosed cases treated under a DOTS (Directly Observed Treatment Strategy) program, that can ensure the eradication of TB, or at least minimize its incidence? Are the undiagnosed cases undermining the efforts of the DOTS strategy with respect to reducing the incidence of TB in Sub-Saharan Africa and effecting proper and efficient treatment policies for patients with active TB? These are the questions to be answered in this paper.

2 Suggested epidemiological model

Data availability and the study objectives generally determine the modelling approach to be used. A finite total population at time t denoted by $N(t)$ is considered and sub-divided into the following mutually exclusive sub-populations:

- S susceptible: healthy people not yet exposed to TB
- E latently infected: exposed to TB but not infectious
- I diagnosed infectious: have active TB confirmed after a sputum examination in a hospital
- J undiagnosed infectious: have not yet been to a hospital for diagnosis but are active for confirmation by a sputum examination
- L lost sight: people who have been diagnosed as having active TB, begun their treatment and quitted before the end
- R recovered: people cured after treatment in the hospital

In Africa, reliable TB tests [KAFM07] are often missing or too expensive. Hence, TB diagnosis based on a single sputum examination can often only be classified as “probable” or “presumed”, and cannot detect cases of less infectious forms of TB [WHO12b]. Therefore, the model is based on the following assumptions, established from behaviours of people in different epidemiological classes.

1. *Mtb* transmission from diagnosed infectious to susceptible population, due to education on the infection is limited. it is therefore modeled using a standard mass action or frequency-dependent force of infection.
2. *Mtb* transmission from undiagnosed infectious to susceptible population, due to their level of education on the disease is modeled by a density-dependent force of infection.

These arguments abide on the fact that diagnosed infectious people are in most cases hospitalized for at least 2 months or are advised to lessen their infectiousness in their residing neighbourhood. Their distribution in the population is not necessarily homogeneous. Since undiagnosed infectious remain inside the population, there is an unlimited possibility of contacts with the susceptible population [BTT⁺11]. We therefore assume a density dependent force of infection for hospital inmates [BBB⁺02].

All recruitment is into the susceptible class and occurs at an average scale Λ . The fixed survey for non-disease related death is μ , thus $1/\mu$ is the average lifetime. Diagnosed infectious, undiagnosed infectious and lost sight population have additional constant death rates due to the disease, defined by d_1 , d_2 and d_3 , respectively. It is Transmission of *Mtb* occurs due to adequate contacts among susceptible and an active

TB case. Thus, susceptible individuals acquire Mtb infection from individuals with active TB and lost sight at a rate $\nu(I, J, L)$ given by

$$\nu(I, J, L) = \beta_1 \frac{I}{N} + \beta_2 \frac{L}{N} + \beta_3 J, \quad (1)$$

where β_i , $i = 1, 2, 3$, are the effective contact rates with diagnosed, lost sight and undiagnosed infectious population sufficient to transmit infection to susceptible people. The effective contact rates β_i in a given population for tuberculosis are measured in effective contacts per unit time. This may be expressed as the product of the total contact rate per unit time (η_i) by the risk of infection (ϕ_i) given contact between an infectious and a susceptible individual,

$$\beta_i = \eta_i \phi_i.$$

This risk is called the transmission risk.

A proportion p of the latently-infected individuals develop fast active TB and the remainder $(1 - p)$ develop latent TB and enter the latent class E . Among latently-infected individuals developing active TB, a fraction f is assumed to undergo a fast progression directly to the diagnosed infectious class I , while the remainder $(1 - f)$ enters the undiagnosed infectious class J . We set

$$p_1 = pf \quad \text{and} \quad p_2 = p(1 - f).$$

Once latently infected with Mtb, an individual will remain so for life unless reactivation occurs. Latently infected individuals are assumed to acquire some immunity as a result of infection, which reduces the risk of subsequent infection but does not fully prevent it.

Due to endogenous reactivation, a fraction $1 - r_1$ of latently infected individuals who did not receive effective chemoprophylaxis become infectious with a constant rate k , and reinfect after effective contact with individuals in the active TB classes or lost sight at a rate

$$\lambda_e = \sigma_1 \nu(I, J, L),$$

where σ_1 is the factor reducing the risk of infection as a result of acquiring immunity for latently infected individuals. Among latently infected individuals who become infectious, the fraction h is diagnosed and treated under the "Stop TB" program, while the remaining $1 - h$ is not diagnosed and becomes undiagnosed infectious J . We assume that after some time suffering from TB, some undiagnosed infectious decide to go to hospital with a rate θ . Also, we assume that among diagnosed infectious who had begun their treatment therapy, a fraction r_2 of I has taken all the dose and has made all the sputum examinations and will be declared cured from the disease. Some diagnosed infectious who have not finished their dose of drugs and sputum examinations or whose treatment was unsuccessful, will not return to the hospital for the rest of sputum examinations and check-up. They will enter the class of lost sight L at a constant rate α . Lost sight can return to the hospital at a constant rate δ .

As suggested by Murray et al. [MSR90], recovered individuals can only have partial immunity. Hence, they can undergo a TB reactivation or relapse with a constant rate γ . The remainder can be reinfected (exogenously) after an effective contact with individuals in the active TB classes and lost sight at a rate

$$\lambda_r = \sigma_2 \nu(I, J, L),$$

where σ_2 is the factor reducing the risk of infection as a result of acquiring partial immunity for recovered individuals. Due to their own immunity, traditional medicine,

natural recovery and drugs bought in the street (practised in sub-Saharan Africa), a fraction of lost sight and undiagnosed infectious can spontaneously recover at constant rates ρ and ω , respectively and enter the latent class E and recovery class R respectively. A description of the parameters is summarized in Table 2. The whole model flow diagram is shown in Fig 1. The flow diagram Figure 1 yields the following

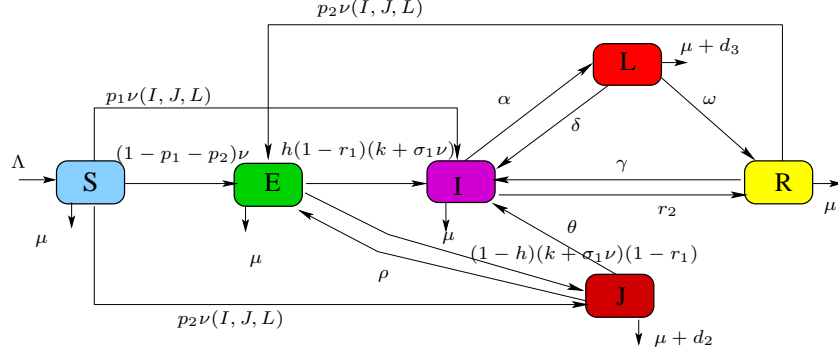


Figure 1: Transfer diagram for a transmission dynamics of tuberculosis.

differential equations,

$$\begin{cases} \dot{S} &= \Lambda - \nu(I, J, L)S - \mu S, \\ \dot{E} &= (1 - p_1 - p_2)\nu(I, J, L)S + \rho J + \sigma_2\nu(I, J, L)R \\ &\quad - \sigma_1(1 - r_1)\nu(I, J, L)E - A_1 E, \\ \dot{I} &= p_1\nu(I, J, L)S + \delta L + \theta J + \gamma R + h(1 - r_1)(k + \sigma_1\nu(I, J, L))E \\ &\quad - A_2 I, \\ \dot{J} &= p_2\nu(I, J, L)S + (1 - h)(1 - r_1)(k + \sigma_1\nu(I, J, L))E - A_3 J, \\ \dot{L} &= \alpha I - A_4 L, \\ \dot{R} &= r_2 I + \omega L - \sigma_2\nu(I, J, L)R - A_5 R, \end{cases} \quad (2)$$

where

$$\begin{aligned} A_1 &= \mu + k(1 - r_1), & A_2 &= \mu + d_1 + r_2 + \alpha, \\ A_3 &= \mu + d_2 + \theta + \rho, & A_4 &= \mu + d_3 + \delta + \omega \quad \text{and} \quad A_5 = \gamma + \mu. \end{aligned}$$

3 Basic properties of the model

Since model (2) monitors a human population, all its associated parameters and state variables should be non-negative and bounded for all $t \geq 0$. It is shown in this section that the model is mathematically well-posed and epidemiologically reasonable [Het00].

Positivity of the solution

The following result shows that state variables are non-negative and dissipative.

Lemma 3.1 *Let the initial values be $S(0) > 0$, $E(0) \geq 0$, $I(0) \geq 0$, $J(0) \geq 0$, $L(0) \geq 0$ and $R(0) \geq 0$. Then, solutions (S, E, I, J, L, R) of model system (2) are non-negative for all $t > 0$. Furthermore,*

$$\limsup_{t \rightarrow \infty} N(t) \leq \frac{\Lambda}{\mu},$$

with $N(t) = S(t) + E(t) + I(t) + J(t) + L(t) + R(t)$.

The proof of this Lemma follows from an obvious adjustment of the result in [MBE12, MMNB11]

Invariant Region

The following steps establish the positive invariance of the set

$$\Omega_\rho = \left\{ (S, E, I, J, L, R) \in \mathbb{R}_+^6, \quad N(t) \leq \frac{\Lambda}{\mu} + \rho \right\}, \quad \rho > 0, \quad (3)$$

i.e. solutions remain in Ω_ρ for all $t \geq 0$. This implies that the trajectories of model system (2) are bounded. On the other hand, integrating the differential inequality $\dot{N} \leq \Lambda - \mu N$ yields

$$N(t) \leq N(0)e^{-\mu t} + \frac{\Lambda}{\mu}(1 - e^{-\mu t}).$$

In particular $N(t) \leq \frac{\Lambda}{\mu}$ if $N(0) \leq \frac{\Lambda}{\mu}$. On the other hand, if $N(0) \geq \frac{\Lambda}{\mu}$, then $\Lambda - \mu N(0) \leq 0$, and

$$\dot{N}(0) \leq \Lambda - \mu N(0) \leq 0,$$

i.e. the total population $N(t)$ will decrease until

$$N(t) \leq \frac{\Lambda}{\mu}.$$

Thus, the simplex Ω_ρ is a compact forward invariant set for model system (2), and for $\rho > 0$ this set is absorbing. So, we limit our study to this simplex for $\rho > 0$. The prevalent existence, uniqueness and continuation results hold for model system (2) in Ω_ρ .

4 Basic reproduction number

The global behavior of the TB model crucially depends on the basic reproduction number, i.e., the average number of secondary cases produced by a single infective individual, who is introduced into an entirely susceptible population. Model system (2) has an evident equilibrium $Q_0 = (x_0, 0)$ with $x_0 = \Lambda/\mu$ when $I = 0$. This equilibrium point is the *disease-free equilibrium* (DFE). We now calculate the basic reproduction number \mathcal{R}_0 , using the next generation method developed in [vdDW02]. For that purpose, let us write system (2) in the form

$$\begin{cases} \dot{x} &= \varphi(x) - \nu(I, J, L)x, \\ \dot{y} &= \mathcal{F}(x, y) + \mathcal{V}(x, y), \end{cases} \quad (4)$$

where

$$\mathcal{F}(x, y) = \nu(I, J, L)B_1x, \quad \mathcal{V}(x, y) = \nu(I, J, L)[B_2\langle e_1 | y \rangle + B_3\langle e_5 | y \rangle] + Ay, \quad (5)$$

$x = S \in \mathbb{R}_+$ is a state representing the compartment of non transmitting individuals (susceptible), $y = (E, I, J, L, R)^T \in \mathbb{R}_+^5$ is the vector representing the state compartment of different infected individuals,

$$\varphi(x) = \Lambda - \mu x$$

is a function that depends of x ,

$$\nu(I, J, L) = \frac{\langle e_1 | y \rangle}{N} + \langle e_2 | y \rangle, \quad N = x + y_1 + y_2 + y_3 + y_4 + y_5$$

is the force of infection. Furthermore, we set

$$\begin{aligned} e_1 &= (0, \beta_1, \beta_2, 0, 0) \in \mathbb{R}^5, & e_2 &= (0, 0, 0, \beta_3, 0) \in \mathbb{R}^5, \\ e_3 &= (1, 0, 0, 0, 0) \in \mathbb{R}^5, & e_4 &= (0, 0, 0, 0, 1) \in \mathbb{R}^5, \\ B_1 &= (1 - p_1 - p_2, p_1, p_2, 0, 0)^T \in \mathbb{R}^5, \\ B_2 &= (-\sigma_1(1 - r_1), h\sigma_1(1 - r_1), \sigma_1(1 - h)(1 - r_1), 0, 0)^T \in \mathbb{R}^5, \\ B_3 &= (-\sigma_2(1 - \gamma), 0, 0, 0, \sigma_2(1 - \gamma))^T \in \mathbb{R}^5. \end{aligned}$$

$\langle \cdot | \cdot \rangle$ is the usual scalar product and A is the constant matrix

$$A = \begin{bmatrix} -A_1 & 0 & \rho & 0 & 0 \\ kh(1 - r_1) & -A_2 & \theta & \delta & \gamma \\ k(1 - h)(1 - r_1) & 0 & -A_3 & 0 & 0 \\ 0 & \alpha & 0 & -A_4 & 0 \\ 0 & r_2 & 0 & \omega & -A_5 \end{bmatrix},$$

with A_1, A_2, A_3, A_4 and A_5 defined as above in equation (2).

We define the Jacobian matrices at the DFE as

$$F = \frac{\partial \mathcal{F}}{\partial y}(Q_0) \quad \text{and} \quad V = \frac{\partial \mathcal{V}}{\partial y}(Q_0).$$

Using the same notations as in [vdDW02], the basic reproduction number is given by the spectral radius of FV^{-1} ,

$$\mathcal{R}_0 = \rho(FV^{-1}). \quad (6)$$

For model (4), one has

$$F = B_1 \left(e_1 + \frac{\Lambda}{\mu} e_2 \right) \quad \text{and} \quad V = -A.$$

Then, according to [KS08, IMST07], the basic reproduction ratio is given by

$$\mathcal{R}_0 = \left\langle e_1 + \frac{\Lambda}{\mu} e_2 \mid (-A^{-1})B_1 \right\rangle. \quad (7)$$

The following result is established from [vdDW02]:

Lemma 4.1 : *The disease-free equilibrium Q_0 of model system (2) is locally asymptotically stable whenever $\mathcal{R}_0 < 1$, and unstable, if $\mathcal{R}_0 > 1$.*

From a biological point of view, Lemma 4.1 implies that TB can be eliminated from the community (when $\mathcal{R}_0 \leq 1$) if the initial size of the population is in the basin of attraction of Q_0 . But if $\mathcal{R}_0 > 1$ the infection will be able to spread in a population. Generally, the larger the value of \mathcal{R}_0 , the harder it is to control the epidemic.

5 Parameter sensitivity and identification

In order to effectively enable the assessment of targeted public health education strategies and chemoprophylaxis against TB spread in a population, we test the suitability of the model by fitting it to data from Cameroon [WHO12b].

5.1 Background

We will briefly describe the mathematical techniques that we use for parameter identification. Formally, the system of differential equation (2) can be written as

$$\begin{cases} \frac{d}{dt}y(t, \mathbf{p}) &= f(t, y, \mathbf{p}), & t \geq 0, \\ y(0, \mathbf{p}) &= y_0, \end{cases} \quad (8)$$

where \mathbf{p} is the vector of parameters, and the right-hand side f depends on both the states, $y \in \mathbb{R}^n$, and the parameter vector, $\mathbf{p} \in \mathbb{R}^q$. The initial condition vector, y_0 , has the same dimension as the state vector y . The TB model equation(2) can be written in the form of equation (8), where

$$y(t, \mathbf{p}) = (S(t, \mathbf{p}), E(t, \mathbf{p}), I(t, \mathbf{p}), J(t, \mathbf{p}), L(t, \mathbf{p}), R(t, \mathbf{p})) \in \mathbb{R}^6$$

and

$$\mathbf{p} = (\Lambda, \beta, \dots, \mu) \in \mathbb{R}^{22}.$$

Assume there are given m experimental measurement time-point τ_1, \dots, τ_m , and corresponding data values $z_j \in \mathbb{R}^n$, $j = 1, \dots, m$ associated with corresponding tolerances $\delta z_j \in \mathbb{R}^n$. Parameter identification consist of solving the least-squares minimization problem

$$g(\mathbf{p}) = \frac{1}{m} \sum_{j=1}^m \| D_j^{-1} \cdot (y(\tau_j, \mathbf{p}) - z_j) \|_2^2 \longrightarrow \min_{\mathbf{p}} \quad (9)$$

with diagonal weighting

$$D_j := \text{diag}((\delta z_j)_1, \dots, (\delta z_j)_n) \in \mathcal{M}_n(\mathbb{R}), \quad j = 1, \dots, m. \quad (10)$$

That means we want to minimize the relative deviation of model and data at the measurement time points τ_j . The minimization problem (9) can be written as an equivalent minimization problem defined by

$$g(\mathbf{p}) := F(\mathbf{p})^T \cdot F(\mathbf{p}) \longrightarrow \min_{\mathbf{p}}, \quad (11)$$

where $F(\mathbf{p}) = (F_1(\mathbf{p}), \dots, F_m(\mathbf{p}))$ is a vector of length $N = m \cdot n$ with entries defined by

$$F(\mathbf{p}) = \begin{bmatrix} D_1^{-1} \cdot (y(\tau_1, \mathbf{p}) - z_1) \\ \vdots \\ D_m^{-1} \cdot (y(\tau_m, \mathbf{p}) - z_m) \end{bmatrix}. \quad (12)$$

$F : \mathbb{R}^q \longrightarrow \mathbb{R}^N$ is a non-linear mapping and structured as a stacked vector. If not all components of a measurement z_j are given, the number N is accordingly made smaller, $N < nm$. The above problem, which is highly nonlinear in \mathbf{p} , can be solved

by affine covariant Gauss-Newton iteration, see [Deu04], where each iteration step i requires the solution of a linear least-squares problem,

$$\begin{aligned} \|J(\mathbf{p}^k) \cdot \Delta\mathbf{p}^k + F(\mathbf{p}^k)\|_2 &\rightarrow \min_{\mathbf{p}^k}, \\ \mathbf{p}^{k+1} &= \mathbf{p}^k + \Delta\mathbf{p}^k \end{aligned} \quad (13)$$

where $J(\mathbf{p}^k) = F'(\mathbf{p}^k)$.

5.2 Sensitivity analysis

The sensitivities

$$s_{ij}(t) = \frac{\partial y_i(t)}{\partial \mathbf{p}_j} \quad (14)$$

are computed at all data time-points and stacked into one large sensitivity matrix. In equation (13), $J(\mathbf{p}) = (s_{ij})$ represents the sensitivity of the solution y with respect to the parameters \mathbf{p} at the time points of measurements. An analysis of the matrix $J(\mathbf{p})$ gives some hints whether the current combination of model and data will allow an identification of a given parameter.

Figure 2(a) shows the column norm ($\|DF(:,j)\|$) of this sensitivity matrix. Obviously, parameters with the largest column norm are γ , \mathbf{p}_2 and ρ , respectively. They have the largest influence on the solution trajectory at the measurement time-points.

5.3 Parameter identification

Parameters with very small sensitivity have nearly no influence on the solution and can therefore not be estimated. In this case the entries of the corresponding column in $J(\mathbf{p})$ (and thus the weighted l^2 column norm) are almost zero. Furthermore, some of the parameters might be linearly dependent, which leads to nearly identical columns in $J(\mathbf{p})$. In both cases the matrix $J(\mathbf{p})$ will be singular or, from a numerical point of view, nearly singular.

The linear least squares problem (13) is solved by QR factorization with column pivoting [DH03]. By a suitable permutation of the matrix columns of $J(\mathbf{p})$, the diagonal elements of the upper triangular matrix R can be ordered in the form $r_{11} \geq r_{22} \geq \dots \geq r_{qq}$. The sub-condition of parameter \mathbf{p}_j is defined by

$$sc_j = \frac{r_{11}}{r_{jj}}. \quad (15)$$

Thus, the permutation of matrix columns corresponds to a new ordering of parameters according to increasing sub-condition. The sub-condition indicates whether a parameter can be estimated from the given data or not. Only parameters for which

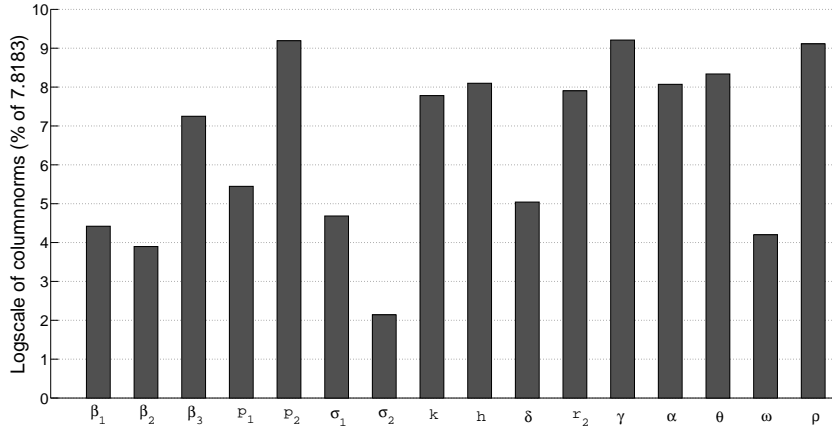
$$sc_j \leq 1/\varepsilon, \quad (16)$$

where ε is the relative precision of the Jacobian $J(\mathbf{p})$ can be estimated [DS80].

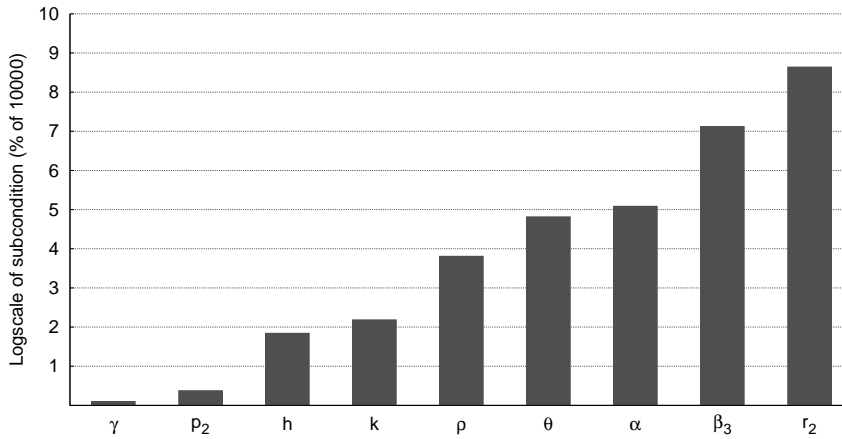
In the global case it comes that

$$\mathbf{p}^{i+1} = \mathbf{p}^i + \lambda_i \Delta\mathbf{p}^i; \quad 0 < \lambda_i \leq 1.$$

The step length $0 < \lambda_i < 1$ is computed successively in each iteration by a trust-region method [Deu04].



(a) Sensitivity norms of parameters



(b) Subcondition numbers (15) of identifiable parameters for $\varepsilon = 10^{-4}$

Figure 2: Subcondition numbers from equation (15) of identifiable parameters for $\varepsilon = 10^{-4}$ and sensitivity norms of parameters

The above described method for solving a non-linear least squares problem is implemented in the software code NLSCON [Deu04] and part of the software package BioPARKIN [DWR11]. Here, a Matlab-based version of this software package, named POEM 2.0, which is especially adapted to parameter identification in ordinary differential equation models, has been used throughout the study.

For model (2), some important parameters are well-known have fixed values according to the National Institute of Statistic of Cameroon (NIS) [BSH96, BOP⁺08]. Other parameters were estimated using the software POEM. As data set for parameter identification, we used figures for diagnosed infectious and total population over the

period 1994-2010 published by WHO [WHO12a]. The data are illustrated in Figure 3.

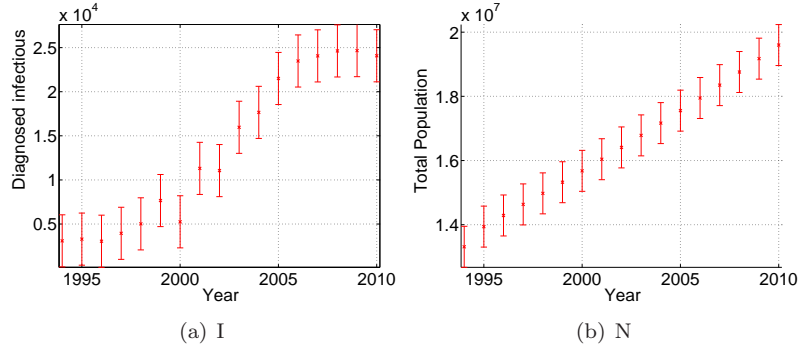


Figure 3: Data from WHO representing the year-by-year trend and variability in yearly case reports over the period 1994-2010.

In fact, Figure 2(b) shows that γ , p_2 and ρ are among the 9 parameters with lowest subcondition, i.e. they can be identified from the given measurements. Fortunately, for some of the non-identifiable parameters, values are already specified for Cameroon or can be derived from other sources, as demonstrated in the following.

5.4 Parameter values of the model

Most demographic parameters are well-known for a given population, others are estimated from the WHO data. In the following, we specify the values for Cameroon.

The natural mortality μ : It is postulated to be equal to the inverse of the life expectancy at birth, which is now about 54.1 years in Cameroon [AH06, oS10]. We fix $\mu = 1/53.6$ per year.

The recruitment Λ : According to NIS [oS10], the average recruitment in the Cameroonian population during the last fifteen year is $\Lambda = 679685$ per year (fixed).

TB mortality d_1 , d_2 and d_3 of undiagnosed infectious and lost sight: Per capita TB-induced mortality rate is 0.193 per year in developed countries, but could be as high as 0.45 per year in some African countries [BPL98]. We fixed the yearly TB-induced mortality rates $d_1 = 0.193$, $d_2 = 0.413$ and $d_3 = 0.20$ TB active cases.

Transmission rates β_i , $i = 1, 2, 3$: Blower et al. (cf. [BPL98]) estimated the contact rates $\beta_i \in [1, 4]$ in the case of a frequency dependent force of infection. In this case, fixed values $\beta_1 = 1$, $\beta_2 = 4$ according to the data of Blower et al. have been used, and we estimated $\beta_3 = 6.05681 \cdot 10^{-06}$ using POEM.

Progression rate parameters p_1 , p_2 and k : For HIV-negative TB people, Bacaer et al. (cf. [BOP⁺08, CFD00]) estimated that people in a South Africa township have 11% annual risk of developing primary TB disease during five years following the first Mtb infection and a 0.03% annual risk of reactivation after five years. In Cameroon, the estimated average TB prevalence for all forms in HIV-positive is about 431 per 100,000 per year. Starting with this order of magnitude, we estimated that $p_1 = 9.36432 \cdot 10^{-04}$, $p_2 = 2.43736 \cdot 10^{-02}$ and $k = 3.31390 \cdot 10^{-04}$ per year. Due to the limited access to hospitals, p_2 is expectedly larger than p_1 and k .

Factors σ_1 and σ_2 : Sutherland et al. [SSR92] estimated that a previous Mtb infection reduces the risk of disease after reinfection by 63% for HIV negative males and by 80% for HIV negative females. We estimate that $\sigma_1 = 2.38390 \cdot 10^{-04}$ and we use the formula from [BOP⁺08] to set $\sigma_2 = 0.7 * (p_1 + p_2)$.

Detection rate h : According to WHO data, $h \in [0.5, 0.9]$ per year. Using POEM, it has estimated $h = 8.28248 \cdot 10^{-01}$ per year.

Diagnosed rate θ : WHO estimated $\theta \in [0.3, 0.6]$ per years. The result obtained with POEM shows that the model is highly sensitive to θ , which was finally estimated from the data as $\theta = 0.495896$ per year.

Proportions r_1 and r_2 of successful treatments: Since the chemoprophylaxis is not practiced in Cameroon, we took $r_1 = 0$ per year and fixed $r_2 = 0.758821$ per year according to [NWF⁺05].

Rate α at which diagnosed infectious population become lost sight: It has been estimated using POEM as $\alpha = 0.216682$ per year.

Rate δ at which lost sight return to the hospital: According to the data of TB in Cameroon, $\delta = 39\%$ per year. This value is fixed in the model.

Natural recovery rate ρ : In [BOP⁺08], the authors estimated that the natural recovery for HIV-negative TB and HIV-positive TB cases are 0.139 and 0.24 per year, respectively. Herein, we took the average of these values as initial guess for the Gauss-Newton algorithm and estimated $\rho = 0.131140$ per year.

Recovery rate ω : We estimated $\omega = 0.5$ per year.

Relapse rate of recovered individuals: The average relapse rate of recovered individuals is estimated as $\gamma = 0.0851257$ per year.

Numerical values of all parameters are summarized in Table 2.

Comparison to measurement data

In order to illustrate the theoretical results of the foregoing analysis, numerical simulations of model system (2) are carried out using a fourth order Runge-Kutta scheme in the software Matlab, version R2009. The total population of Cameroon in 1994 is given by $N = 13,240,337$ [oS10]. The initial values of the other variables were set as in Table 1.

Using the aforementioned data, model (2) gives a very good fit to the Cameroonian data for the period 1994-2010 [WHO12a], as depicted in Figure 4. The agreement is not perfect, but there are no evident consistent patterns in the discrepancy. Forward solutions of the deterministic model follow fairly well the observed TB patterns of incidence. With the estimated transmission parameters, the deterministic model appears to capture all the qualitative properties of the observed pattern. Hence, model (2) can be used to gain realistic insight into tuberculosis transmission dynamics at least for a limited period.

6 Effects of increased access to treatment

Herein, we investigate the impact of the time variation of some specific parameters on the dynamics of model (2). For this purpose, some model parameters are considered as time dependent variables to reflect their possible change within time. However, the variation is assumed to be slow over time.

Effects of increasing the access to TB treatment as a result of infrastructures and education are explored by taking into account the following expressions of model parameters:

$$\begin{aligned} \theta(t) &= \theta + \frac{(1-\theta)t}{\theta_\delta + t}, & \delta(t) &= \delta + \frac{(1-\delta)t}{\delta_\delta + t}, \\ p_1(t) &= p_1 + \frac{p_2 t}{p_\delta + t}, & p_2(t) &= p_2 - \frac{p_2 t}{p_\delta + t}. \end{aligned} \tag{17}$$

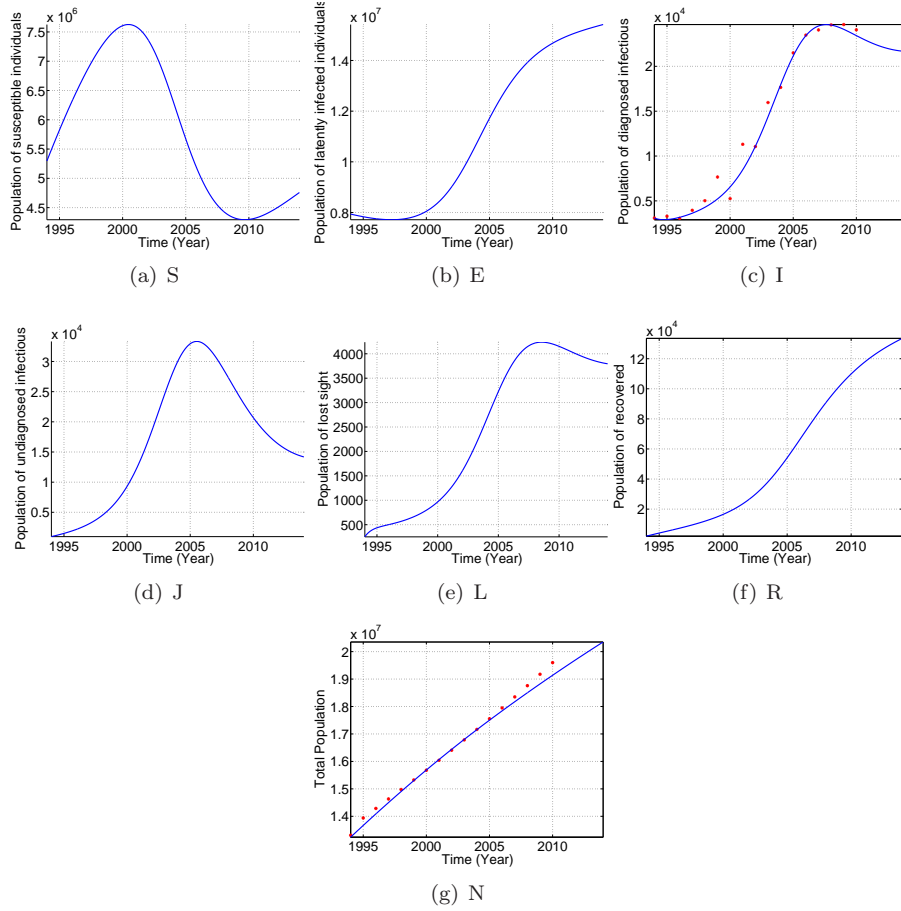


Figure 4: Evolution of model (2) showing the estimated state trajectories of susceptible, latently infected, diagnosed infectious, undiagnosed infectious, lost sight, recovered and total population classes. The dot plots represent the year-by-year trend and variability in yearly case reports over the period 1994-2010. Parameter values are defined in Table 2 and initial values are presented in Table 1.

These parameters are assumed to be control functions for the dynamics of TB. Thus, model (2) becomes a non-autonomous controlled system. Herein, θ_δ , δ_δ and p_δ are positive constants to be estimated. The increase (or decrease) in these parameters can be interpreted as the result of change in treatment access, diagnosis campaign or large scale education via social networks, TV, radio etc.

The goal now is to find two different sets of values for parameters θ_δ , δ_δ and p_δ such that the following two scenarios can be achieved:

- a) a reduction in the population of undiagnosed infectious (J) and lost sight (L) by 20% until 2025,
- b) a reduction in the population of undiagnosed infectious (J) and lost sight (L) by 60% until 2025.

To estimate the parameter values, we generated two artificial data sets. Both data sets contain the values of J and L from the previous simulation (Figure 4) at the adaptively chosen time-points in the time interval [1994, 2035]. However, the values

in the time interval [2010, 2035] have been reduced by 20% and 60% to obtain the first and second data set, respectively. These data sets were then used in the Gauss-Newton algorithm to obtain the following estimates:

a) reduction by 20%: $p_\delta = 8.56043 \cdot 10^7$, $\theta_\delta = 81.2807$, $\delta_\delta = 37.2240$

b) reduction by 60%: $p_\delta = 85.6660$, $\theta_\delta = 81.3256$, $\delta_\delta = 37.1301$

The simulation results are presented in Figures 5 and 6. They show the dynamic of TB inside the population in the presence and absence of continuous effort to diagnose the population. In particular, a relatively small increase in the access to TB treatment could generally result in a decrease in the number of susceptible (S), diagnosed infectious (I), undiagnosed infectious (J), lost sight (L), and recovered (R) individual, and an increase in the number of latently infected (E) individuals. We also observe that the number of diagnosed infectious (I) increases at the beginning, but decreases after a few years prior to the beginning of the control strategies. Thus, TB can be reduced within 15 years if some efforts are made to increase the treatment access for rural population, and TB prevention and education for fast and immuno-compromised people.

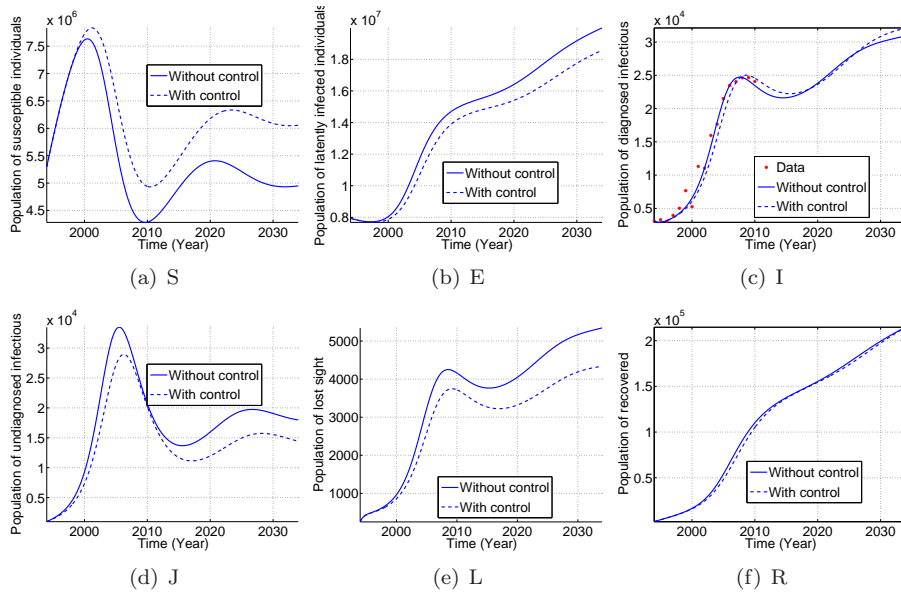


Figure 5: Time series of model (2) showing the impact of a slow change on parameter values θ , δ , p_1 and p_2 with respect to time in order to reduce the TB burden by 20% within 15 years. (plain lines) present the model predictions for TB dynamics using parameter values of Table 2 and the (dashed lines) present the estimated trajectories for parameters θ , δ , p_1 and p_2 set as in equation (17). Parameter identification with artificial data gave $p_\delta = 8.56043 \cdot 10^7$, $\theta_\delta = 81.2807$, $\delta_\delta = 37.2240$. All other parameters are defined as in Table 2.

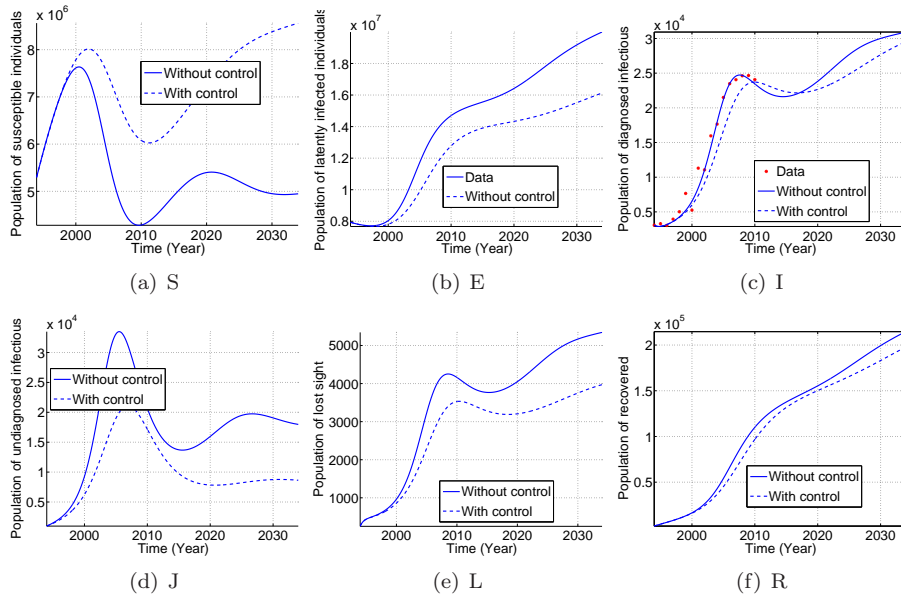


Figure 6: Impact of a slow change of parameter values θ , δ , p_1 and p_2 with respect to time in order to reduce the TB burden by 60% within 15 years. Model predictions (plain) for TB dynamics using parameter values of Table 2 and the estimated trajectories (dashed) for parameters θ , δ , p_1 and p_2 set as in equation (17). Parameter identification with artificial data gave $p_\delta = 85.6660$, $\theta_\delta = 81.3256$, $\delta_\delta = 37.1301$.

7 Conclusion

In this paper, a deterministic model for the transmission dynamics of TB in sub-Saharan Africa has been presented. The objective is to determine the role of TB diagnosis, treatment, lack of information about the epidemiological status of some people, and the role of traditional medicine and natural recovery on the dynamics of TB. In contrast to other TB models in the literature, the model includes three infective classes emanating from diagnosed infectious, undiagnosed infectious, and lost sight individuals. The undiagnosed and lost sight subclasses are shown to be of particular importance in TB modelling in developing countries like sub-Saharan Africa where public health is under-developed. Model parameters have either been fixed according to data published in literature, or they have been estimated with a Gauss-Newton method using data published by WHO and the National Institute of Statistics of Cameroon (NIS). A sensitivity analysis revealed that parameters representing the proportion of individuals having access to medical facilities have a large impact on the dynamics of the disease. We showed that a change in these parameters over time can significantly reduce the disease burden in the population within the next 15 year. These parameters can be used to measure the success of educational and diagnosis campaigns that encourage individuals to go for TB screening. In future work, optimal control strategies could be applied to determine the optimal dynamics of these parameters in order to achieve the highest possible reduction of TB in shortest time at low costs. In addition, the model might be extended towards the inclusion of co-infection between TB and HIV, or the resistance to treatment.

8 Acknowledgement

The first author wishes to thank Dr. S. Bowong for his advices on the model construction and the IMU Einstein foundation for funding the project.

References

- [ACCC02] J.P. Aparicio, A. F. Capurro, and C. Castillo-Chavez. Markers of disease evolution: the case of tuberculosis. *J. Theo. Biol.*, 215:227–237, 2002.
- [AH06] J. P. Aparicio and J.C. Hernandez. *Preventive Treatment of Tuberculosis through Contact Tracing. Mathematical Studies on Human Disease Dynamics: Emerging paradigms and Challenges*. AMS Contemporary Mathematics Series, 2006.
- [AM91] R. M. Anderson and R. May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, 1991.
- [BBB⁺02] M. Begon, M. Bennett, R. G. Bowers, N. P. French, S. M. Hazel, and J. Turner. A clarification of transmission terms in host-microparasite models : numbers, densities and areas. *Epidemiol. Infect.*, 129:147–153, 2002.
- [BD02] S M Blower and C L Daley. Problems and solutions for the Stop TB partnership. *Lancet Infect Dis*, 2(6):374–6, June 2002.
- [BGMZ08] C. Bhunu, W. Garira, Z. Mukandavire, and M. Zimba. Tuberculosis transmission model with chemoprophylaxis and treatment. *Bulletin of Mathematical Biology*, 70(4):1163–1191, 2008. 10.1007/s11538-008-9295-4.
- [BOP⁺08] Nicolas Bacaër, Rachid Ouifki, Carel Pretorius, Robin Wood, and Brian Williams. Modeling the joint epidemics of TB and HIV in a South African township. *J Math Biol*, 57(4):557–93, October 2008.
- [BPL98] S. M. Blower, T. C. Porco, and T. M. Lietman. Tuberculosis: The evolution of antibiotic resistance and the design of epidemic control strategies. In M. A. Horn, G. Simonett, and G. F. Webb, editors, *Mathematical Models in Medical and Health Science*. Vanderbilt University Press and Nashville, 1998.
- [BSH96] S. M. Blower, P. Small, and P. Hopewell. Control strategies for tuberculosis epidemics: new method for old problem. *Science*, 273:497–500, 1996.
- [BTT⁺11] Yanina Balabanova, Vladimir Tchernyshev, Igor Tsigankov, Svetlana Maximova, Natalya Mikheeva, Ljudmila Fedyukovitch, Sergey Kuznetsov, Ivan Fedorin, and Francis Drobniowski. Analysis of undiagnosed tuberculosis-related deaths identified at post-mortem among hiv-infected patients in russia: a descriptive study. *BMC Infectious Diseases*, 11(1):276, 2011.
- [CCF97] C. Castillo-Chavez and Z. Feng. To treat or not to treat: the case of tuberculosis. *J. Math. Biol.*, 35:629–635, 1997.

- [CFD00] C.P. Chaulk, M. Friedman, and R. Dunning. Modeling the epidemiology and economics of directly observed therapy in baltimore. *Int. J. Tuberc. Lung. Dis.*, 4:201–207, 2000.
- [CM99] C. Chintu and A. Mwinga. An african perspective of tuberculosis and hiv/aids. *Lancet*, 353:997–1005, 1999.
- [Deu04] P. Deuffhard. *Newton Methods for Nonlinear Problems: Affne Invariance and Adaptive Algorithms*, volume 35. Springer Series in Computational Mathematics. Springer Verlag, Berlin, 2004.
- [DH03] P. Deuffhard and A. Hohmann. *Numerical Analysis in Modern Scientific Computing. An Introduction*. Springer, New York., second edition, 2003.
- [DS80] P. Deuffhard and W. Sautter. On rank-deficient pseudoinverses. *Lin. Alg. Appl.*, 29:91–111, 1980.
- [DWR11] T. Dierkes, M. Wade, U. Nowak, and S. Röblitz. Bioparkin-biology-related parameter identification in large kinetic networks. arXiv:1303.4928, December 2011.
- [FCCC00] Z Feng, C Castillo-Chavez, and A F Capurro. A model for tuberculosis with exogenous reinfection. *Theor Popul Biol*, 57(3):235–47, May 2000.
- [Fer67] S. Ferebee. An epidemiological model of tuberculosis in the united states. *Bulletin of the National Tuberculosis Association*, 53:4, 1967.
- [Het00] Herbert W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42:599–653, 2000.
- [IMST07] A. Iggidr, J. Mbang, G. Sallet, and J. J. Tewa. Multi-compartment models. *Discrete Contin. Dyn. Syst. Ser. B*, 1:506–519, 2007.
- [KAFM07] Vinay Kumar, Abul K. Abbas, Nelson Fausto, and Richard N. Mitchell. *Basic Pathology (8th ed.)*, volume ISBN 978-1-4160-2973-1. Saunders Elsevier, 2007.
- [KS08] Jean Claude Kamgang and Gauthier Sallet. Computation of threshold conditions for epidemiological models and global stability of the disease-free equilibrium (DFE). *Math Biosci*, 213(1):1–12, May 2008.
- [MBE11] D.P. Moualeu, S. Bowong, and Y. Emvudu. Global properties of a tuberculosis model with n latents classes. *JAMI*, 29(5-6):1097–116, 2011.
- [MBE12] D. P. Moualeu, S. Bowong, and Y. Emvudu. Analysis of the impact of diabetes on the dynamics transmission of tuberculosis. *Mathematical modeling of natural phenomena*, 7(03):117–146, 2012.
- [MMNB11] D. P. Moualeu, J. Mbang, R. Ndoundam, and S. Bowong. Modeling and analysis of hiv and hepatitis c co-infections. *Journal of Biological Systems*, 19(04):683–723, 2011.
- [MSK03] B.M. Murphy, B. H. Singer, and D. Kirschner. On the treatment of tuberculosis in heterogeneous populations. *J. Theor. Biol.*, 223:391–404, 2003.

- [MSR90] C.J.L. Murray, K. Styblo, and A. Rouillon. Tuberculosis in developing countries: burden, intervention, and cost. *Bull. Int. Union Tuberc. Lung Dis.*, 65:6–24, 1990.
- [Mur02] J. D. Murray. *Mathematical biology: I. an introduction*. Springer-Verlag, 2002.
- [NWF⁺05] P. Nunn, B. G. Williams, K. Floyd, C. Dye, G. Elzinga, and M. C. Raviglione. Tuberculosis control in the era of hiv. *Nat. Rev. Immunol.*, 5:819–826, 2005.
- [oFAT01] National Comity of Fight Againts Tuberculosis. *Guide de personnel de la santé*. NCFIT, 2001.
- [OK07] D. Okuonghae and A. Korobeinikov. Dynamics of tuberculosis: The effect of direct observation therapy strategy (dots) in nigeria. *Math. Mod. Nat. Phen.*, 2:101–113, 2007.
- [oS10] National Institute of Statistics. Rapport sur la présentation des résultats définitifs. Technical report, Bureau Central des Recensements et des Etudes de Population, 2010.
- [RH04] M. G. Roberts and J. A. P. Heesterbeek. Mathematical models in epidemiology. Encyclopedia of Life Support Systems (EOLSS), 2004.
- [RKL⁺07] Wood R, Middelkoop K, Myer L, Grant AD, Whitelaw A, Lawn SD, Kaplan G, McIntyre J Huebner R, and Bekker LG. Undiagnosed tuberculosis in a community with high hiv prevalence: implications for tuberculosis control. *Am J Respir Crit Care Med*, 175(1):87–93, Jan. 2007.
- [RLF67] C S ReVelle, W R Lynn, and F Feldmann. Mathematical models for the economic allocation of tuberculosis control activities in developing nations. *Am. Rev. Respir. Dis.*, 96(5):893–909, November 1967.
- [SSR92] I. Sutherland, E. Svandova, and S. Radhakrishna. The development of clinical tuberculosis following infection with tubercle bacilli. *Tubercle*, 63:255–268, 1992.
- [TDB09] J.J. Tewa, J.L. Dimi, and S. Bowong. Lyapunov functions for a dengue disease transmission model. *Chaos, Solitons and Fractals*, 39:936–941, 2009.
- [Thi03] H. R. Thieme. *Mathematics in Population Biology*. Princeton University Press, Princeton, NJ, 2003.
- [vdDW02] P. van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosc.*, 180:29–28, 2002.
- [VF97] E. Vynnycky and P. E. M. Fine. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. *Epidemiol. Infect.*, 119:183–201, 1997.
- [Waa68] H T Waaler. The economics of tuberculosis control. *Tubercle*, 49:Suppl:2–4, March 1968.

- [WGA62] H WAALER, A GESER, and S ANDERSEN. The use of mathematical models in the study of the epidemiology of tuberculosis. *Am J Public Health Nations Health*, 52:1002–13, June 1962.
- [WHO12a] WHO. Fact sheet number 104. Technical report, World Health Organization, Geneva, Switzerland, 2012.
- [WHO12b] WHO. Global tuberculosis control: surveillance, planning, financing. Technical report, World Health Organization, Geneva, Switzerland, 2012.

A Initial values

Symbol	Initial value	Source	Symbol	Initial value	Source
S	5576135	Estimated	E	8357382	Estimated
I	3092	Fixed (WHO)	J	1037	Estimated
L	251	Estimated	R	2140	Estimated
N	13240337	Fixed [oS10]			

Table 1: Initial values of state variables of the TB model.

B Estimated parameter values

Parameters	Symbol	Estimate/yr	Source
Recruitment rate of susceptible	Λ	679685	Fixed, [oS10]
Transmission rate	β_1, β_2	1, 4	Fixed [BSH96]
Transmission rate	β_3	$6.05681 \cdot 10^{-06}$	Estimated
Fast route to infectious class	p_1	$9.36432 \cdot 10^{-04}$	Estimated
Fast route to undiagnosed infectious class	p_2	$2.43736 \cdot 10^{-02}$	Estimated
Reinfection parameter of latently infected individuals	σ_1	$2.38390 \cdot 10^{-04}$	Estimated
Reinfection parameter of recovered individuals	σ_2	$0.7 * (p_1 + p_2)$	Fixed, [BOP+08]
Slow route to active TB	k	$3.31390 \cdot 10^{-04}$	Estimated
Natural mortality	μ	1/53.6	Fixed, [BSH96, oS10]
TB mortality of diagnosed infectious	d_1	0.139	Fixed,[BSH96]
TB mortality of undiagnosed infectious	d_2	0.413	Fixed
TB mortality of lost sight	d_3	0.20	Fixed
Chemoprophylaxis of latently infected individuals	r_1	0	Fixed, [NWF+05]
Detection rate of active TB	h	0.828248	Estimated
Recovery rate of diagnosed infectious	r_2	0.758821	Fixed, [NWF+05]
Recovery rate of lost sight	ω	0.5	Fixed
Recovery rate of undiagnosed infectious	ρ	0.131140	Estimated
Relapse of recovered individuals	γ	$8.51257 \cdot 10^{-02}$	Estimated
Diagnosed infectious route to the lost sight class	α	0.216682	Estimated
Lost sight route to the diagnosed infectious class	δ	0.39	Fixed
Diagnosed rate	θ	0.495896	Estimated

Table 2: Estimated numerical values of the TB model parameters