Impact of tuberculosis treatment length and adherence under different transmission intensities

S. T. R. Pinho[a,*], P. Rodrigues[b], R. F. S. Andrade[a], H. Serra[c], J.S. Lopes[d], M. G. M. Gomes[d]

[a]*Instituto de Física, Universidade Federal da Bahia, Campus Universitário de Ondina, 40210-340, Salvador, Brazil*
[b]*Departamento de Matemática and Centro de Matemática e Aplicações, Universidade Nova de Lisboa, Quinta da Torre, 2829-516, Caparica, Portugal*
[c]*Faculdade de Educação, Universidade Federal da Grande Dourados, Rodovia Dourados, Itahum, km 12, 79804-070 - Dourados, Brazil*
[d]*Instituto Gulbenkian de Ciência, Oeiras, Portugal*

Abstract

Tuberculosis (TB) is a leading cause of human mortality due to infectious disease. Treatment default is a relevant factor which reduces therapeutic success and increases the risk of resistant TB. In this work we analyse the relation between treatment default and treatment length along with its consequence on the disease spreading. We use a stylized model structure to explore, systematically, the effects of varying treatment duration and compliance. We find that shortening treatment alone may not reduce TB prevalence, especially in regions where transmission intensity is high, indicating the necessity of complementing this action with increased compliance. A family of *default functions* relating the proportion of defaulters to the treatment length is considered and adjusted to a particular dataset. We find that the epidemiological benefits of shorter treatment regimens are tightly associated with increases in treatment compliance and depend on the epidemiological background.

Keywords:
tuberculosis, treatment, default, reinfection, mathematical model

*Corresponding author, Tel: + 5571 32836640; fax: + 5571 32357730
Email addresses: suani@ufba.br (S. T. R. Pinho), pcpr@fct.unl.pt (P. Rodrigues), randrade@ufba.br (R. F. S. Andrade), hiraldoserra@ufgd.edu.br (H. Serra), jmlopes@igc.pt (J.S. Lopes), ggomes@igc.pt (M. G. M. Gomes)

1. Introduction

Tuberculosis (TB), a communicable disease transmitted by *Mycobacterium tuberculosis* (*Mtb*), is the second leading cause of death worldwide [1], with approximately two million deaths per year around the world. The basis of the World Health Organization (WHO) TB control programs recommends a treatment strategy that consists on a 6- to 8-month regimen, termed directly observed treatment (DOTS) [2]. Treatment default remains a challenge to the success of this strategy, contributing to treatment failure and increasing the risk of multidrug-resistant TB [3]. Under DOTS programs, the proportion of individuals dropping out of treatment can range from 6% to 30% [4, 5], with available data indicating that defaulters more often leave treatment in its later stages [4]. Current thesis asserts that an improvement of the patient health conditions soon after treatment initiation associated with long treatment duration is an important cause of default. Because of this, some efforts have been placed to reduce the length of treatment based on new short-term efficient drugs [6, 7].

Mathematical models have been proposed to analyse the effect of novel therapies with the focus on treatment shortening, taking into account treatment default [8, 9, 10]. However, the relation between treatment length and proportion of defaulters, and its impact on disease spreading has not been fully addressed in these models.

In this work we focus on the influence of treatment length and treatment default on TB dynamics and on the possible relation between these two processes. We assume that individuals while in treatment cannot be reinfected and do not contribute to disease spread. This leads to a conflicting situation whereby shortening treatment is beneficial to TB control due to reduction in default while, simultaneously, it reduces the time during which individuals are protected against subsequent reinfections. Whether the net effect is beneficial or detrimental is the question that we address herein by means of mathematical models.

This study is systematized by first analyzing the impact of shortening treatment independently of changes in the proportion of default, and then recognizing that variations in treatment length can also induce a change in treatment adherence and analyzing the combined effects of both changes. There are many works reporting data on TB treatment default, but just a few in relation to treatment shortening. Examples include timing of default for cities in 10 countries [4], default after re-treatment in India (2006) [11],

factors predicting treatment adherence in Southern Ethiopia (2002-2004) [3], and default during diagnosis in Pakistan (2008) [12]. TB compliance is a complex multifactorial phenomenon that has motivated many qualitative studies compiled in review articles, e.g. [13]. Quantitative studies are fewer and include the analysis of TB default in Uzbekistan (2005) [14] and in the Netherlands (1993-1997) [15]. These data are typically analyzed by calculating the cumulative probability of non-adherence using Kaplan-Meier survival curves. However, the rate of defaults over time has limited information regarding the effect of treatment length on the rate of default. TB surveillance systems that are in place in many countries are likely to contain this information.

Surveillance data on TB in Portugal for the period of 2002-2009 was provided by the Portuguese National Directorate of Health (Portuguese TB Surveillance System – SVIG-TB). These data, collected by medical practitioners, consists of predicted treatment length, varying from 2 to 12 months, and treatment outcome of well-identified patients, which distinguishes the following possibilities: i) treatment completed, with or without laboratory confirmation; ii) treatment failure; iii) treatment default, i.e., lost to follow up; iv) death from tuberculosis or other causes. From the comparison of default rates for different predicted treatment lengths, we can analyze the dependence between these variables. Although available information on the interdependency between default rates and treatment lengths is restricted to three different durations (2, 6 and 12 months), the data consistently indicates an inverse relationship between these quantities. Therefore we set up a family of default functions capable of describing this trend. Function paramaters are estimated by fitting to the Portuguese data. The function is then introduced into a TB transmission model to assess the epidemiological consequences of the estimated relation.

In contrast to other models [8, 9, 10], we use a minimal mathematical model with few compartments and parameters. Aspects that are mainly important when TB dynamics are confronted with data for structured populations (such as age, gender, social conditions) or other epidemiological aspects (such as smear status) are not detailed herein. We discuss how changes in treatment protocols might impact the dynamics of TB, considering both TB prevalence (proportion of infectious individuals) and the basic reproductive number R_0 as outcomes. We report results for a wide range of contact rates and focus on general trends, hoping to motivate future research to assess specific settings in more detail conditional to data availability.

This work is organized as follows: In Section 2 we introduce the model, in

context with others in the literature: in Subsection 2.1 we present the system of nonlinear differential equations, and the model is analyzed in Subsection 2.2. In Section 3, the results are presented in terms of the impact of reducing treatment length on TB prevalence and on R_0: in Subsection 3.1 the default and treatment length are independent from each other; in Subsection 3.2 we consider a family of functions that couples default and treatment length. Finally, in Section 4, we present the discussion and concluding remarks.

2. Tuberculosis transmission and treatment model

The mathematical model for TB transmission considered herein extends the well known SIR and SIS systems by incorporating latency, the effect of treatment, dropout from it (treatment default), and reinfection with partial immunity. Our purpose is to describe the non-negligible treatment default that is relevant in some already quoted regions of the world [4], together with the processes of relapse and reinfection.

We chose a minimal model, that is able to take into account the quoted effects, with five distinct compartments, which correspond to proportions of the population in the following groups:

S - **Susceptible**: individuals that have never been in contact with *Mtb*.

P - **Primary**: individuals that have been infected by *Mtb* but for whom disease progression is still uncertain.

L - **Latent**: infected individuals that have contained the infection and remain latent.

I - **Infectious**: individuals that have progressed to active TB and are not yet in treatment.

T - **Under Treatment**: individuals that presented active TB and are currently submitted to treatment.

The dynamics of transitions among these compartments are based on the following assumptions, as illustrated by a diagram in Figure 1:

a - All individuals have the same death rate μ except for **I** individuals who, in general, may suffer additional TB-related mortality, this is $\mu_I \geq \mu$.

b - **S** individuals become **P** infected due to the contact between **S** and **I** individuals with effective contact rate β. To keep the population constant in time, the birth rate compensates deaths $b = \mu(1-I)+\mu_I I$.

c - **P** state results from the contact process described above, but also from the contact between **L** and **I** individuals with rate $\sigma\beta$ due to reinfection, where $\sigma \in [0, 1]$ refers to partial protection to reinfection. We assume that **P** individuals are not infectious. A fraction ϕ of individuals in **P** progress to active disease and the remaining consolidate as latently infected, at rates $\delta\phi$ and $\delta(1-\phi)$, respectively.

d - **L** individuals are latently infected. They are able to contain infection, whether this has never taken the form of active disease or has been active and cured, either self-cured directly from class **I** at rate τ_{SC} or after successful treatment. Besides the reinfection process described before, **L** individuals are subject to relapse at a constant rate ω.

e - **I** individuals are responsible for the disease spreading. They have active disease requiring treatment to reestablish normal health conditions. We assume that **I** individuals can leave the compartment when no longer spread disease, after being detected and start treatment to class **T** at a constant rate τ, or by self-cure at a constant rate τ_{SC} to class **L**.

f - Individuals in class **T** are under treatment. They no longer contribute to disease spread and are protected from reinfection. Individuals leave this compartment at rate δ_T, depending on treatment outcome: a proportion ϕ_T defaults back to the infectious state and the remaining progress to the latent state as a result of successful treatment.

These assumptions are formalized by a system of ordinary differential equations (ODE) as follows:

$$\begin{cases} \dfrac{dS}{dt} &= \mu(1-I) + \mu_I I - \beta I S - \mu S \\ \dfrac{dP}{dt} &= \beta I(S + \sigma L) - (\delta + \mu)P \\ \dfrac{dL}{dt} &= (1-\phi)\delta P + (1-\phi_T)\delta_T T + \tau_{SC} I - \sigma\beta I L - (\omega + \mu)L \\ \dfrac{dI}{dt} &= \phi\delta P + \omega L + \phi_T \delta_T T - (\tau + \tau_{SC} + \mu_I)I \\ \dfrac{dT}{dt} &= \tau I - (\delta_T + \mu)T \end{cases} . \quad (1)$$

The proposed model is derived from a previous formulation [16], with modifications as follows. In order to address specifically treatment length and default we introduce compartment **T** in the current model to represent individuals undergoing treatment. Furthermore, we do not consider any treatment actions for **P** and **L** individuals, as we are not assessing the effect of latent treatment in this study. Note that, as before, we assume that even the patients that conclude treatment and become cured are considered to be latent **L**, since *Mtb* is not completely eliminated from the organism [17]. It is important to note the distinction between relapse and reinfection: the first process is described by a direct transition from **L** to **I** at a constant rate, while the second it is proportional to βI and it requires a first transition to class **P** where it follows the same dynamics used to describe the new infections.

A specificity of this model is that patients, while in treatment (on class **T**), are neither infectious nor subject to reinfection. Hence, class **T** acts as a protected class when compared to the others. Previous models have different assumptions, considering that individuals under treatment can be partially infectious [8] or can be subject to reinfection at a reduced rate [9]. Our assumption can be relaxed to include these cases, as we comment on these model differences and their consequences in forthcoming sections. Another distinguishing feature of our model is to formulate the average treatment duration ($1/\delta_T$) independently of the proportion of treatment defaulters (ϕ_T), while the average time spent in treatment (class **T**) is different for defaulters ($1/\phi_T\delta_T$) and successfully treated individuals ($1/(1-\phi_T)\delta_T$). This allows us to explore the consequences of changing independently the recommended duration of treatment and the fraction of defaulters under that treatment

schedule. The rate of default is then a combination of these two, $\phi_T \delta_T$. Alternatively, other authors have considered two different rates from **T** to **I** (rate of default) and from **T** to **L** (rate of successful treatment) [8, 9]. Although either formulation involves two parameters, the interpretation seems more direct as performed here.

Model parameters along with their typical values used herein are listed in Table 1. Our main focus is on the parameters related to treatment, specifically the average treatment length and the proportion of individuals who drop out of treatment. These parameters will be varied to explore different scenarios. Sensitivity to some of the other parameters will also be provided, with all results presented for a realistic range of transmission intensities (represented by β in the model). In particular, we have concluded that the results are quite insensible to the influence of self cure as well as TB-related death. Therefore from now on, we consider $\tau_{SC} = 0$ and $\mu_I = \mu$.

2.1. Basic analysis of the model

System (1) has one disease free equilibrium, $E_0 = (1, 0, 0, 0, 0)$ and one endemic equilibrium, $E_1 = (S_1, P_1, L_1, I_1, T_1)$. The local stability analysis of E_0 indicates that a transcritical bifurcation occurs when

$$\beta_c = \frac{\mu(\mu+\delta)[(\mu+\tau+\delta_T)(\mu+\omega)+\delta_T\tau(1-\phi_T)]}{\delta(\omega+\phi\mu)(\mu+\delta_T)}. \tag{2}$$

Indeed, for $\beta < \beta_c$, E_1 does not exists in the positive cone and E_0 is stable, while for $\beta > \beta_c$, E_0 becomes unstable and E_1 exists as a stable equilibrium. The critical value of the effective contact rate β_c corresponds to the epidemic threshold, i.e., the minimal transmission potential required for the persistence of infection.

The same result can be express in terms of the basic reproduction number R_0, considering that the threshold condition for endemicity corresponds to $R_0 = 1$, with the disease dying out if $R_0 < 1$, and becoming endemic if $R_0 > 1$. Applying to model (1) the next generation matrix approach as in [18], we obtain

$$R_0 = \frac{\beta\delta(\omega+\phi\mu)(\mu+\delta_T)}{\mu(\mu+\delta)[(\mu+\tau+\delta_T)(\mu+\omega)+\delta_T\tau(1-\phi_T)]}. \tag{3}$$

The coordinates of E_1 are shown in Appendix A; I_1 is the positive real root of the polynomial that exists for $R_0 > 1$, when the disease becomes

endemic. Due to the relevant effect of reinfection on our analysis, we calculate the reinfection threshold (RT) as in [16, 19, 20], further details of the calculation are provided in Appendix B. The RT is attained at

$$R_0 = \frac{1}{\sigma}\frac{\omega+\phi\mu}{\phi\mu}\frac{(\mu+\tau+\delta_T)\mu+\delta_T\tau(1-\phi_T)}{(\mu+\tau+\delta_T)(\mu+\omega)+\delta_T\tau(1-\phi_T)}(=R_0^{RT}). \tag{4}$$

Note, in particular, that for very small values of ω this is close to the classical RT attained at $R_0 = 1/\sigma$. Moreover, when σ is close to zero the value of R_0 at which RT is attained goes to infinity ($R_0^{RT} \to \infty$).

3. Evaluation of different treatment regimens

Our aim is to explore how disease spread is affected by the shortening of treatment regimens, whether this comes alone or linked to a change in the default proportion. The two key parameters are the inverse of treatment length and the proportion of defaulters, respectively represented in the model by δ_T and ϕ_T. The analysis is set to investigate how this impact depends on intensity of transmission as captured by the baseline R_0, the composite parameter given by (3). Although ranges of R_0 are obtained by varying the effective contact rate β in the model, we chose to present the results in terms of R_0 to benefit from its dimensionless nature and clear interpretation. As the value of R_0 itself is subjected to change as new treatment regimens are introduced, we denote its baseline value by R_0^{base} and its value under new treatments by R_0^{new}.

We assume baseline values for δ_T and ϕ_T and assess how modifying these parameters affects TB prevalence. We define a measure of effectiveness E as

$$E \equiv 1 - \frac{I_{new}}{I_{base}}, \tag{5}$$

where I_{base} is the proportion of infectious individuals at equilibrium with the baseline parameter values, while I_{new} indicates the equilibrium proportion of infectious individuals after the parameters have changed to describe a new treatment regimen. Note that the values of E depend on R_0 through I_{base} and I_{new}, as it is shown in Appendix A. Moreover, note that for a fixed I_{base}, effectiveness E is maximum when the intervention is able to eliminate the disease ($I_{new} = 0$) and is null or negative when intervention does not reduce TB incidence ($I_{new} \geq I_{base}$).

Results are presented in two subsections. First, the analysis is performed in terms of the default proportion and inverse treatment length (ϕ_T and δ_T), by varying one parameter and keeping the other fixed. Second, we set up a family of functions relating these parameters $\phi_T(\delta_T)$, termed as default functions. We then perform the analysis of E taking into account the proposed functional dependence. For both cases, we point out how the results are related to the reinfection threshold.

3.1. Sensitivity to default proportion and treatment length

By assigning different values to δ_T and ϕ_T independently, it is possible to analyze how the dynamics of TB transmission depends on the treatment lenght and default parameters when these are assumed to be uncorrelated. In particular, we focus on the effects of changing treatment regimen as manifested as follows: (i) change in disease prevalence, measured by the effectiveness function E, in terms of R_0^{base}; (ii) dependence of R_0 expression on treatment parameters ϕ_T and δ_T.

Figure 2(a) comparatively shows prevalence curves for active TB with baseline parameters (bold gray), reduced treatment length (dashed black), and reduced default proportion (solid black). For illustration, we choose as baseline values $1/\delta_T = 1/2$ yr and $\phi_T = 0.05$, according to [21]. As expected, a decrease in treatment default leads to reduction in active TB prevalence for any value of R_0^{base}. Reducing the treatment length, on the other hand, tends to increase active TB prevalence for higher values of R_0^{base}. This behavior is better observed in Figure 2(b) that shows the corresponding effectiveness functions E on R_0^{base}. When treatment default is reduced (solid line), E is positive and peaks near the reinfection threshold (vertical dotted line), when the reduction of TB prevalence is more noticeable. Shortening treatment, on the contrary, leads to negative values of E above the reinfection threshold (dashed line), indicating that an increase in TB prevalence may result.

The relevant role of reinfection is evident when we compare panels (a) and (b) ($\sigma = 0.5$) with panels (c) and (d) of Figure 2, for which there is no reinfection ($\sigma = 0$). In this latter case, although reducing treatment default still reduces TB prevalence, there is no longer a peak on the effectiveness function, and E becomes a monotonic decreasing function (solid black lines in panels (c)-(d)). Furthermore, when only treatment length is reduced there is virtually no change on disease prevalence and $E \approx 0$ (dashed lines). This suggests that assumptions about protection from reinfection while individuals

are under treatment are likely to be important determinants for the outcomes of alternative treatment regimens.

Consistent results are obtained if we analyze the dependence of R_0 expression on treatment parameters ϕ_T and δ_T. These can be confirmed by inspection on the following partial derivatives:

$$\frac{\partial R_0}{\partial \phi_T} = \Gamma_0 \delta_T (\mu + \delta_T), \tag{6}$$

$$\frac{\partial R_0}{\partial \delta_T} = \Gamma_0 (\phi_T \mu + \omega), \tag{7}$$

where

$$\Gamma_0 = \frac{\beta\delta(\omega + \phi\mu)\tau}{\mu(\mu+\delta)\left[(\mu+\omega)(\delta_T+\tau+\mu)+\tau\delta_T(1-\phi_T)\right]^2}.$$

Both derivatives are positive, since all model parameters are positive. The positive sign in (6) clearly corresponds to the expectations of reducing R_0 as ϕ_T is reduced. However, the positive sign in (7) is puzzling, as it suggests that more time under treatment (smaller δ_T) would be preferable.

We can understand better these results by looking into the interpretation of R_0 in terms of the contribution of individuals under treatment to the average infectious period. These individuals are not infectious but they can go back to the infectious state without further contact with infectious individuals: directly by default or by relapse after successful recovery (see Figure 1). In Appendix C, we show that the derivatives of R_0 (equations (6) and (7)) are related to these two phenomena (see (C.2) and (C.3)). Hence, we can show that when decreasing ϕ_T, there is a shift from the decreasing contribution through default to the increasing contribution through reactivation. Overall, the infectious period decreases since the first effect is stronger, decreasing R_0, which explains the positive sign in (6). However, when treatment length decreases (increasing δ_T), there is an increase in the contribution of both paths, as the time spent in **T** is shorter, increasing the infectious period, which leads to the positive sign in (7) (see further details in Appendix C).

In order to obtain a more informative approach to discuss the effect of treatment length and default, the next section considers a dependence between the two parameters ϕ_T and δ_T [3] and analyzes the combined effects.

3.2. Dependence between default proportion and treatment length

Now we assume that the proportion of defaulters is not independent of the prescribed treatment length, and proceed to investigate how a default function $\phi_T(\delta_T)$ leads with the conflicting effects described above.

Let us consider the total derivative of R_0 in relation to δ_T

$$\frac{dR_0}{d\delta_T} = \Gamma_0[(\omega + \mu\phi_T) + (\mu + \delta_T)\delta_T\phi_T'], \tag{8}$$

where $\phi_T' = \dfrac{d\phi_T}{d\delta_T}$. Treatment shortening leads to a reduction in R_0 when the right hand side of (8) has a negative sign, now occurring when

$$\phi_T' < -\frac{\omega + \mu\phi_T}{(\mu + \delta_T)\delta_T}, \ \forall\delta_T > 0, \tag{9}$$

which is to say that a decrease in the treatment length can only lead to a decrease in R_0 if combined with a sufficiently large decrease on the default proportion.

From these considerations we now impose a positive impact on R_0, i.e., $dR_0/d\delta_T < 0$ to set up a family of default functions. By integrating both sides of (9) from a certain $\delta_T^* > 0$ for which $\phi_T(\delta_T^*) = \phi_T^*$ until $\delta_T > \delta_T^*$, we obtain

$$\phi_T(\delta_T) \leq B + \frac{C}{\delta_T} \tag{10}$$

with $B = (\delta_T^*\phi_T^* - \omega)/(\mu + \delta_T^*)$ and $C = \delta_T^*(\omega + \mu\phi_T^*)/(\mu + \delta_T^*)$. This inequality provides a more concrete idea of the many possible functions that meet condition (9). We want to consider functions that also verify the following conditions:

(i) $\phi_T(\delta_T \leq \delta_1) = 1$;

(ii) $\lim\limits_{\delta_T \to \infty} \phi_T = 0$.

The first condition sets up δ_1 as the maximum value until which the proportion ϕ_T assumes its maximum value 1, and the second assures that instantaneous treatments result in no default. Based on the expression obtained in (10) we consider a simple family of default functions

$$\phi_T(\delta_T) = \min\left\{1, \frac{\delta_1 - \Delta}{\delta_T - \Delta}\right\}. \tag{11}$$

From now on, we study the impact on TB prevalence by modifying the treatment length and proportion of defaulters according to this default function, which obeys both inequality (10) and conditions (i)-(ii). Other functions obeying these conditions would lead to similar results but, in the interest of concreteness, we proceed to estimate parameters δ_1 and Δ by fitting the default function (11) to data obtained from the Portuguese TB Surveillance System (SVIG-TB) for the period 2002-2009. From this dataset, which included predicted treatment length and outcome of TB patients (Table 2), we were able to extract default proportions of 5.99%, 4.80% and 2.44% for prescribed treatment lengths of 12, 9 and 6 months, respectively. The rate of treatment incompleteness (refereed to as rate of default in this study) was calculated by pooling together cases of default and treatment failure. Note, in particular, that the proportion of defaulters decreases with the prescribed treatment length, consistently with the family of default functions adopted here.

The most common predicted treatment lengths are 6, 9 and 12 months. Using only these three points makes the analysis of $\phi_T(\delta_T)$ somewhat limited. Nevertheless, we used a Gauss-Newton algorithm to calculate the curve that best-fits the data according to the least-squares criteria. The values of the parameters that minimize the residual sum of squares are of $\delta_1 = 0.2188$ yr^{-1} and $\Delta = 0.1677$ yr^{-1}. Using the parameterized default function, we calculated a predicted default rate of 6.14%, 4.38%, 2.79% and 0.88% for treatment lengths of 12, 9, 6 and 2 months, respectively. Figure 3 shows the resulting function with parameter values $\delta_1 = 0.2188$ yr^{-1} and $\Delta = 0.1677$ yr^{-1}. The small number of data points in Figure 3 may rise a conjecture that they could be adjusted by a simple linear function with a negative slope. However, we emphasize that satisfying the condition expressed by equation (9) was the adopted guideline to obtain the explicit functional dependence between ϕ_T and δ_T. We refer to this question in more detail in Appendix E.

Figure 4 shows an analysis based on TB prevalence and effectiveness function E as previously presented in Figure 2, now changing simultaneously treatment duration and treatment default according to the parameterized default function and shown in Figure 3. Four treatment lengths and corresponding treatment defaults are considered: 2, 6, 9 and 12 months with 0.88%, 2.79%, 4.38% and 6.14% treatment default, respectively. We compare effectiveness of a decrease in treatment length from 12 to 9, 9 to 6 and from 6 to 2 months (dotted, dashed and solid lines in panel (b)). The reduction in treatment length and treatment default simultaneously leads to a slight

decrease in TB prevalence for low to moderate transmission levels. The positive impact on the reduction of TB prevalence decreases from the reinfection threshold (dotted vertical line) to higher values of R_0^{base}. For high transmission levels, there is a value of R_0^{base} above which the change in treatment regimen results in a slight increase on the disease prevalence. This negative impact for higher transmission rates becomes more evident when looking to the effectiveness function ($E < 0$ in panel (b)). When transmission is sufficiently intensive, even the increase in treatment adherence and consequent decrease of the directly infected proportion of the population, cannot compensate the overall enhancement of TB dissemination by reinfection, due to faster recovery from treatment.

As we stressed before, we chose a particular function that decreases R_0 and that corresponds to a positive correlation between decrease of treatment duration and proportion of default, according to the Portuguese dataset. A more general picture is drawn in Figure 5, where we superimpose the contours of equilibrium TB prevalence (per 100,000, $I(\delta_T, \phi_T) \times 100,000$), on the parameter space (δ_T, ϕ_T), with the particular default function (gray curves, $\phi_T(\delta_T)$). Below the reinfection threshold (panel (a)), the contours are almost flat so even a slightly decreasing function would have the benefit of reducing disease prevalence. Above the reinfection threshold a steeper default function would be required to this effect (panel (b)). As as example, for $R_0^{base} = 3.2$ (referring to $\beta = 250$ yr^{-1}) a decrease from six to two months treatment (from $\delta_T = 2$ yr^{-1} to $\delta_T = 6$ yr^{-1}) would lead to an increase in prevalence for the parameterized default function.

Our results depend on the assumption that individuals under treatment (class **T**) are in a pseudo-quarantine state, they are neither infectious nor susceptible to reinfection. Shortening this state has the potential to increase TB prevalence, specially in regions where transmission intensity is already high (above the reinfection threshold). In Appendix D, we extend our model in order to accommodate different assumptions on infectiousness/susceptibility of individuals under treatment. If they are assumed partially infectious, following Salomon *et al.* [8], shortening treatment becomes more beneficial in terms of reducing disease prevalence, but the impact still decreases for hight transmission where reinfection becomes more frequent. If class **T** is subject to reinfection (with the same partial protection as latent individuals), as in Abu Raddad *et al.* [9], then the detrimental effect of treatment shorten-

ing in terms of increasing disease prevalence for higher endemicities is not as pronounced, since both classes **T** and **L** are subject to reinfection. If treatment duration and proportion of defaulters are reduced simultaneously as informed by the Portuguese dataset, treatment shortening only results in increased disease prevalence for very high transmission scenarios. The qualitative behavior is the same as that described in Figure 4(b).

Other modeling studies have considered other values for δ [21] as well as non-zero self cure (τ_{SC}) and TB-related death ($\mu_I - \mu$) [22, 23, 24]. Sensitivity to these parameters shows that the qualitative behavior described in this paper is preserved (Appendix E).

4. Discussion and conclusions

The epidemiological consequences of introducing shorter regimens to treat active tuberculosis have been previously addressed by elaborated mathematical models [8, 9]. Here we use a minimal model structure where only the main disease mechanisms are present, to gain mathematical tractability while keeping comparability with more detailed studies. We explore, systematically, the effects of varying treatment length and compliance under different epidemiological scenarios as represented by the basic reproduction number (R_0^{base}).

Consistent among the various studies is the conviction that shorter treatment regimens are likely to be associated with increased compliance, and this duality is likely to reduce TB prevalence. We suggest, in this work, a family of *default functions* to make evidence of this duality. We use data from TB surveillance in Portugal to set up parameter values. Other datasets, such as one from the Netherlands that refers to different treatments lengths [15], may also be considered and most likely lead to variation in parameter estimates. What is unique to the present study is that we do not necessarily assume an increase in compliance as an effect of shortening prescribed treatment length but, instead, these two parameters can be varied independently. We find that shortening treatment alone is not expected to reduce TB prevalence, indicating the necessity of complementation with increased compliance. Intuitively, this follows from the fact that reducing only treatment duration and keeping the fraction of defaulters fixed, implies an increased rate out of the treatment class **T** into the latent class **L** by successfull treatment ($(1-\phi_T)\delta_T$), but also increases an increase in the rate of defaulting, from **T** to **I** ($\phi_T\delta_T$). Implicit in previous model formulations is that only the rate from **T** to **L** is increased

in a controlled manner by reducing treatment duration, making it difficult to systematically trace the default fraction (see [8, 9]).

Another advantage of our analysis is the possibility of exploring different epidemiological scenarios instead of focusing on a particular region. By varying the values of our model parameter β (linearly related to R_0^{base}) we analyze these effects for low, moderate and high transmission intensities as well as how they change near the reinfection threshold [20]. This finding has important practical relevance given the wide range in tuberculosis transmission intensities across the globe, with some regions below and others above the reinfection threshold [21] as informed by analysis of molecular typing data [25]. Our results show that the effectiveness of changing treatment length and/or default proportion is highly sensitive to the epidemiological background.

One important assumption we make is that individuals in compartment $\mathbf{T}$ are neither infectious nor subject to reinfection, which is not an universal assumption in the tuberculosis literature. We relax this assumption (details in Appendix E) and consider the cases for which individuals under treatment are partially susceptible to reinfection as in [9] or partially infectious as in [8]. The qualitative behavior is not changed but the effectiveness of reducing treatment length is enhanced when considering that individuals under treatment are infectious until the end of treatment.

We summarize our results based on active TB prevalence and effectiveness function E in three cases:

a) Reducing treatment default, for fixed treatment length, leads to positive effectiveness (reduction in disease prevalence) for low, moderate and high transmission intensities, with the highest impact near the reinfection threshold.

b) Reducing treatment length, for fixed treatment default, leads to negative effectiveness (increase in disease prevalence) for high transmission intensities (above the reinfection threshold), as this shortens the period of time when individuals are under treatment and therefore not subject to reinfection.

c) Assuming a combined scenario with a decreasing default function $\phi_T(\delta_T)$ coupling treatment default and treatment length, whose parameters are based on data of treatment default for three different treatment length (12, 9 and 6 months), shortening treatment may also have negative effectiveness (increase disease prevalence) especially in regions where

transmission intensity is already high (above the reinfection threshold). Results are, in principle, robust to the particular choice of a default function that adjusts the trend of decreasing default with treatment length found in the data.

While the results relative to reducing treatment default are intuitive, those concerning the reduction of treatment length require further investigation into the assumptions regarding the characteristics of individuals under treatment. Shortening treatment alone may not reduce the burden of active TB under high transmission rates, but this is sensitive to whether a reduction in prescribed treatment length brings with it a reduction in the default proportion. This is also depending on whether individuals under treatment contribute to the transmission dynamics, either by being capable of infecting or being reinfected by others. Clinical data is necessary for the development of a smoother description of the transitions into and out of treatment, as the effects of changing treatment regimens are so sensitive to different assumptions that seem equally plausible and pave the literature.

Acknowledgements

The authors are grateful to M. L. Barreto and C. Rebelo for critical reading of the manuscript. The authors are also grateful to S. Pereira, J. M. Marinho and R. Duarte for helpful discussions, and to A. Fonseca-Antunes for providing data on default rates in Portugal used for parameterizing the default function. This work was partially supported by Fundação de Amparo à Pesquisa do Estado da Bahia - FAPESB (contract PNX 0006/2009), Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (contract 410498/2006-8), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES (contract AUX-PE-FCT 1171/2009), Fundação Ciência e Tecnologia (FCT) and the European Commission (grant EC-ICT-231807). It is part of INCT-Citecs (contract no. 57386/2008-9) and INCT-SC funded by the National Institutes of Science and Technology Programme (MCT-CNPq, Brazil). PR is supported by the Fundação para a Ciência e Tecnologia through PEst-OE/MAT/UI0297/2014 (CMA) and EXPL/MAT-CAL/0794/2013. HS is supported by FAPESB through BOL0143/2008.

Tables

Table 1: Model parameters

symbol	definition	value	**Reference/Comment**
β	effective contact rate	77.3-386.6 yr^{-1}	referring to R_0 (1-5) [21]
μ	death rate for noninfectious individuals	$1/80$ yr^{-1}	
μ_I	death rate for infectious individuals	$1/80$ yr^{-1}	assumed $\mu_I \approx \mu$
δ	inverse of primary infection length	12 yr^{-1}	[16]
ϕ	fraction of *Mtb* infected population developing active TB (the $1-\phi$ fraction remain latent)	0.05	[27]
σ	factor of exogenous reinfection for latent	0.5	[25]
ω	rate of endogenous reactivation for latent infections	0.0003 yr^{-1}	[17]
τ	rate at which infectious individuals enter treatment	6 yr^{-1}	[21]
τ_{SC}	rate of self-cure	0 yr^{-1}	neglected self-cure
δ_T	inverse of treatment length	2 yr^{-1} or varying	[21]
ϕ_T	whole fraction of individuals that drop out treatment (the $1-\phi_T$ fraction are successfully treated)	0.05 or varying	[21]

Table 2: Outcome of TB treatment in Portugal for different predicted treatment lengths.

predicted δ_T	$1\ \mathrm{yr}^{-1}$	$4/3\ \mathrm{yr}^{-1}$	$2\ \mathrm{yr}^{-1}$
completed	877 (83.44%)	3546 (88.67%)	5722 (93.16%)
failure	4 (0.38%)	11 (0.28%)	4 (0.07%)
defaults	59 (5.61%)	181 (4.53%)	146 (2.38%)
death	111 (10.56%)	261 (6.53%)	270 (4.40%)
total	1051 (100.00%)	3999 (100.00%)	6142 (100.00%)

Figure captions

Figure 1: Tuberculosis model. Individuals are classified according to disease history into susceptible (**S**), primary infected (**S**), latent (**L**), infectious (**I**), or under treatment (**T**). Parameters are described in Table 1. The force of infection is represented by $\beta I(t)$, where β is the efficient contact rate.

Figure 2: Dependence of equilibrium infectious proportion I **(a)** on R_0. The heavy gray curve corresponds to the baseline scenario with treatment duration of six months ($\delta_T = 2$ yr^{-1}) and treatment failure of 5% ($\phi_T = 0.05$). Dashed and thin solid black curves correspond to a change in treatment duration to two months ($\delta_T = 6$ yr^{-1}) and in treatment failure to 1% ($\phi_T = 0.01$), respectively. **(b)** Dependence of the effectiveness function E on R_0 when decreasing treatment length from six to two months (dashed black curves) and when decreasing treatment default from 5% to 1% (light full black line). The reinfection threshold for the baseline values of parameters ϕ_T and is marked by the vertical dotted line. A similar analysis of Figures **(a)** and **(b)** assuming $\sigma = 0$ is presented in **(c)** and **(d)** respectively.

Figure 3: Default function $\phi_T(\delta_T)$, given by (11). Using the Portuguese dataset, the parameter values for the default function (11) are: $\delta_1 = 0.2188$ yr^{-1} and $\Delta = 0.1677$ yr^{-1}. Data on treatment default for predicted treatment length are marked by open circles $\circ$.

Figure 4: **(a)** Equilibrium curve for the proportion of infectious: light solid curve, bold gray curve, dashed and dotted curves correspond to treatment duration of 2, 6, 9 and 12 months and treatment default of 0.88%, 2.79.15%, 4.38% and 6.14%, respectively, according to the default function given by (11) with parameter values $\delta_1 = 0.2188$ yr^{-1} and $\Delta = 0.1677$ yr^{-1}. **(b)** Effectiveness of decreasing treatment length and default proportion from: 12 to 9 months that corresponds to 6.14% to 4.38% (dotted line); 9 to 6 months that corresponds to 4.38% to 2.79% (dashed line); from 6 to 2 months that corresponds to 2.79% to 0.88% (solid line). The reinfection threshold is marked by the vertical dotted line.

Figure 5: Contour plot for the infectious prevalence per 100 000 at equilibrium when varying δ_T (inverse of treatment duration) and ϕ_T (proportion default), for different transmission scenarios $R_0 = 1.9$ and 3.9 (below and above the reinfection threshold). Gray curves correspond to the default function $\phi_T(\delta_T)$ adjusted to data.

Figure 6: Dependence of the effectiveness function E on R_0, when varying the treatment length δ_T, for the three different assumptions on the under treatment class: not infectious neither susceptible to reinfection (solid curves); partially infectious (dashed curves) as in [8], and partially susceptible to reinfection (gray curves) as in [9]. **(a)** Case where treatment duration is reduced from six months to two months and treatment default is 5% ($\delta_T = 2$ yr^{-1} to $\delta_T = 6$ yr^{-1}, $\phi_T = 0.05$) for all three scenarios. **(b)** Effectiveness curves when treatment duration is reduced from six months to two months and treatment default is 5% for different levels of infectiousness: $\theta = 0.25, 0.1, 0.05, 0.01$ and 0. **(c)** Case where treatment length and treatment default are reduced simultaneously, from six months to two months and from 5% to 1%. ($\delta_T = 2$ yr^{-1} to $\delta_T = 6$ yr^{-1} and $\phi_T = 0.05$ to $\phi_T = 0.01$).

Appendix A. Equilibrium equations

System (1) has one disease free equilibrium, $E_0 = (1, 0, 0, 0, 0)$ and one endemic equilibrium, $E_1 = (S_1, P_1, L_1, I_1, T_1)$. The coordinate of E_1 are expressed in terms of the positive root of a second order polynomial

$$a_2(R_0)I_1^2 + a_1(R_0)I_1 + a_0(R_0) = 0, \tag{A.1}$$

with

$$\begin{aligned} a_2(R_0) &= \sigma R_0^2 \frac{AB^2}{C^2} \\ a_1(R_0) &= R_0 \frac{B}{C}\left[\sigma R_0 \frac{B}{C}\delta\phi(\mu+\delta_T) + \mu\sigma A + \frac{B}{\mu}\right] \\ a_0(R_0) &= B(1-R_0) \end{aligned} \tag{A.2}$$

where $A = (\mu+\delta\phi)(\mu+\delta_T+\tau)+\delta_T\tau(1-\phi_T)$, $B = \mu(\mu+\delta)[(\mu+\omega)(\mu+\delta_T+\tau)+\delta_T\tau(1-\phi_T)]$ and $C = \delta(\omega+\phi\mu)(\mu+\delta_T)$ and the basic reproduction number R_0. In (A.1), we stress the dependency of a_i on R_0 as this is the parameter most directly related to disease spreading. Finally, I_1 is the positive real root of the polynomial that exists for $R_0 > 1$, when the disease becomes endemic.

Appendix B. Reinfection threshold

Following Rodrigues [26], the reinfection sub-model, where the population is considered to be partially immunized and infection processes other then reinfection are removed ($\omega = 0$) is represented by the following system of ordinary differential equations (ODE):

$$\begin{cases} \dfrac{dP}{dt} &= \sigma\beta IL - (\delta+\mu)P \\ \dfrac{dL}{dt} &= \mu + (1-\phi)\delta P + (1-\phi_T)\delta_T T - \sigma\beta IL - \mu L \\ \dfrac{dI}{dt} &= \phi\delta P + \phi_T\delta_T T - (\tau+\mu)I \\ \dfrac{dT}{dt} &= \tau I - (\delta_T+\mu)T \end{cases}. \tag{B.1}$$

System (B.1) has one disease free equilibrium, $e_0 = (0, 1, 0, 0)$ and one endemic equilibrium, $e_1 = (P^*, L^*, I^*, T^*)$ where the coordinates can be derived

from the equilibrium equations to be expressed in terms of

$$I^* = \frac{\frac{\beta}{\beta_{RT}} - 1}{\sigma\beta((\mu + \delta\phi)(\mu + \delta_T + \tau) + \delta_T\tau(1 - \phi_T))}. \tag{B.2}$$

The linear stability analysis of e_0 indicates that a transcriptical bifurcation occurs when

$$\beta_{RT} = \frac{1}{\sigma}\frac{(\delta + \mu)[\mu(\delta_T + \tau + \mu) + (1 - \phi_T)\delta_T\tau]}{\phi\delta(\mu + \delta_T)}$$

Indeed, for $\beta < \beta_{RT}$, e_1 does not exists in the positive cone and e_0 is stable, while for $\beta > \beta_{RT}$, e_0 becomes unstable and e_1 exists as a stable equilibrium. The critical value of the effective contact rate β_{RT} corresponds to the invasion threshold [19], i.e., the minimal transmission potential required for the persistence of reinfection and it marks a change in the behavior of the complete model. The reinfection threshold is then attained for $\beta = \beta_{RT}$ at

$$R_0 = \frac{1}{\sigma}\frac{\omega + \phi\mu}{\phi\mu}\frac{(\mu + \tau + \delta_T)\mu + \delta_T\tau(1 - \phi_T)}{(\mu + \tau + \delta_T)(\mu + \omega) + \delta_T\tau(1 - \phi_T)} (= R_0^{RT}).$$

Appendix C. Interpretation of R_0 in relation to treatment

It is reasonable to rewrite R_0 as

$$R_0 = \Gamma_1\frac{1}{(\tau + \mu)}\frac{1}{(1 - A)}, \tag{C.1}$$

where Γ_1 is given by the transmission rate times the average fraction surviving class **P** directly to **I** plus the fraction going to **I** through **L** and surviving the latent period:

$$\Gamma_1 = \beta\Big[\frac{\phi\delta}{\delta + \mu} + \frac{(1 - \phi)\delta}{\delta + \mu}\frac{\omega}{\omega + \mu}\Big].$$

The factor $1/[(\tau + \mu)(1 - A)]$ is equivalent to the average infectious period related to the process ($\mathbf{I} \to \mathbf{T} \to \mathbf{I}$ or $\mathbf{I} \to \mathbf{T} \to \mathbf{L} \to \mathbf{I}$) repeated as many times as possible:

$$\frac{1}{(\tau + \mu)}\frac{1}{(1 - A)} = \frac{1}{\tau + \mu}(1 + A + A^2 + \ldots)$$

with

$$A = \frac{\tau}{\tau + \mu}(A_1 + A_2).$$

Note that A_1 and A_2 take into account the two possible paths from $\mathbf{T}$ back to $\mathbf{I}$:

a) directly by defaulting, when surviving the $\mathbf{T}$ class, $A_1 = \dfrac{\phi_T \delta_T}{\mu + \delta_T}$;

b) by relapsing back to the infectious class after recovering to the latent state, surviving both $\mathbf{T}$ and $\mathbf{L}$ classes, $A_2 = \dfrac{(1-\phi_T)\delta_T}{\mu + \delta_T}\dfrac{\omega}{\omega + \mu}$.

Therefore

$$A = \frac{\tau}{\tau + \mu}\left[\frac{\phi_T \delta_T}{\delta_T + \mu} + \frac{(1-\phi_T)\delta_T}{\delta_T + \mu}\frac{\omega}{\omega + \mu}\right],$$

completing the derivation of (C.1).

Now we rewrite (6,7) as

$$\frac{\partial R_0}{\partial \phi_T} = \Gamma_2 \frac{\partial A}{\partial \phi_T} = \Gamma_2 \frac{\tau}{\tau + \mu}\Big[\frac{\partial A_2}{\partial \phi_T} + \frac{\partial A_1}{\partial \phi_T}\Big], \tag{C.2}$$

$$\frac{\partial R_0}{\partial \delta_T} = \Gamma_2 \frac{\partial A}{\partial \delta_T} = \Gamma_2 \frac{\tau}{\tau + \mu}\Big[\frac{\partial A_2}{\partial \delta_T} + \frac{\partial A_1}{\partial \delta_T}\Big], \tag{C.3}$$

where

$$\Gamma_2 = \Gamma_1 \frac{1}{(1-A)^2}\frac{1}{\tau + \mu}.$$

Now, we can infer the behaviour of R_0 relative to ϕ_T and δ_T by analysing the contribution of individuals in class $\mathbf{T}$ to the infectious period. When treatment default (ϕ_T) decreases, there is a shift from the contribution through default (A_1), which decreases, to the contribution through reactivation, which increases (A_2). We can check by the following derivatives that the first effect is stronger, which results in a lower value of R_0, as :

$$\frac{\partial A_1}{\partial \phi_T} = \frac{\partial}{\partial \phi_T}\left[\frac{\phi_T \delta_T}{\mu + \delta_T}\right] = \frac{\delta_T}{\mu + \delta_T} > 0, \tag{C.4}$$

$$\frac{\partial A_2}{\partial \phi_T} = \frac{\partial}{\partial \phi_T}\left[\frac{(1-\phi_T)\delta_T}{\mu + \delta_T}\frac{\omega}{\omega + \mu}\right] = -\frac{\delta_T}{\mu + \delta_T}\frac{\omega}{\omega + \mu} < 0, \tag{C.5}$$

and

$$\left|\frac{\partial A_2}{\partial \phi_T}\right| < \left|\frac{\partial A_1}{\partial \phi_T}\right|.$$

However, when treatment length decreases, there is an increase in the contribution of both paths, as the time spent in class **T** decreases:

$$\frac{\partial A_1}{\partial \delta_T} = \frac{\partial}{\partial \delta_T}\left[\frac{\phi_T \delta_T}{\mu+\delta_T}\right] = \frac{\mu \delta_T}{(\mu+\delta_T)^2} > 0. \tag{C.6}$$

$$\frac{\partial A_2}{\partial \delta_T} = \frac{\partial}{\partial \delta_T}\left[\frac{(1-\phi_T)\delta_T}{\mu+\delta_T}\frac{\omega}{\omega+\mu}\right] = \frac{\mu(1-\phi_T)}{(\mu+\delta_T)^2}\frac{\omega}{\omega+\mu} > 0, \tag{C.7}$$

Appendix D. The compartment of under treatment individuals in TB models

We consider a more general model that can accommodate different assumptions for the infectiousness/susceptibility of under treatment class **T**, in the literature [8, 9]. We consider the following system of differential equations

$$\begin{cases} \frac{dS}{dt} &= \mu - \lambda S - \mu S \\ \frac{dP}{dt} &= \lambda(S + \sigma L + \sigma_T T) - (\delta+\mu)P \\ \frac{dL}{dt} &= (1-\phi)\delta P + (1-\phi_T)\delta_T T - \sigma\lambda L - (\omega+\mu)L \\ \frac{dI}{dt} &= \phi\delta P + \omega L + \phi_T\delta_T T - (\tau+\mu)I \\ \frac{dT}{dt} &= \tau I - \sigma_T\lambda T - (\delta_T+\mu)T \end{cases}, \tag{D.1}$$

where $\lambda = \beta(I+\theta T)$ for $\theta \in [0,1]$ and $\sigma_T \in [0,1]$.

We consider three possible scenarios:

a) for $\lambda = \beta I$ ($\theta = 0$), $\sigma_T = 0$, individuals under treatment are neither infectious nor susceptible to reinfection, system (D.1) corresponds to a simplified version of system (1), with $\mu_I = \mu$ and $\tau_{SC} = 0$;

b) for $\lambda = \beta I$ and $\sigma_T = \sigma$, individuals under treatment are assumed susceptible to reinfection with partial protection similar to latent individuals as in [9];

c) finally, for $\lambda = \beta(I + \theta T)$ and $\sigma_T = 0$ individuals in class **T** are considered partially infectious, as in [8], parameter $\theta \in [0, 1]$ controls infectiousness of individuals under treatment.

For all scenarios, system (D.1) has two possible equilibria: a disease free and an endemic equilibrium. We compare the endemic equilibria and the correspondent effectiveness function E, under the three assumptions described above. If only the duration of treatment is changed from six ($\delta_T = 2$ yr^{-1}) to two months ($\delta_T = 6$ yr^{-1}) then treatment effectiveness is null or negative when assuming that individuals under treatment are not infectious (see Figure 6 (a) full and dotted lines). However, when individuals in class T are partially infectious we only see this negative impact for very high endemicities. In this case, shorter time in class **T** corresponds to a lower force of infection and consequently to a lower disease level, for low endemic regions. But faster treatment also means a faster change to a latent state where reinfection is possible. So, when reinfection is strong, it can have a negative impact. Figure 6 (b) shows that the behavior described above is enhanced by infectiousness (for higher values of θ).

Finally, if we improve both treatment duration ($\delta_T = 2$ yr^{-1} to $\delta_T = 6$ yr^{-1}) and default ($\phi_T = 5\%$ to $\phi_T = 1\%$) then all three assumptions give similar qualitative results (see Figure 6 (c)). Again, effectiveness is higher for the case where individuals under treatment are partially infectious.

Appendix E. Supplementary material

In the first section, a sensitivity analysis to parameter δ is done to illustrate how the magnitude of our results depend on this parameter.

In second section, a sensitivity analysis to parameters τ_{SC} and μ_I is conducted to justify our choice to simplify the model and consider these parameters $\tau_{SC} = 0$ and $\mu_I = \mu$, respectively.

Due to the small data set analysed, a simple linear fit of the data was performed, in the third section, and the restrictions of the use of this linear function as a default function is further discussed.

The supplementary material here described can be found online at

References

[1] T.R. Frieden, T.R. Sterling, S.S. Munsiff, C.J. Watt, C. Dye C 2003. Tuberculosis. *The Lancet* **362**:887-899.

[2] WHO. 2008. The Stop strategy – Building on and enhancing DOTS to meet the TB-related Millennium Development Goals. WHO/HTM/TB/2006.368

[3] E.B. Shargie, B. Lindtjorn, 2007. Determinants of treatment adherence among smear-positive pulmonary tuberculosis patients in Southern Ethiopia. *PLOS Medicine* **4**; *e*37.

[4] M.E. Kruk, N.R. Schwalbe, C.A. Aguiar 2008. Timing of default from tuberculosis treatment: a systematic review. *Trop. Med. Inter Health* **13**:703-712.

[5] M.S.P.L. Souza, S.M. Pereira, J.M. Marinho, M.L. Barreto. 2009. Characteristics of healthcare services associated with adherence to tuberculosis treatment. *Rev Saúde Pública* **43**: 1-8.

[6] R.J. O'Brien,M. Spigelman, 2005. New drugs for tuberculosis: Current status and future prospects.*Clin Chest Med* **26**:327-340.

[7] TB Alliance 2008. Confronting TB: What it takes, 2008 Annual Report (Global Alliance for TB Drug Development, New York).

[8] J.A. Salomon, J.O. Lloyd-Smith, W.M. Getz, S. Resch, M.S. Sánches, T. C. Porco, M.W. Borgdorff. 2006. Prospects for Advancing Tuberculosis Control Efforts through Novel Therapies. *PLOS Medicine* **3**:1302-1309.

[9] L.J. Abu-Raddad, L. Sabatelli, J.T. Achterberg, J.D. Sugimoto, I.M. Longini, C. Dye, M.E. Halloran. 2009. Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics. *PNAS*, **106**:13980-985.

[10] Christopher Dye, Geoffrey P Garnett, Karen Sleeman, Brian G Williams. 1998. Prospects for worldwide tuberculosis control under the WHO DOTS strategy *The Lancet* **352**:1886-1891.

[11] U.M. Jha, S. Satayanarayana, P.K. Dewan, S. Chadha, F. Wares, S. Sahu, D. Gupta, L.S. Chauhan, 2010. Risk Factors for Treatment Default among Re-Treatment Tuberculosis Patients in India, 2006. *PLOS One* **5**: *e*8873.

[12] M.S. Khan, S. Khan, P. Godfrey-Faussett, 2009. Default during TB diagnosis: quantifying the problem. *Trop. Med. Int. Health* **14**:1437-1441.

[13] S.A. Munro, S.A.Lewin, H.J. Smith, M.E. Engel, A. Fretheim, J. Volmink. 2007. Patient Adherence to Tuberculosis Treatment: A Systematic Review of Qualitative Research *PLOS Medicine* **4**(7):1230-1245.

[14] E. Hasker, M. Khodjikhanov1, S. Usarova, U. Asamidinov, U. Yuldashova, M. J van der Werf, G. Uzakova, J. Veen, 2008. Default from tuberculosis treatment in Tashkent, Uzbekistan; Who are these defaulters and why do they default? *BMC Infectious Diseases* **8**: 97.

[15] M.W. Borgdorff, J. Veen, N.A. Kalisvaart, J.F. Broekmans, N.J.D. Nagelkerke, 2000. Defaulting from tuberculosis treatment in the Netherlands: rates, risk factors and trend in the period 1993-1997. *Eur. Respir. J.* **16**::209-213.

[16] M.G.M. Gomes, P. Rodrigues , F. Hilker, N.B. Mantilla-Beniers, M. Muehlen, A. Paulo, G.F. Medley. 2007. Implications of partial immunity on the prospects for tuberculosis control by post-exposure interventions. *J. Theor. Biol.* **248**: 608-17.

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

[18] P. van den Driessche, J. Watmough. 2002. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission *Math Biosci.* **180**:29-48.

[19] M.G.M. Gomes, L.J. White, G.F. Medley. 2004. Infection, reinfection, and vaccination under suboptimal immune protection: epidemiological perspectives. *J Theor Biol.* **235**(2):151-2.

[20] M.G.M. Gomes, A.O. Franco, M.C. Gomes, G.F. Medley. 2004a. The reinfection threshold promoted variability in tuberculosis epidemiology and vaccine efficacy. *Proc. R. Soc. B* **271** 617-623.

[21] Lopes J.S., Rodrigues P., Pinho S.T.R., Andrade R.F.S., Duarte R., Gomes M.G.M. 2014. Interpreting simple measures of tuberculosis transmission: a case study on the Portuguese population. *BMC Infect Dis.* **14**:340.

[22] Coljin C, Cohen T, Murray M 2007. Emergent heterogeneity in declining tuberculosis epidemics. *J Theor Biol*, **247**:765-774.

[23] Cohen T, Coljin C, Finklea B, Murray M 2007. Exogenous re-infection and the dynamics of tuberculosis epidemics: Local effects in a network model of transmission. *J R Soc Interface*, **4**:523-531.

[24] Dye C, Garnett GP, Sleeman K, Williams BG. 1998. Prospects for worldwide tuberculosis control under the WHO DOTS strategy. *The Lancet.* **352**:1886-91.

[25] M. G. M. Gomes, R. Águas, J. S. Lopes, M. C. Nunes, C. Rebelo, P. Rodrigues and C. J. Struchiner, 2012. How host heterogeneity governs tuberculosis reinfection. *Proc. R. Soc. B* **279**: 2473-2478.

[26] P. Rodrigues. 2009. Modeling tuberculosis: a compromise between biological realism and mathematical tractability. *PhD Thesis*, University of Lisbon.

[27] P.M. Small, M.D. Fujiwara, 2001. Management of tuberculosis in the United States. *N. Engl. J. Med.* **345**(3): 189-200.

Figure1

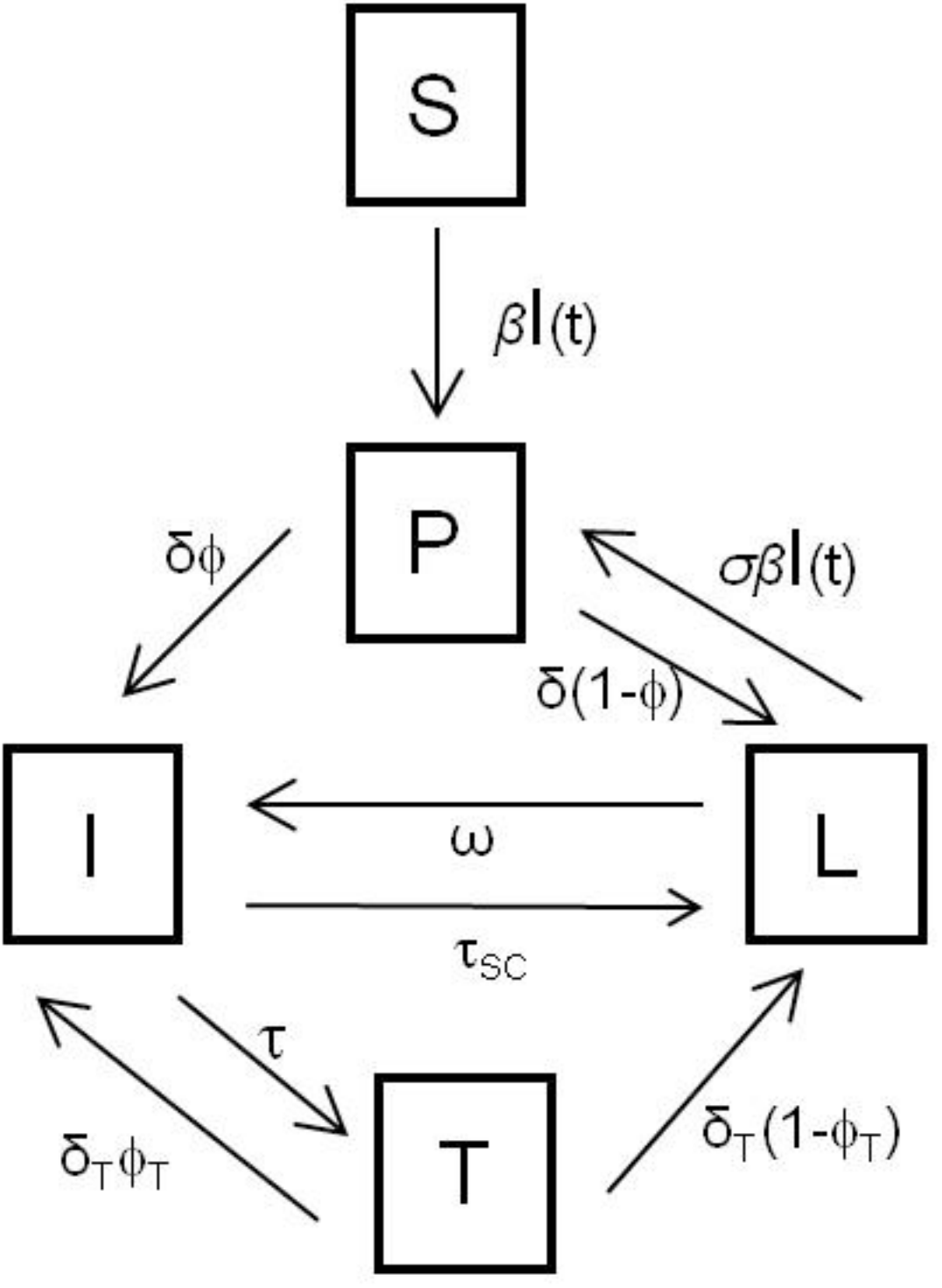


Figure 2

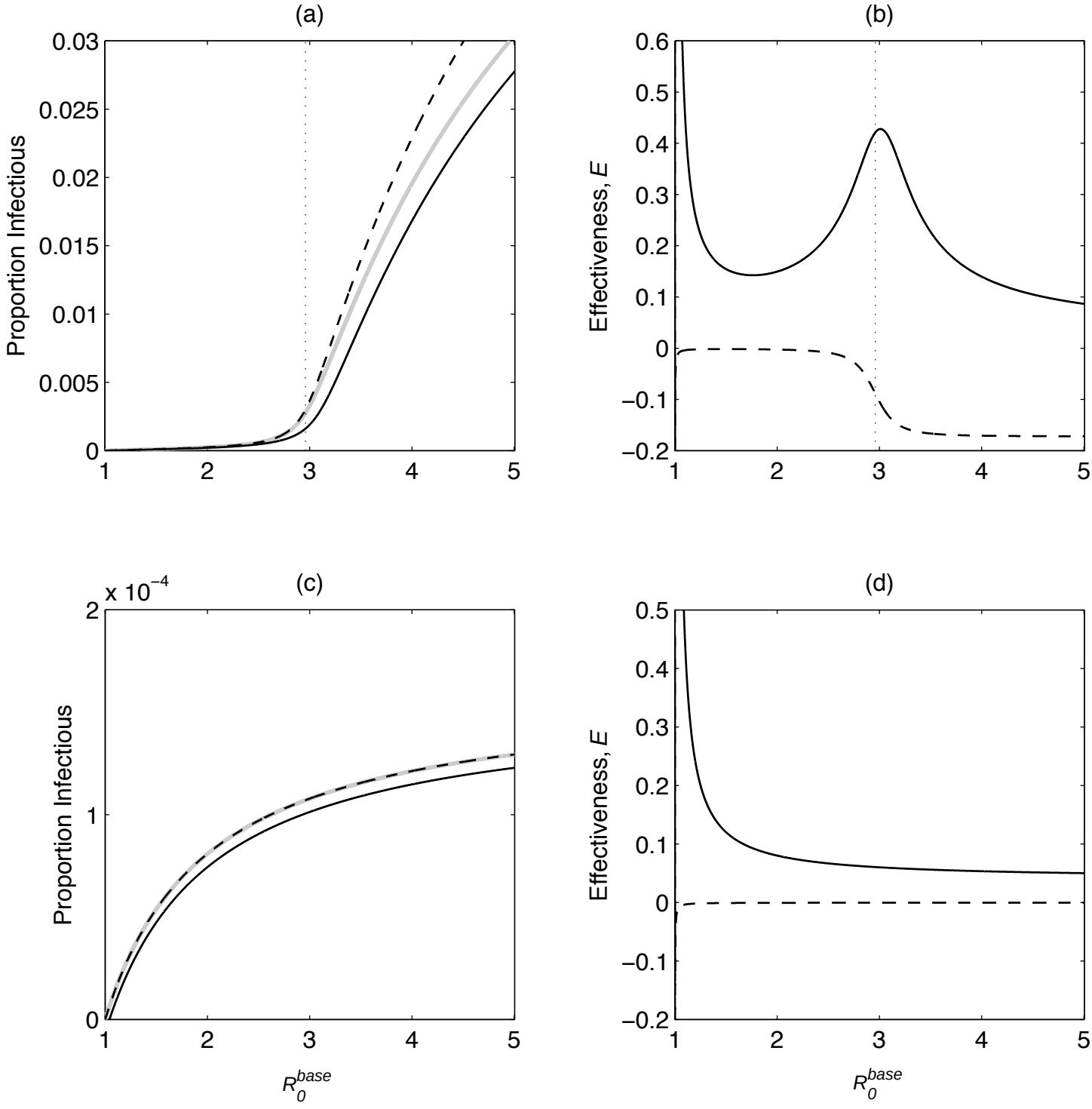

(a)
Proportion Infectious
0.03
0.025
0.02
0.015
0.01
0.005
0
1
2
3
4
5
(b)
Effectiveness, E
0.6
0.5
0.4
0.3
0.2
0.1
0
−0.1
−0.2
1
2
3
4
5
(c)
x 10^−4
Proportion Infectious
2
1
0
1
2
3
4
5
R_0^base
(d)
Effectiveness, E
0.5
0.4
0.3
0.2
0.1
0
−0.1
−0.2
1
2
3
4
5
R_0^base

Figure3

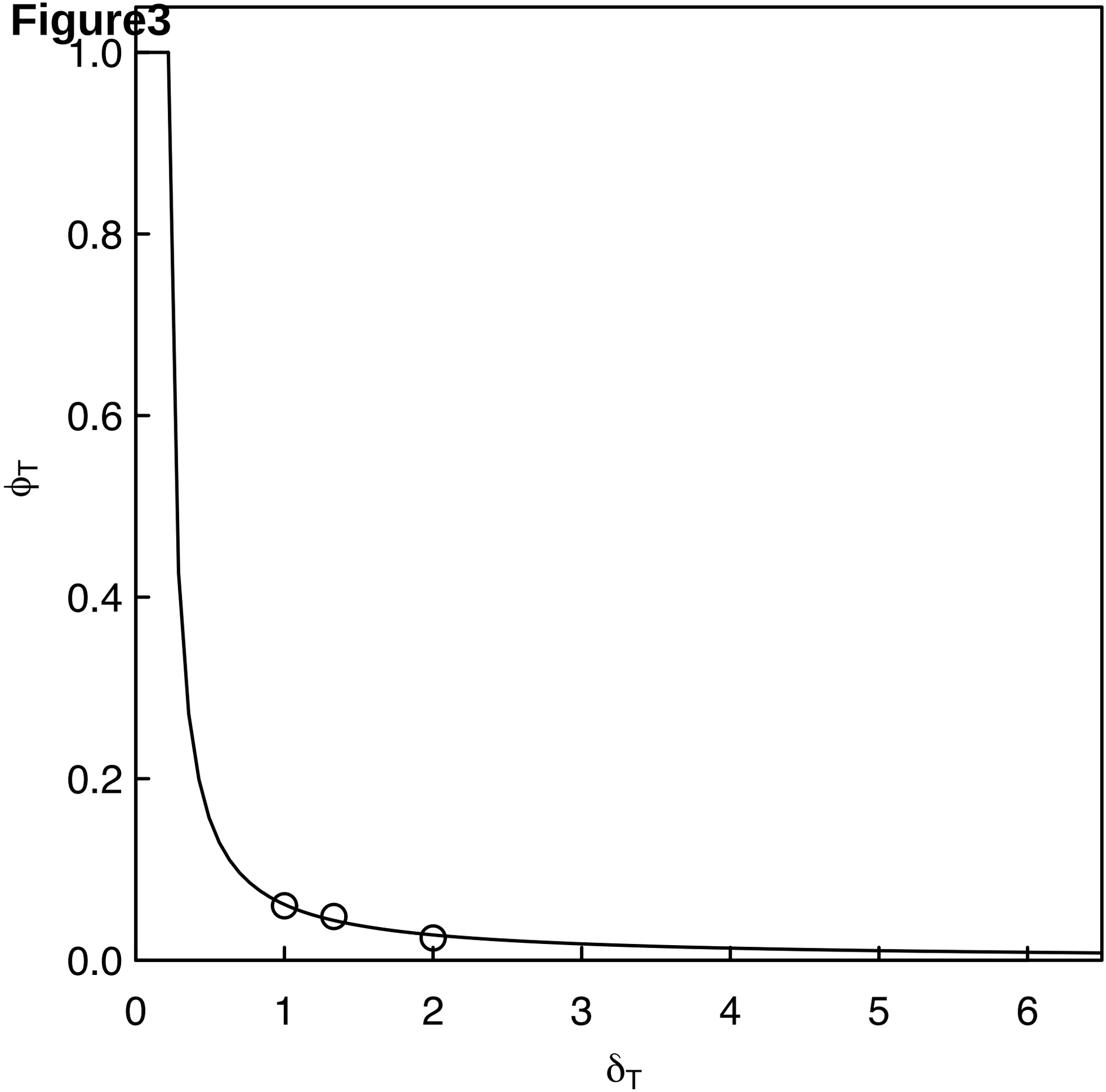


Figure 4

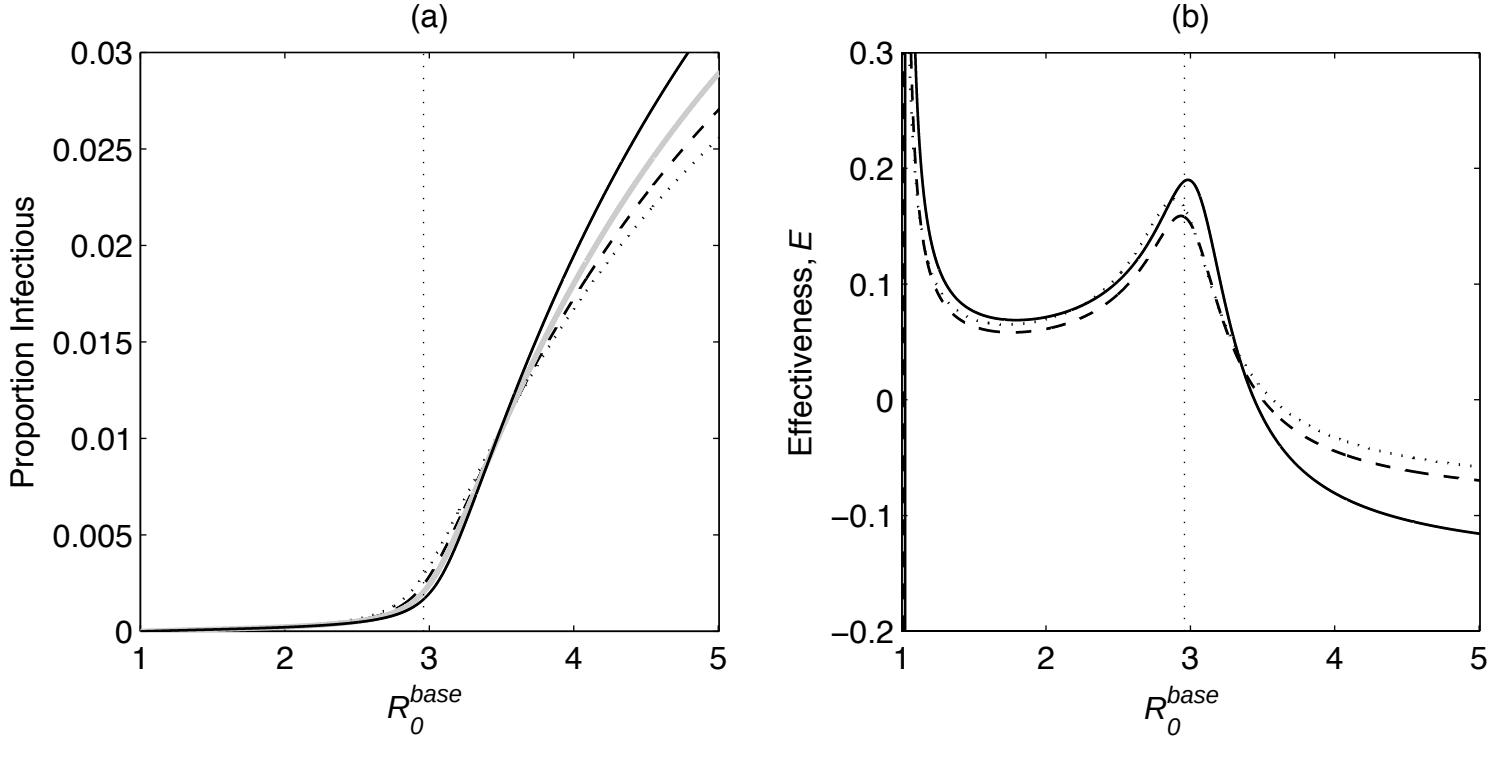

(a)
Proportion Infectious
0.03
0.025
0.02
0.015
0.01
0.005
0
1
2
3
4
5
R_0^{base}
(b)
Effectiveness, E
0.3
0.2
0.1
0
−0.1
−0.2
1
2
3
4
5
R_0^{base}

Figure5

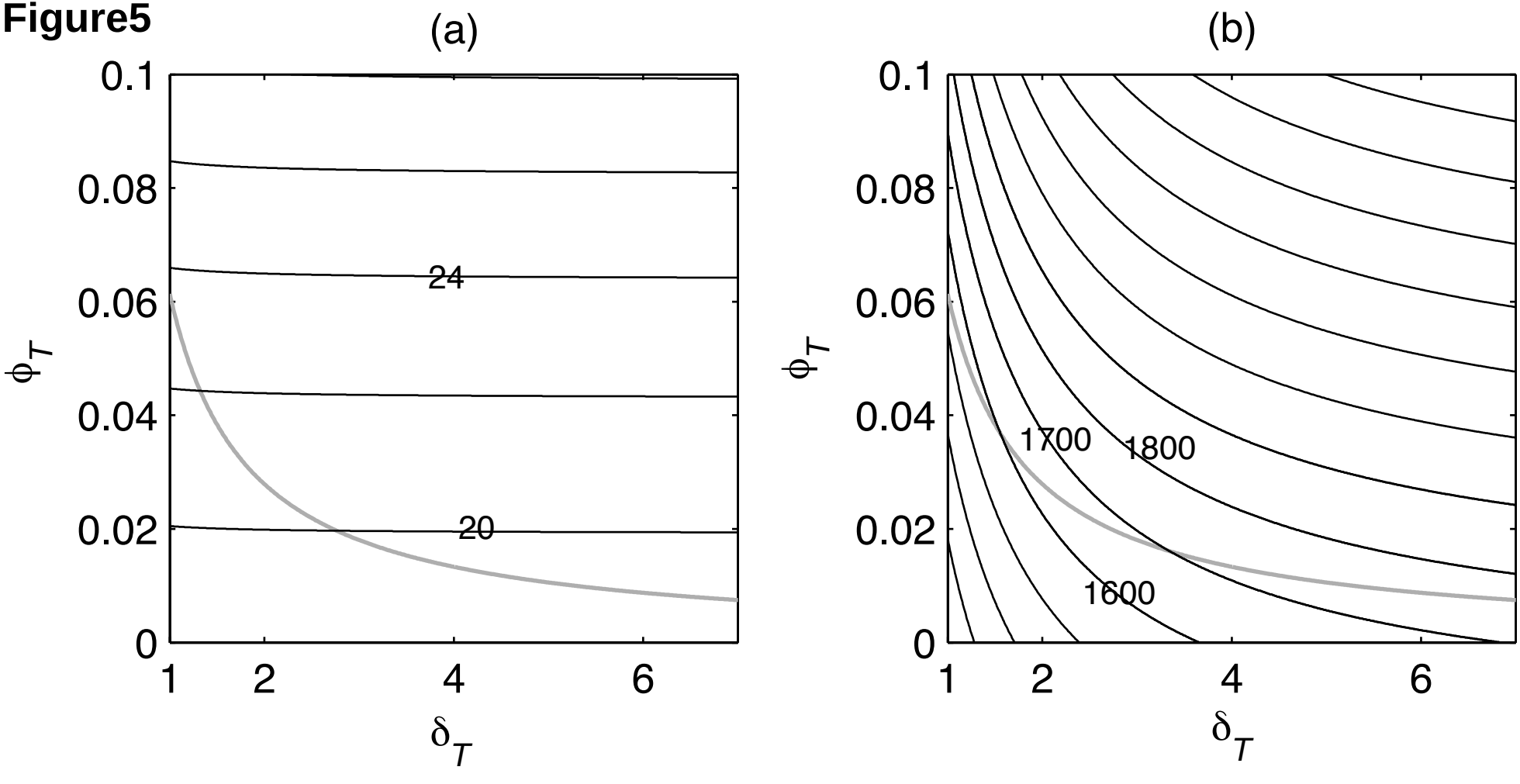
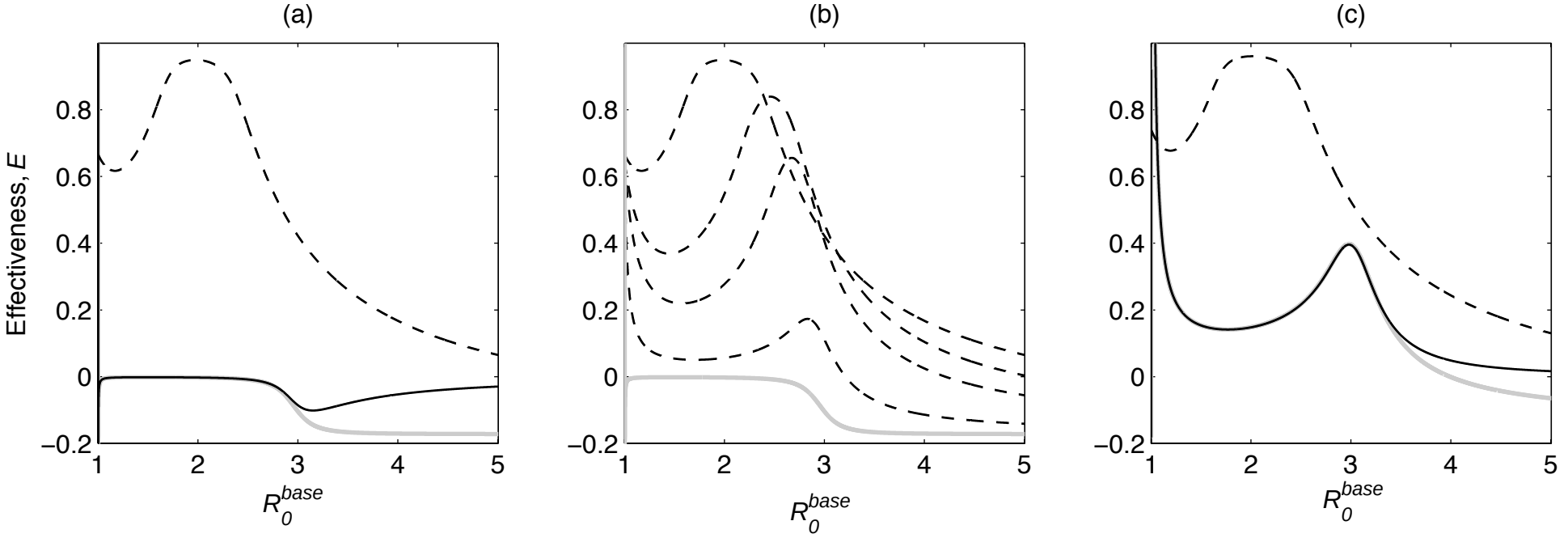


Figure D6



Supplementary Material for on-line publication only

Click here to download Supplementary Material for on-line publication only: SM-Pinho-TPB-D-13-00101.pdf