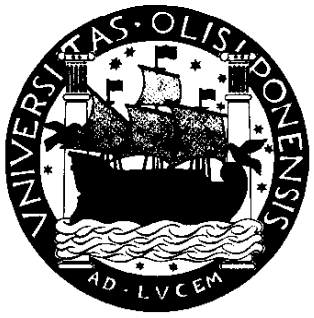
UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE MATEMÁTICA



Modeling tuberculosis:

a compromise between biological realism and mathematical tractability

Paula Cristiana Costa Garcia da Silva Patrício Rodrigues

Tese orientada pela Doutora Maria Gabriela Gomes
e Prof.ª Doutora Maria Carlota Rocha Xavier Rebelo Gonçalves

DOUTORAMENTO EM MATEMÁTICA
(Análise Matemática)

2009

Agradecimentos

Ao chegar ao fim deste longo caminho não poderei deixar de endereçar vários agradecimentos.

Em primeiro lugar, às minhas orientadoras, Gabriela e Carlota, pelo apoio, disponibilidade e solidariedade que sempre demonstraram, mesmo nos momentos mais difíceis. O seu envolvimento e interesse foram inspiradores.

Às Instituições que possibilitaram todas as condições para a realização deste trabalho quero demonstrar o meu maior apreço e satisfação, nomeadamente ao Departamento de Matemática da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, na pessoa da Presidente do Departamento, Prof.ª Doutora Maria Helena Santos; ao Instituto Gulbenkian de Ciência, Fundação Calouste Gulbenkian, na pessoa do Director Prof. Doutor António Coutinho; ao Centro de Matemática e Aplicações Fundamentais da Universidade de Lisboa e à Fundação para a Ciência e Tecnologia. Sem eles este trabalho não teria sido possível. O seu apoio, suporte e solidariedade foram essencias para o concretizar desta tese. Não posso deixar passar este momento sem agradecer a todos aqueles que contribuíram para a minha formação pessoal e académica, ao longo dos anos de estudo na Faculdade de Ciências da Universidade de Lisboa, instituição a que estou reconhecida.

Aos Prof. Herbet Hethcote e Prof. Cláudio Struchiner, por partilharem o seu vasto conhecimento, num espírito de cooperação que contribuiu determinantemente para o enriquecimento das ideias e do trabalho, consolidando as dúvidas e questionando as certezas.

Aos colaboradores, com quem tive o prazer de trabalhar mais de perto: Ana Cristina Paulo, Alessandro Margheri, Frank Hilker, Graham Medley, Marion Muehlen, Marta Nunes e Ricardo Águas, pelo tempo comum e pela partilha das tristezas e alegrias, dos sucessos e insucessos, das angústias e expectativas.

Aos meus colegas e amigos do IGC pelas longas discussões científicas e pelas horas bem passadas com muitos bolos, muito em especial à Cristina, Natália e Ricardo, pela amizade que foi crescendo. Fizeram deste espaço-tempo mais do que um momento para o tornarem especial e único.

À minha Família e Amigos, por estarem sempre 'lá',

e muito em epecial aos meus *três 'ursinhos' TODOS preferidos*, Francisco, Catarina e Pedro, pelo seu Amor *até à Lua... e de volta até cá abaixo*!

e ao Francisco por TUDO!

Obrigada.
Paula Rodrigues

Sumário

Esta tese tem como motivação um problema de interesse actual – transmissão da tuberculose, do qual focamos sobretudo dois aspectos: o papel da reinfecção na tuberculose recorrente (recidivas) e na propagação de estirpes resistentes a antibióticos. O objectivo é desenvolver explicações teóricas para os diversos fenómenos, capturando os mecanismos subjacentes através de modelos matemáticos simples de modo a permitirem um profundo estudo analítico. Os mecanismos explorados são a imunidade parcial e a heterogeneidade. A conjugação destes dois mecanismos mostrou ser um importante factor na determinação de diferentes aspectos da epidemilogia da Tuberculose.

A propagação de estirpes resistentes é abordado como a competição entre estirpes sensíveis e resistentes a antibióticos. O objectivo é caracterizar o impacto da reinfecção na região de coexistência. Observamos que a reinfecção impõe um novo limiar na transmissão, acima do qual a dissiminação de estirpes resistentes é facilitada. Consequentemente, as medidas de controlo beneficiaríam de uma alteração do seu focus, passando da tentativa de redução de casos de resistência adquirida para a interrupção das cadeias de transmissão, dependendo do enquadramento epidemiológico.

Motivado por estudos epidemiológicos que indicam que as taxas de tuberculose por reinfecção são mais altas que as taxas de tuberculose primária, propomos uma explicação assente na heterogeneidade da população. Esta hipótese permite reconciliar estas observações com a ideia consensual de que a infecção confere protecção parcial, reduzindo o risco individual de infecções subsequentes. Assim, postulamos que alguns indivíduos têm *a priori* um risco acrescido de infecção, sendo por isso mais afectados. Isto contribui para uma acumulação de individuos de elevado risco entre o grupo de pacientes, o que inflaciona as taxas de reinfecção. Esta hipótese é formulada matematicamente e confrontada com dados epidemiológicos referentes a várias regiões endémicas. Propomos um critério alternativo de validação desta hipótese.

O estudo das consequências da heterogeneidade na susceptibilidade à infecção é alargado a modelos epidemiológicos mais gerais que assumem imunidade parcial - os modelos Susceptíveis-Infecciosos-Recuperados-Infecciosos (SIRI).

Palavras Chave: Modelos matemáticos; análise de bifurcação e estabilidade; imunidade parcial; heterogeneidade; tuberculose

Abstract

This thesis is a compromise between biological realism and mathematical tractability. Our aim is to provide theoretical explanations for observed phenomena by capturing the underlying mechanisms involved, with simple mathematical models that allow a deep analytic investigation. The biological motivation is an epidemiological problem of major current interest –tuberculosis transmission, for which we focus on two questions: the role of reinfection in recurrent disease and the spread of drug resistant strains. The underlying mechanisms explored are partial immunity and heterogeneity. In conjunction these two prove to be important factors in the determination of tuberculosis epidemiological landscape.

We address the problem of drug resistance as a competition between drug-sensitive and resistant strains. Our objective is to characterize how reinfection modifies the conditions for strain coexistence. Reinfection imposes a new threshold for transmission, above which resistant strains dissemination is facilitated. Consequently, drug resistance control would benefit from a change in the interventions focus, from drug acquisition reduction to transmission blocking, depending on the epidemiological setting.

Motivated by molecular epidemiology studies indicating that rates of reinfection tuberculosis are higher than rates of new tuberculosis, we propose the selection hypothesis to reconcile these observations with the consensual view that tuberculosis infection confers partial protection that reduces the individual susceptibility to reinfection. We postulate that some individuals are *a priori* more likely to develop the disease. As infection tends to affect the more susceptible individuals, the distribution of recovered individuals is skewed towards higher susceptibility, inflating the rates of reinfection. The hypothesis is formulated mathematically and confronted with data for different endemic regions. We propose a new criterion for further validation.

The consequences of host heterogeneity in susceptibility to infection is also explored in the context of more general epidemiological models assuming partial immunity (SIRI models).

Keywords: Mathematical models; stability and bifurcation analysis; partial immunity; heterogeneity; tuberculosis

Resumo

A motivação para esta tese é um problema epidemiológico de interesse actual – a transmissão da Tuberculose (TB), do qual focamos sobretudo dois aspectos: o papel da reinfecção na TB recorrente (recidivas) e na propagacção de estirpes resistentes a antibióticos. Ao longo da tese, dois aspectos transversais são analisados: imunidade parcial, para a qual a proteccção conferida pela primeira infeccção não é total mas reduz o risco de reinfecção; e heterogeneidade, quer seja na susceptibilidade dos indivíduos à infeccção ou na susceptibilidade do agente infeccioso a antibióticos.

O nosso objectivo é propôr explicacções teóricas para os fenómenos observados em TB, capturando os mecanismos que os explicam através de modelos matemáticos simples, passíveis de um profundo estudo analítico. Na nossa abordagem, procuramos estudar o comportamento desses fenómenos para diferentes regiões em vez de nos focarmos numa só região ao longo do tempo. Priviligiamos os modelos matemáticos simples em determimento dos mais complexos, fortemente dependentes de dados epidemiológicos. Usamos metodologias das áreas da teoria de sistemas dinâmicos e teoria da bifurcação para desenvolver e analizar os modelos. Em paralelo ao trabalho desenvolvido especificamente em TB, propomos metodologias mais gerais que possam ser aplicadas a outras doenças infecciosas. O trabalho tem como base modelos anteriores que caracterizam sistematicamente a reinfeccção em sistemas epidemiológicos em geral e especificamente em TB (Gomes *et al.*, 2004a,b, 2005).

O Capítulo 2 é uma introdução aos conceitos de imunidade parcial e de limiar de reinefecção. O conteúdo dos Capítulos 3 e 5 faz parte de manuscritos já publicados (Rodrigues *et al.*, 2007, 2009). O Capítulo 4 resulta de trabalho ainda em curso que esperamos melhorar e completar no futuro próximo. Finalmente, o Capítulo 6 constitui um manuscrito completo recentemente submetido para publicação.

Há muito que os modelos matemáticos são usados no estudo de doenças infecciosas. Os modelos clássicos SIR (Susceptíveis-Infecciosos-Recuperados) e SIS (Susceptíveis-Infecciosos-Susceptíveis) têm sido extensamente estudados e aplicados com sucesso (Kermack & McKendrick, 1927; Anderson & May, 1991). São modelos em que se considera que a população está dividida em compartimentos ou classes de acordo com o seu historial de infecção: susceptível, infeccioso ou recuperado com imunidade. A dinâmica destes modelos é descrita por sistemas de equações diferenciais. Nos modelos SIR e SIS assume-se que a infecção confere imunidade que protege totalmente ou não confere qualquer protecção. A primeira forma de imunidade é típica das chamadas doenças infantis, como o sarampo, a papeira ou a rubéola, enquanto que a segunda forma refere-se a infecções que permitem múltiplas reinfecções ao longo da vida como a malária. Contudo, a maior parte das infecções situar-se-á entre estes dois extremos. Uma das extensões propostas é o modelo com imunidade parcial - modelo

SIRI (Susceptíveis-Infecciosos-Recuperados-Infecciosos) (Gomes *et al.*, 2004b), no qual se assume que os indivíduos estão protegidos enquanto infectados mas voltam a adquirir alguma susceptibilidade após recuperarem. Nestes modelos observamos dois tipos de comportamento endémico: baixo e potencialmente oscilatório, como nos modelos SIR ou alto e não-oscilatório, como nos modelos SIS. Estes dois comportamentos estão separados por um limiar na transmissão que depende da dinâmica de reinfecção - o limiar de reinfecção (LR). Os modelos SIRI e o conceito do limiar de reinfecção, são conceitos centrais para o trabalho desenvolvido nesta tese.

O Capítulo 2 surge da necessidade de estender o conceito do limiar da reinfecção a um conjunto mais alargado de modelos e à definição de um método generalizado para o seu cálculo. Primeiro, estendemos o conceito LR, incluído no modelo SIRI, a outros fenómenos como a imunidade temporária, a latência e a reactivação. As características inicialmente descritas para o SIRI podem variar com a importância da reinfecção relativamente aos outros fenómenos, mas o LR estará sempre associado a mudanças na dinâmica da doença e reduzido impacto da vacinaçã. Na Secção 2.5 propomos um método para o cálculo do LR, para qualquer modelo de transmissão descrito por um sistema de equações diferenciais. O método é baseado na definição de um sub-modelo de reinfecção, para o qual o processo de reinfecção é isolado dos restantes fenómenos. Uma bifurcação no parâmetro de transmissão, para o equilíbrio trivial, quando existe, corresponde ao LR. Assim, o LR pode ser mais geralmente interpretado como o nível de transmissão a partir do qual a infecção pode ser mantida endémica numa população parcialmente imunizada. O método desenvolvido é usado no Capítulo 5 para um modelo SIRI heterogéneo, para o qual mais do que um LR é identificado. De facto, quando existe mais do que um grupo da população com níveis distintos de susceptibilidade, a dinâmica do modelo é determinada por mais do que um LR. A sua identificação pode não ser evidente até que o sistema seja perturbado e os perfis de susceptibilidade modificados. Um outro exemplo deste fenómeno pode ser encontrado em Gomes *et al.* (2007).

A TB é uma doença rica em paradoxos. Foi uma das primeiras doenças para a qual o agente causador foi identificado (Robert Koch, 1882), e uma das primeiras para as quais uma vacina foi desenvolvida (por Albert Calmette e Camille Guérin, 1906). Antibióticos eficazes estão disponíveis há mais de meio século (Iseman, 2002). No entanto mantem-se uma das doenças infecciosas mais comuns e mortais. Estima-se que um terço da população mundial esteja actualmente infectada. TB é uma doença infecciosa causada pelo *Mycobacterium tuberculosis* que afecta sobretudo os pulmões. Apenas 10% dos indivíduos infectados progridem directamente para doença e tornam-se infecciosos nos dois anos após o contágio. Nos restantes, a infecção é controlada, mas mantida num estado de latência, podendo ser reactivada após um período variável de tempo. Pode ainda haver reinfecção, aquando do contacto com uma nova estirpe. Estas duas características, reactivação e reinfecção, colocam várias dificuldades na modelação desta doença e são também aqui estudadas em diferentes momentos.

Em conjunto com a co-infecção por HIV a resistência aos antibióticos constitui uma das preocupações principais no controlo da TB. Coloca importantes restrições à eficácia dos tratamentos disponíveis e o problema continua a agravar-se com o aparecimento de estirpes multiresistentes. No Capítulo 3, abordamos o problema da propagação das estirpes resistentes. É proposto um modelo com dois tipos de estirpes: susceptíveis e resistentes a antibióticos. O nosso objectivo é caracterizar como a reinfecção modifica as condições de coexistência. Conclui-se que a reinfecção impõe um novo limiar na transmissão, acima do qual a disseminação de estirpes resistentes é facilitada. Consequentemente, as medidas de controlo

beneficiariam de uma alteração do seu focus, passando da tentativa de redução de casos de resistência adquirida para a iterrupção das cadeias de transmissão, dependendo do equadramento epidemiológico.

Estudos de epidemiologia molecular sugerem que a coexistência de estirpes com diferentes perfis de susceptibilidade é possível não apenas ao nível da população mas também ao nível do indivíduo. Na Secção 3.5, adaptamos o modelo proposto anteriormente de modo a incluir a possibilidade de infecções mistas. Todavia, a necessidade de melhor compreender como a competição de estirpes sensíveis e resistentes dentro do hospedeiro influência a capaciade da estirpe ser transmistida na população, motiva a integração de dinâmicas em duas escalas - população e indivíduo. No Capítulo 4, apresentamos trabalho ainda em curso sobre modelos para a tuberculose usando estas duas escalas de modelação. São propostos modelos simples para dinâmica da infecção dentro do hospedeiro que descrevem o progresso das infecções mistas. A duração das infecções mistas e a frequência de estirpes resistentes durante estas infecções são relacionadas com o fitness relativo das estirpes. As relações obtidas são então usadas no modelo epidemiológico. Deste modo, obtemos uma caracterização do cenário epidemiológico com base na fitness da estirpe equanto competição dentro do indivíduo.

Como a heterogeneidade, nas suas diferentes formas, altera a dinâmica das doenças infecciosas, tem sido uma questão importante na sua modelação. A principal motivação para a introdução de heterogeneidade nos modelos é necessidade de melhor fazer corresponder os resultados teóricos aos dados epidemiológicos. No Capítulo 5, propomos uma metodologia para incluir a heterogeneidade do hospedeiro nos modelos SIRI, capaz de captar os seus efeitos, sem que seja perdida capacidade de análise matemática dos modelos. Analizamos o impacto da heterogeneidade na prevalência da doença e comparamos os perfis de susceptibilidade das sub-populações em risco de primeira infecção e de reinfecção. De acordo com o descrito na literatura para os modelos SIRI e SIS com heterogeneidade, também estes modelos tendem a gerar prevalências mais baixas do que os correspondentes modelos homogéneos. Para além disso, observamos que a heterogeneidade na susceptibilidade à infecção gera um mecanismo de selecção nos grupos de risco elevado, através da transmissão, o que pode explicar taxas de reinfecção inesperadamente altas. Este mecanismo de selecção é especialmente notório em regiões de baixa ou moderada transmissão onde, sob a hipótese de imunidade parcial, se poderia esperar que a reinfecção fosse rara. As vantagens das estratégias de controlo focadas nos grupos de elevado risco são exploradas através do estudo da vacinação não uniforme e de intervenções capazes de alterar o perfil de susceptibilidade da população. Estas últimas, ao contrário das primeiras, têm o potencial de poder reduzir ou mesmo eliminar a doença para populações acima do LR.

O tópico no Capítulo 5, é na verdade motivado por um problema em TB. Estudos de epidemiologia molecular relatam que, para certas regiões, a taxa de tuberculose por reinfecção, entre doentes tratados, é maior que a taxa de tuberculose por novas infecções (Verver *et al.*, 2005). Estas observações podem ser explicadas admitindo que o risco individual de infecção é maior depois de tratamento (Uys *at al.*, 2009). Baseando-nos nas conclusões do modelo SIRI com heterogeneidade, formulamos uma explicação alternativa, assente na heterogeneidade da população. Assim, postulamos que alguns indivíduos têm *a priori* um risco mais elevado de infecção. Estes indivíduos vão ser seleccionados pelo processo de transmissão contribuindo para uma acumulação de indivíduos de elevado risco entre o grupo de doentes tratados. Consequentemente as taxas de reinfecção observadas neste grupo aparecem inflac-

cionadas. Esta hipótese permite reconciliar a ideia consensual de imunidade parcial conferida pela infecção de tuberculose, com o aumento da taxa de reinfecção observado. No Capítulo 6, as duas hipóteses alternativas para este fenómeno são integradas em dois modelos assumindo a mesma estrutura para a transmissão da TB. Dados publicados sobre a proporção de reinfecção na tuberculose recorrente (recidivas) para diversas regiões endémicas são usados para parametrizar os modelos, os quais são analisados e comparados. Concluímos que só a hipótese de selecção é compatível com os critérios epidemiológicos descritos para a TB. Um critério alternativo para distinguir os modelos, baseado em dados epidemiológicos de taxas de reinfecção e tuberculose primária, é proposto e discutido.

Contents

Chapter 1

Introduction

1.1 Motivation and history

This thesis is a compromise between biological realism and mathematical tractability. The motivation for this work is an epidemiological problem of major current interest – Tuberculosis (TB) transmission. We address mainly two questions: the role of reinfection in recurrent infections and the spread of drug resistance. Throughout the work two transversal aspects are analyzed: partial immunity, protection against subsequent infections conferred by first infection is not fully protective; and heterogeneity in host susceptibility to infection or in strain sensitivity to drugs. Our aim is to provide theoretical explanations for observed phenomena in TB. We intend to capture the underlying mechanisms involved in the different manifestations across regions, instead of focusing in a certain region/country over time. We develop simple mathematical models for TB transmission that allow a deep analytic investigation, in opposition to data driven, highly computational complex models. We make use of methodologies from dynamical systems theory and bifurcation theory to develop and analyze these models. In parallel, we propose general theoretical frameworks that can easily be applied to other infectious diseases (chapters 2 and 5). The work builds on original models that systematically characterize reinfection in epidemiological systems in general, and specifically in tuberculosis epidemiology (Gomes *et al.*, 2004a,b, 2005).

Mathematical models have long been used in the study of the transmission of infectious diseases. One of the first examples is the work developed by Daniel Bernoulli on smallpox (1760), where he used a mathematical model to evaluate the effectiveness of variolation against smallpox, with the aim of informing public health policy. Also determinant was the work of Ronald Ross on the transmission of malaria (1908). He is responsible for a continuous-time framework using mass action principle, which assumes that the rate of infection is proportional to the product of the density of susceptible individuals and the density of infectious individuals. Few years latter Kermak and McKendrick established the theory of the epidemic threshold according to which the turning point of an epidemic occurs when the density of susceptible individuals crosses a certain critical value (Kermack & McKendrick, 1927). This theory in conjunction with the principle of mass action is the cornerstone of the modern theoretical epidemiology.

The SIR (Susceptible-Infectious-Recovered) and SIS (Susceptible-Infectious-Susceptible) frameworks for infectious diseases have been extensively studied and successfully applied

(Kermack & McKendrick, 1927; Anderson & May, 1991). These are compartmental models, where the host population is considered to be divided into different classes depending on disease state: susceptible, infectious and recovered with immunity. The dynamics of the various compartmental models are deterministically determined by systems of ordinary differential equations. These models correspond to two extreme situations and have been the bases for many extensions, motivated by the need to describe the specific disease dynamics in more detail. In this work we are particularly interested in extending these compartmental models in two directions by allowing for variation in susceptibility to infection due to innate factors or acquired immunity.

In the SIR and SIS models infection induces an immune response that protects totally or nothing, respectively. The former is classically applied to childhood diseases like measles, mumps or rubella, and the latter refers to infections that have repeated reinfections throughout life, such as malaria. But most infections are somewhere between these two cases. A natural extension of this framework was developed by Gomes *et al.* (2004b). We can consider several mechanism to interpolate between the two extremes: temporary immunity, where the infection confers total immunity but for just a certain period of time; partial immunity, where immunity is not fully protective but reduces the risk of subsequent infections; and combinations of these two in which infection confers partial immunity for a certain period of time. When these extensions are considered new dynamics appear. In the case of partial immunity models, we can have two types of endemic behaviour, low and potential oscillatory or high and steady. These two types of endemic behaviour are separated by a second threshold of transmissibility that depends on the reinfection dynamics - reinfection threshold (RT). The importance of the RT in marking the emergence of new dynamical behavior will to be discussed throughout this work.

How host heterogeneity, in its different forms, changes the dynamics of infectious diseases has been an important question in infectious diseases modeling. One the one hand, models ought to be simple and tractable but, on the other hand, key traits cannot be neglected under the risk of not capturing important features of the disease behavior. Discussion on this subject is scattered through the literature sometimes in different areas and apparently unrelated problems. Mostly, the incorporation of host heterogeneities into the models is driven by disease related questions and data interpretation. Structuring work on the subject was proposed by Anderson & May (1991) and Diekmann *et al.* (1990) and also by a series of related papers (Ball, 1985; Anderson & Britton, 1998; Britton, 1998).

Heterogeneity in the host population can be based on biological factors such as genetic susceptibility or resistance, duration of the infectious period or chemotherapy and also on social, economic or demographic factors that affect frequency or intensity of contacts. Different frameworks have been proposed to include host heterogeneity into SIR or SIS models, from compartmental models (Anderson & May, 1991; Hethcote, 1996), to distributed parameter systems (Coutinho *et al.*, 1999; Diekmann & Heesterbeek, 2000; Diekmann *et al.*, 1990; Dushoff, 1999; Veliov, 2005) and more recently, using network models (Miller, 2007).

It generally concluded that the inclusion of host heterogeneity in the models reduces the epidemic sizes and alters the effects on the control effort. However, it is important to note that it depends on the phenomena analyzed and that comparison criteria between homogeneous and heterogeneous models can alter these conclusions (Anderson & Britton, 1998).

Tuberculosis is a disease rich in paradoxes. It was one of the first diseases for which the causative agent was identified (by Robert Koch, 1882), and one of the first for which a vaccine

was developed (by Albert Calmette and Camille Guérin, 1906). Effective antibiotics have been widely available for half a century (Iseman, 2002). Yet it remains one of the most common infectious disease and a great killer. One third of the world's population (approximately two billion individuals) is believed to be currently infected. The World Health Organization (WHO) estimates that 9.27 million new cases of TB occurred in 2007, killing 1.32 million among individuals with no human immunodeficiency virus (HIV) infection and 456,000 HIV-positive people (WHO, 2009). The South-East Asia and Western Pacific regions account for 55% of global cases and the African Region for 31% (WHO, 2009). The magnitude of the TB burden within countries can also be expressed as the number of incident cases per 100 000 population. The world incidence of new TB cases, during 2007, was 139 per 100 000 population (WHO, 2009). Some regions have extremely high incidence rates reaching 1,000 per 100,000 population, while the usual incidence is considered to vary between 5 (low-incidence countries) and 200 (high- incidence countries). Among the 15 countries with the highest estimated TB incidence rates, 13 are in Africa, a phenomenon linked to high rates of HIV coinfection (WHO, 2009). Also resistance to the anti-tuberculosis antibiotics poses serious difficulties to TB control. In 2007, among all cases of TB, around 5% were multi-drug resistant (WHO, 2009). The only vaccine in current use, the bacille Calmette-Guérin (BCG), is cost-effective, but its efficacy is highly variable, ranging from 0% to 80%.

Tuberculosis is an infectious disease caused by *Mycobacterium tuberculosis* (*Mtb*). Although tuberculosis is primarily a pulmonary disease, the bacterium can infect and cause disease in almost all organs and tissues, including the central nervous system and bone. TB is transmitted by airborne particles spread through the air when people who have the disease cough, sneeze, or spit. Infection usually results when the *mycobacterium* is deposited in the lungs of exposed persons. From all infected individuals, less than 10% progress directly to active disease and become infectious within two years upon infection – primary disease. For the remainder, infection is successfully controlled but not eliminated – latent infection. Latently infected individuals are asymptomatic and do not contribute to transmission. Complex cellular structures, called granulomas, ensure mycobacterial containment in a dormant stage that can last for variable periods of time. Latent infection can reactivate to cause active tuberculosis by endogenous reactivation, mainly due to immunossupress ion, or by exogenous reinfection, which is typically caused by a different genotype. Different models have been used in the study the long-term dynamics of tuberculosis. The selection of a model is intimately connected to the particular question one wants to address. We can say that in the tuberculosis modeling literature there is a core group of model structures that have been adapted to the different questions, such as interpretation of historical data, effectiveness assessment of different control strategies, spread of drug-resistance or HIV coinfection. The main contributions can be tracked to the initial models from different research groups such as Blower *et al.* (1995); Castillo-Chavez & Feng (1997); Vynnycky and Fine (1997); Dye *et al.* (1998); Murray & Salomon (1998); Murphy *et al.* (2003); Gomes *et al.* (2004a) or Cohen & Murray (2004).

Tuberculosis is a complex disease and this work is not meant to be exhaustive. We focus on two aspects: the role of reinfection in tuberculosis transmission and the spread of drug-resistance. Throughout the present work we introduce the required biological information and modeling formalism.

1.2 Summary

Chapter 2 is an introduction on the concepts of partial immunity and the reinfection threshold. The contents of Chapters 3 and 5 are published manuscripts (see Section 1.3). Chapter 4 is ongoing work that we hope to complete in the near future. And Chapter 6 is a complete manuscript that has been recently submitted for publication in a shorter version.

Chapter 2 is based on the reinfection threshold concept initially introduced in the Susceptible-Infectious-Recovered-Infectious (SIRI) context where infection induces a partially protective immune response (Gomes *et al.*, 2004b). The inclusion of partial immunity in tuberculosis models, in particular for the work on post-exposure interventions in Gomes *et al.* (2007), motivated the generalization of the RT to a wider range of models and to the definition of a generalized method for its computation. We first extend the concept of the RT, by including into the SIRI model other processes involved in disease transmission: temporary immunity, latency and reactivation/relapse. The characteristics described for the simple SIRI can vary with the importance of reinfection in comparison to the other disease processes, but it is always associated with changes in the disease dynamics. In Section 2.5, we propose a generalized method for the computation of the RT for any general compartmental disease transmission model described by a system of ordinary differential equations. The method is based on the definition of a sub-model, the reinfection sub-model, for which the reinfection process has been isolated from the remaining disease processes. A bifurcation on the transmission parameter of the disease-free equilibrium, if it exists, will correspond to the RT. Hence, the RT can be more generally interpreted as the transmission level above which (re)infection can sustains transmission in a partially immunized population. The method developed is used throughout this work. For systems where more than one partially immune class exists, multiple reinfection thresholds can be defined. These thresholds show an increased importance in the context of interventions that affect the population risk profile, where stability of multiple equilibria can occur in regions limited by these thresholds. In Section 5.5.2 such an intervention is analyzed.

Together with coinfection with HIV, drug resistance constitutes the main concern in tuberculosis control. It poses important restrictions in treatment management and the problem continues to aggravate with the emergence of strains that are resistant, each time, to a higher number of anti-tuberculosis drugs. In Chapter 3, we address the problem of drug-resistance spread. We construct a model with two types of strains, drug-susceptible and resistant strains. Our goal is to characterize how reinfection modifies the conditions for coexistence by giving another opportunity of resistant strains to spread, independently of drug acquisition. Different control strategies are discussed and it is shown that intervention effectiveness is highly sensitive to the baseline epidemiological setting.

Molecular epidemiological studies suggest that coexistence of strains with different susceptibility profiles are possible not only at the population level but also within an individual. In Section 3.5, we extend the drug resistance TB model to include the possibility of mixed infections. However, the need to better understand how competition of sensitive and resistant strains within-host influences the ability of strains to be transmitted in the population, motivates the integration of dynamics at two scales. In Chapter 4, we present ongoing work on multi-scale models for tuberculosis. Simple within-host models describing the progression of mixed infections are proposed. The duration of mixed infections and the frequency of resistant strains during these infections are related to the relative fitness of strains. The obtained

relations are then used in the epidemiological model. This way, one can obtain a theoretical epidemiological landscape on coexistence of strains, based on within-host relative fitness. The models developed, so far, still require further study and validation.

Heterogeneity in susceptibility and infectivity is inherent to infectious disease transmission in nature. In Chapter 5 we use a simple framework to include host heterogeneity into mathematical models that is able to capture the essence of heterogeneity effects maintaining a simple structure suitable of mathematical analysis. We explore the consequences of host heterogeneity in the susceptibility to infection, for epidemiological models for which immunity conferred by infection is partially protective (SIRI models). We analyze the impact of heterogeneity on disease prevalence and contrast the susceptibility profiles of the sub-populations at risk for primary infection and reinfection. Heterogeneity in susceptibility to infection generates a selection mechanism on the high-risk groups driven by transmission, that can explain unexpectedly high reinfection rates. The enhanced benefits of control strategies that target the more high-risk groups are explored.

The topic explored in Chapter 5 is, in fact, motivated by a concrete problem in TB epidemiology. Molecular epidemiological studies report a rate of reinfection TB higher than the rate of new TB among treated patients, for transmission community in Cape Town (Verver *et al.*, 2005). Based on the conclusions of the simple SIRI heterogeneous model, we formulate an hypothesis that explains the observations noting that infection imposes a selection mechanism whereby the risk profile of the recovered compartment is skewed towards high risk. This hypothesis can reconcile the widely accepted idea of partial immunity conferred by TB infection with the increased reinfection rate observed, without assuming an increased risk of subsequent infection after treatment, as proposed by others (Uys *at al.*, 2009). In Chapter 6 the alternative hypotheses for this phenomenon are then integrated into two models assuming the same structure for tuberculosis transmission. Published data on the reinfection proportion in recurrent TB, for different endemic regions, is used to parameterize the models, which are analyzed and compared. An alternative measure that can further distinguish both models based on epidemiological data on reinfection and primary TB rates, is discussed.

Possible extensions and applications of the work developed in this thesis are discussed in Chapter 7.

1.3 Publications

The contents of this thesis relate to the following publications:

- Chapter 2 is a summary of the effort along different problems and models to generalize the definition and computation of the reinfection threshold. One of the determinant problems for this generalization is published in

  Gomes M.G.M., RODRIGUES P, Hilker F., Mantilla-Beniers NB, Muehlen M, Paulo A, Medley G. 2007. Implications of partial immunity on the prospects for tuberculosis control by post-exposure interventions, *J. Theor. Biol.* **248**,(4):608-17.

- Chapter 3 is adapted with minor changes from

  RODRIGUES P, Gomes MG, Rebelo C. 2007. Drug resistance in tuberculosis - a reinfection model, *Theor. Pop. Biol.* **71**: 196-212.

- Chapter 5 is adapted, with exception of the inclusion of section 5.5.2, from the accepted manuscript

  RODRIGUES P, Margheri A, Rebelo C, Gomes MGM, 2009. Heterogeneity in susceptibility to infection can explain high reinfection rates, *J. Theor. Biol.* **259**: 280-290.

- Chapter 6 is an unpublished manuscript. Recently, an adapted version of this manuscript has been submitted for publication.

  RODRIGUES, P, Águas R, Nunes M, Rebelo C, Gomes MGM. 2009. High rates of reinfection tuberculosis: the selection hypothesis.

1.4 Software

Throughout the work, numeric calculations and some analytical manipulations are obtained using MATLAB 6.5®. In Chapter 3, equilibrium curves are computed with MATCONT continuation package of MATLAB 6.5® (Dhooge *et al.*, 2003). In Chapter 6, the fitting procedure is performed using Berkeley Madonna software v8.3.6c.

Chapter 2

Partial immunity and thresholds in transmission - the Reinfection Threshold

2.1 Introduction

The basic reproduction number R_0 is one of the most important quantities in the study of epidemics and in the comparison of population dynamical effects of control strategies. It is present in almost all papers that use mathematical modeling for the study of the spread of infections in populations. R_0 is, biologically, defined as the expected number of new infections caused by a typical infected individual in a totally susceptible population, during his/her infectious period. In demography and ecology, R_0 has an equivalent interpretation as the expected number of female offspring born to one female during her entire life. This concept was fully formed much earlier in the demography context and despite several opportunities to cross over between demography, ecology and epidemiology (Heesterbeek, 2002), it took until the 90s to become fully developed and its applicability realized in theoretical epidemiology. For that, it was determinant the extensive use of the concept and estimates from serological data by Anderson & May (1991) and the development of a mathematical theory of R_0 for heterogeneous populations by Diekmann *et al.* (1990) and Dietz (1993).

The value $R_0 = 1$ defines a threshold in disease transmission – the epidemic threshold. Below this threshold, an infectious agent will not invade a totaly susceptible population. Above the threshold, the pathogen can invade and it is expected that for $R_0 > 1$ the disease becomes endemic. R_0 has been used to characterize different populations according to transmission potential. It is also used as a measure of the control/elimination effort – if transmission can be reduced by the control measures implemented and subsequently R_0 decreased below one, then the pathogen can be eliminated. Mathematically, for models expressed by deterministic dynamical systems, $R_0 = 1$ is associated with a bifurcation in the transmission parameter that marks the transition of the disease free steady state from stable to unstable. Most of the times, it is also associated with the emergence and/or stability of an endemic steady state. R_0 can be formally defined as the dominant eigenvalue of the so called *next generation operator* (Diekmann *et al.*, 1990).

The epidemic threshold is not the only threshold to affect transmission and disease control. It has been argued that for diseases for which immunity is not fully protective, allowing

reinfection to occur after recovery at a reduced rate, another important threshold is associated with disease level and control effectiveness – the reinfection threshold (Gomes *et al.*, 2004b, 2005).

The RT was initially defined in the context of the SIRI model. This model assumes that individuals are protected while infected but regain some susceptibility upon recovery. Compared to susceptibility prior to infection, this susceptibility is reduced by a factor σ. The SIR and SIS models can be regarded as two extremes of the SIRI model by setting σ to zero or one, respectively. Associated with the reinfection threshold it was described a steep increase in disease prevalence, a drop in vaccination impact and a shift from low and steady equilibrium to high and possible oscillatory as the threshold is crossed, corresponding to a shift from a SIR-type of behavior, below threshold, to SIS-type of behavior above. The SIRI model has been further investigated in a spatial stochastic version (Stollenwerk *et al.*, 2007) and the RT implications have been studied for different diseases for which immunity is not fully protective (Águas *et al.*, 2006; Gökaydin *et al.*, 2007; Gomes *et al.*, 2004a, 2007).

This chapter is motivated by the need to extend the concept of the RT to more general models and to define a generalized method for its computation. Initially, we extend the concept of the RT, by including into the SIRI model other processes involved in disease transmission. These processes are analyzed independently. As in Gomes *et al.* (2004b), we include temporary immunity as another example of imperfect immunity. Motivated by the tuberculosis models that are central in this work, we also consider latency and endogenous reactivation/relapse. We focus on two characteristics of the change in behavior induced by the RT in the simple SIRI model: (1) steep increase in the disease prevalence, (2) vaccine impact variation. Both are considered in relation with the basic reproduction number. We proceed by describing which of these characteristics are maintained and which are changed, for the different sub-models. Finally, we formulate a generalized method for the computation of the RT for any general compartmental disease transmission model described by a system of ordinary differential equations.

2.2 Generalized model

We construct a generalized model where, in addition to SIRI transmission other processes are included: latency, temporary immunity and endogenous reactivation/relapse. We analyze each process separately. The total population is divided into disease related classes S, L, I, R and V, that stand for the proportion of susceptible, latently infected, infectious, recovered and vaccinated individuals. We use the following set of differential equations to describe the generalized model:

$$\begin{cases} S' &= (1-v)\mu + \alpha(R+V) - \lambda S - \mu S \\ L' &= \lambda S + \sigma\lambda(R+V) - (\nu+\mu)L \\ I' &= \nu L + \omega R - (\tau+\mu)I \\ R' &= \tau I - \sigma\lambda R - (\mu+\omega+\alpha)R \\ V' &= v\mu - \sigma\lambda V - (\mu+\alpha)V \end{cases} \tag{2.1}$$

A schematic version of the model is represented in Figure 2.1 and Table 2.1 summarizes the parameter definitions. The model is parameterized by the transmission coefficient (β), which differentiates regions/countries according to socioeconomic and environmental factors and re-

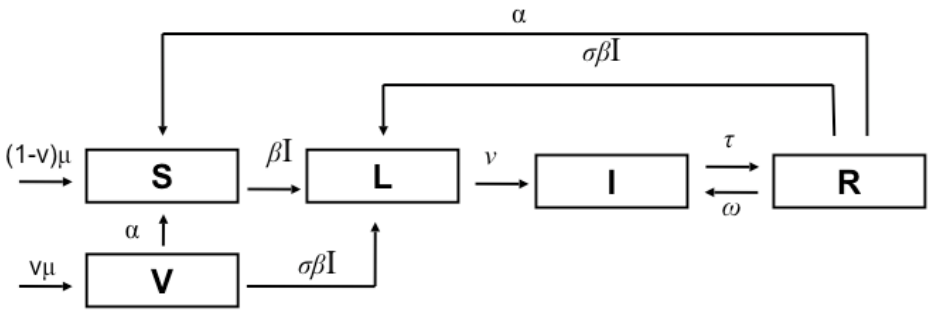


Figure 2.1: SIRI Generalized Model. S, L, I and R stand for the proportion of susceptible, latent infected, infectious and recovered individuals in the population. The model parameters are described in Table 2.1.

flects itself on the force of infection $\lambda = \beta I$. For all sub-models we assume that infection

Table 2.1: SIRI generalized model parameters

parameter	definition
μ	birth and death rates
β	transmission coefficient
σ	factor reducing the risk of reinfection
τ	rate of recovery
ν	rate at which latent individuals progress to the infectious state
α	rate of loss of acquired immunity
ω	rate of endogenous reactivation/relapse after treatment
v	vaccination coverage

confers partial immunity against subsequent reinfections. Parameter σ is the factor reducing the risk of infection as a result of acquired immunity. The birth and death rates are assumed to be equal (μ), rendering the total population constant over time. Before entering the infectious state individuals can remain latent for a period $1/\nu$, where ν is the rate of progression from the latent to the infectious stage. Infectious individuals recover from the infectious state at a rate τ to the recovered class. For the temporary immunity model, recovered individuals loose their protection at a rate α, moving from the recovered back to the susceptible class. Finally, for the ractivation/relapse process, recovered individuals have a chance of relapse back to the infectious stage at a rate ω, independently of any infectious contact. We also include in the model the possibility of vaccinating a proportion v of the population at birth. In all cases, vaccinated individuals are assumed to have the same protection as that conferred by natural infection.

2.3 The epidemic threshold

2.3.1 The basic reproduction number, R_0

The basic reproduction number can be formally define as the dominant eigenvalue of the so called *next generation operator* associated with the system (Diekmann *et al.*, 1990). We use the method described in van den Driessche & Watmough (2002) to compute the next generation matrix, corresponding to this operator. First, we must identify new infections and distinguish the infected classes from all other class transitions in the population. The infected classes are L, I and also R if $\omega \neq 0$. So we write system (2.1) as

$$X' = f(X) \Leftrightarrow X' = \mathcal{F}(X) - \mathcal{V}(X), \tag{2.2}$$

where $X = (L, I, R, S)$, $\mathcal{F} = (\beta IS, 0, 0, 0)^T$ is the rate of appearance of new infections in each class and the disease-free equilibrium is $X_0 = (0, 0, 0, 1)$.

Derivatives $D\mathcal{F}(X_0)$ and $D\mathcal{V}(X_0)$ can be partitioned as

$$D\mathcal{F}(X_0) = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix}, \quad D\mathcal{V}(X_0) = \begin{bmatrix} V & 0 \\ J_3 & J_4 \end{bmatrix},$$

where F and V correspond to the derivatives of $\mathcal{F}$ and $\mathcal{V}$ with respect to the infected classes:

$$F = \begin{bmatrix} 0 & \beta & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} \nu + \mu & 0 & 0 \\ -\nu & \tau + \mu & -\omega \\ 0 & -\tau & \mu + \omega + \alpha \end{bmatrix}.$$

Now, the basic reproduction number is defined as the spectral radius of the next generation matrix, FV^{-1}:

$$R_0 = \beta \frac{\nu(\mu + \omega + \alpha)}{(\nu + \mu)[(\mu + \tau)(\mu + \alpha) + \mu\omega]}. \tag{2.3}$$

Sometimes, the basic reproduction number can be easily computed as the contact number (Hethcote, 2000), by multiplying the contact rate, the probability of becoming infectious and the average infectious period. In this case, the contact rate is β, the average fraction surviving the latent period and arriving to I is $\nu/(\nu + \mu)$ with an infectious period of

$$\frac{1}{\tau + \mu}\Big[1 + \frac{\tau}{\tau + \mu}\frac{\omega}{\mu + \omega + \alpha} + \Big(\frac{\tau}{\tau + \mu}\frac{\omega}{\mu + \omega + \alpha}\Big)^2 + \dots\Big] = \frac{\mu + \omega + \alpha}{(\mu + \tau)(\mu + \alpha) + \mu\omega}, \tag{2.4}$$

which stands for the usual infectious period $1/(\tau + \mu)$, times the multiple opportunities to go back to the infectious class by endogenous reactivation (ω).

Note that if we do not consider endogenous reactivation ($\omega = 0$) these calculation are no longer valid. For the next generation method, the R class is not an infected class and for the direct method the infectious period is just $1/(\tau + \mu)$. Making the corresponding changes the basic reproduction number becomes

$$R_0(\omega = 0) = \beta \frac{\nu}{(\nu + \mu)(\mu + \tau)}. \tag{2.5}$$

For the case $\nu \to +\infty$ the calculations should also be adapted. When there is no latency period the probability to survive this stage is one, hence the resulting expression for R_0 is

$$R_0(\nu \to +\infty) = R_0 = \beta \frac{\mu + \omega + \alpha}{(\mu + \tau)(\mu + \alpha) + \mu\omega}. \tag{2.6}$$

The expression of R_0 is model dependent, so it should be interpreted as a way to classify different populations according to their potential for transmission, under the assumptions made in the model.

2.3.2 Stability of the disease-free equilibrium

The stability properties of the disease-free equilibrium relate to the threshold condition $R_0 = 1$ and are given in the following theorem.

Theorem 2.3.1. *The disease-free equilibrium* $(1, 0, 0, 0)$ *of system (2.1) with no vaccination* $(v = 0)$ *is locally asymptotically stable if* $R_0 < 1$ *and it is unstable for* $R_0 > 1$.

Proof. By theorem 2 in van den Driessche & Watmough (2002), for a system $\dot{X} = f(X)$ it is sufficient to prove conditions:

(A1) if $X \geq 0$, then $\mathcal{F}$, $\mathcal{V}^+$, $\mathcal{V}^- \geq 0$

(A2) if $X_i = 0$ then $\mathcal{V}_i^- = 0$ (where i refers to a vector component)

(A3) $\mathcal{F}_i = 0$ for the components that correspond to uninfected classes

(A4) if X^* is a disease-free equilibrium then $\mathcal{F}_i(X^*) = 0$ and $\mathcal{V}_i^+(X^*) = 0$ for the components that correspond to uninfected classes

(A5) If $\mathcal{F}$ is set to zero then all eigenvalues of $Df(X_0)$ have negative real parts

with $\mathcal{V}(X) = \mathcal{V}^-(X) - \mathcal{V}^+(X)$, where $\mathcal{V}^+$ is the rate of transfer into each class by all other means and $\mathcal{V}^-$ is the rate of transfer out of each class.

For $\omega > 0$, the verification of (A1)-(A4) is straightforward using $\mathcal{F}$, $\mathcal{V}$ and $X^* = X_0$ defined as before. As for condition (A5), the Jacobian of f at X_0 with $\mathcal{F}$ set to zero, as

$$Df_{(\mathcal{F}=0)}(X_0) = \begin{bmatrix} -(\nu+\mu) & 0 & 0 & 0 \\ \nu & -(\tau+\mu) & \omega & 0 \\ 0 & \tau & -(\mu+\omega+\alpha) & 0 \\ 0 & \beta & \alpha & -\mu \end{bmatrix}.$$

The eigenvalues are $-\mu$, $-(\gamma+\mu)$ and the solutions of equation $p(\lambda) = 0$, where $p(\lambda) = \lambda^2 + a_1\lambda + a_0$ and $a_1 = 2\mu + \alpha + \tau + \omega$, $a_0 = (\mu+\tau)(\mu+\alpha) + \mu\omega$. Since a_1 and a_0 are positive, all eigenvalues have negative real part and the result follows for the case $\omega > 0$.

If $\omega = 0$, we define $\mathcal{F}$, $\mathcal{V}$ and $X^* = X_0$ as in the previous case, just F and V are different since R is no longer considered an infected class. Hence, the verification of (A1)-(A4) is as before. As for condition (A5), the Jacobian of f at X_0 with $\mathcal{F}$ set to zero, is now

$$Df_{(\mathcal{F}=0)}(X_0) = \begin{bmatrix} -(\nu+\mu) & 0 & 0 & 0 \\ \nu & -(\tau+\mu) & 0 & 0 \\ 0 & \tau & -(\mu+\alpha) & 0 \\ 0 & \beta & \alpha & -\mu \end{bmatrix}.$$

The matrix is triangular, so the eigenvalues correspond to the diagonal entries which are all real and negative. □

2.4 The reinfection threshold

2.4.1 The SIRI sub-model

By using $\omega = 0$, $\alpha = 0$ and $\nu \to +\infty$, we recover the simple SIRI model

$$\begin{cases} S' &= (1-v)\mu - \lambda S - \mu S \\ I' &= \lambda S + \sigma\lambda(R+V) - (\tau+\mu)I \\ R' &= \tau I - \sigma\lambda R - \mu R \\ V' &= v\mu - \sigma\lambda V - \mu V. \end{cases} \tag{2.7}$$

System (2.7) without vaccination has a disease-free equilibrium $E_0 = (1, 0, 0, 0)$ and an endemic equilibrium E_1. At the epidemic threshold $R_0 = 1$, as we define in the previous section, the system undergoes a transcritical bifurcation. As we have showed for the generalized model, the disease-free equilibrium is stable for $R_0 < 1$ and for $R_0 > 1$ a stable endemic equilibrium emerges. Note that in the epidemiological context we refer to endemic equilibrium as a positive solution of the system at equilibrium and we give no meaning to possible negative states. In the next theorem we prove the existence and stability of the endemic equilibrium.

Theorem 2.4.1. *If $R_0 > 1$, system (2.7) has exactly one endemic equilibrium E_1 that is stable for $R_0 > 1$.*

Proof. Existence. From the first, third and forth equations of system (2.7) at equilibrium, we get a relation between S, R, V and I:

$$S = \frac{\mu}{\mu+\beta I} = s(I), \quad R = I\frac{\tau}{\mu+\sigma\beta I} = r(I)I, \quad \text{and} \quad V = 0.$$

From the second equation of the system (2.7) we get

$$I\Big(\beta s(I) + \sigma\beta r(I)I - (\tau+\mu)\Big) = 0 \Leftrightarrow -\mu I\frac{P(I)}{Q(I)} = 0, \tag{2.8}$$

where P and Q are polynomials of second degree such that:

$$\begin{aligned} Q(I) &= (\mu+\beta I)(\mu+\sigma\beta I), \\ P(I) &= p_2(\beta)I^2 + p_1(\beta)I + p_0(\beta), \end{aligned}$$

with $p_2(\beta) = \sigma\beta^2 > 0$, $p_1(\beta) = \beta(-\sigma\beta + \tau + \mu + \sigma\mu)$, and $p_0(\beta) = \mu((\tau+\mu) - \beta)$. If $I = 0$ we get the disease-free equilibrium. For $I > 0$ then $Q(I) \neq 0$ and we look for positive solutions of $P(I) = 0$. If $R_0 \leq 1$, then $p_0(\beta) \geq 0$ but also $p_1(\beta) \geq 0$, for $0 < \sigma < 1$. So, there are no positive solutions of $P(I)$. If $R_0 > 1 \Leftrightarrow \beta > \tau + \mu$, then $p_0(\beta) < 0$ and we have exactly one positive solution of $P(I)$:

$$I^* = \frac{-p_1 + \sqrt{p_1^2 - 4p_2p_0}}{2p_2}. \tag{2.9}$$

Stability. Lets compute the Jacobian matrix of system (2.7) at $E_1 = (S^*, I^*, R^*, 0)$

$$J = \begin{bmatrix} -\beta I^* - \mu & -\beta S^* & 0 & 0 \\ \beta I^* & \beta S^* + \sigma\beta R^* - (\tau+\mu) & \sigma\beta I* & 0 \\ 0 & \tau - \sigma\beta R^* & -\sigma\beta I^* - \mu & 0 \\ 0 & 0 & 0 & -\sigma\beta I^* - \mu \end{bmatrix}$$

The eigenvalues of J are $-\mu$,$-(\sigma\beta I^* + \mu)$, that are real and negative, and the solution of the polynomial $p(\lambda) = \lambda^2 + a_2\lambda + a_0$, where $a_1 = \beta I^*(\sigma + 1) + 2\mu + \tau - \sigma\beta R^* - \beta S$ and $a_0 = -\sigma\beta^2 I^*(S^* + R^*) - \mu\beta S^* - \sigma\mu\beta R^* + \sigma\beta^2 I^{*2} + (\beta\tau + \sigma\beta\mu)I^* + \mu(\mu + \tau)$. From second equation of the system at equilibrium, we can rewrite $a_1 I^*$ as $\beta I^{*2}(\sigma + 1) + \mu I^* > 0$, from which we conclude that a_1 is positive. Using the fact that $P(I^*) = 0$ and $1 = S^* + R^* + I^*$, we get $a_0 = -\sigma\beta^2 I^*(1 - I^*) - \mu\beta S^* - \sigma\beta R^* + (P(I^*) + \sigma\beta^2 I^* + \beta\mu)$. Finaly, by substituting $1 - S^* = I^* + R^*$, we conclude that $a_0 = \sigma\beta^2 I^{*2} + \mu\beta I^* + \mu\beta R^*(1 - \sigma)$ which is positive for $0 < \sigma < 1$. Hence $p(\lambda)$ has solutions with negative real part and the result follows. □

The reinfection threshold is associated with the epidemic threshold in the reinfection sub-model (Gomes *et al.*, 2004b, 2005)

$$\begin{cases} I' &= \sigma\lambda(R + V) - (\tau + \mu)I \\ (R + V)' &= \mu + \tau I - \sigma\lambda(R + V) - \mu(R + V), \end{cases} \tag{2.10}$$

which is obtained from system (2.7) by setting $v = 1$. The bifurcation and therefore the RT takes place at

$$R_0 = \frac{1}{\sigma}. \tag{2.11}$$

that corresponds to the transmission level above which transmission can be sustained in a population for which individuals are born with partial protection σ. One of the characteristics

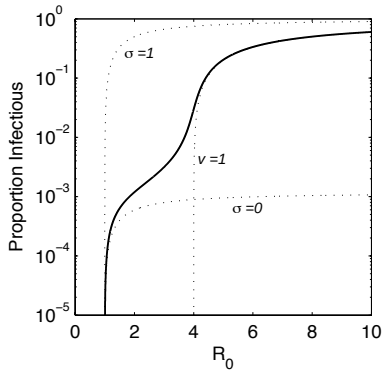


Figure 2.2: Equilibrium curve for the SIRI model. Dotted lines correspond to the SIS model ($\sigma = 1$), SIR model ($\sigma = 0$) and the reinfection sub-model ($v = 1$). Parameters used are $\mu = 1/70$, $\tau = 12$ and $\sigma = 0.25$.

of the SIRI model is the steep increase in disease prevalence that occurs when the reinfection threshold is crossed. The SIRI model changes from a low endemic level typical of an SIR model (lower dotted curve in Figure 2.2 obtained for $\sigma = 0$) to a high endemic level typical of an SIS model (top dotted curve in Figure 2.2 obtained for $\sigma = 1$).

The difference in the disease prevalence levels, below and above the RT, depends on population and disease factors. Figure 2.3 describes how this difference changes with two of these factors: the average life span $(1/\mu)$ and the average duration of the infectious period $(1/\tau)$. Later on, we comment on the impact of other disease related factors. To characterize the potential for variation due to partial immunity we define the ratio

$$k = \log_{10} \frac{I^1(R_0)}{I^0(R_0)}, \tag{2.12}$$

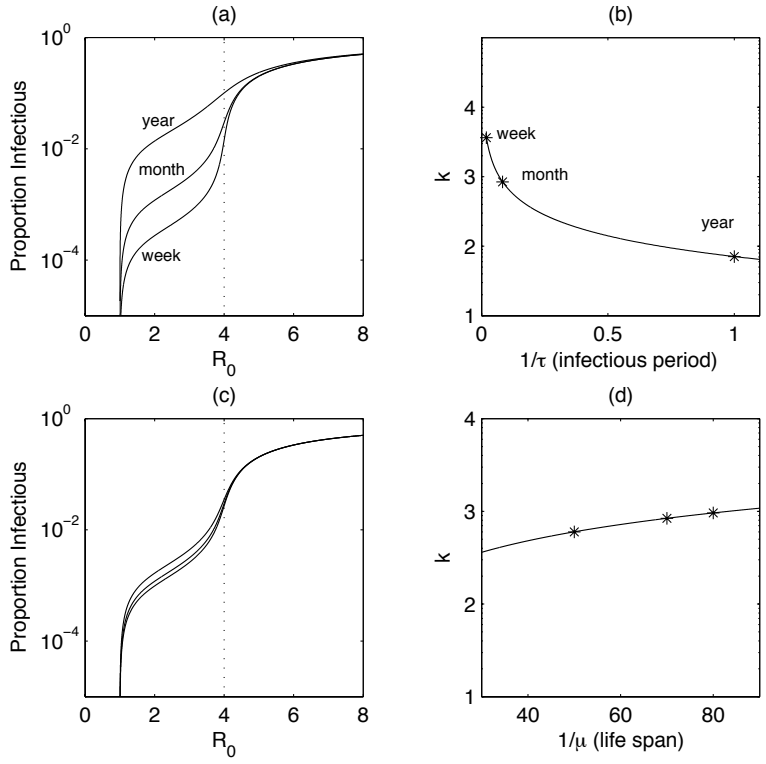


Figure 2.3: Impact of the infectious period and average life expectancy on the disease level above and below the RT.**(a)** and **(c)** Equilibrium infectious proportion for different values of $\tau : 1, 12, 52$ (corresponding to an average duration of infection of one year one month and one weak, respectively) and $\mu = 1/50, 1/70, 1/80$. **(b)** and **(d)** How the difference in the disease prevalence below and above the RT is affected by the infectious period and the average life expectancy.

where $I^0(R_0)$ and $I^1(R_0)$ correspond to the equilibrium proportions of infectious for the SIR ($\sigma = 0$) and SIS ($\sigma = 1$) model extremes. We can obtain analytic expressions for I^0 and I^1 from the proof of the endemic equilibrium existence for the SIRI model (Theorem 2.4.1) as the roots of polynomial $P(I)$:

$$I^0(R_0) = \frac{\mu}{\tau + \mu}\Big(1 - \frac{1}{R_0}\Big) \quad \text{and} \quad I^1(R_0) = 1 - \frac{1}{R_0}. \tag{2.13}$$

Interestingly the ratio between them results independent of R_0, $k \equiv \log_{10} \dfrac{\tau + \mu}{\mu}$.

Figure 2.3 (a) and (c) show the equilibrium curves for different values of μ and τ and the corresponding (b) and (d) panels show the theoretical ratio, k. The variation in disease prevalence decreases with the infectious period duration but it increases with the average life expectancy. In the second case, the impact is much less for biologicaly meaningful values than in the first case, for which differences can go from 2 to 4 orders of magnitude for infectious periods of one year or one weak, respectively.

For the purpose of illustration in the remaining plots in this section we fix $\mu = 1/70$, $\tau = 12$ and $\sigma = 0.25$, corresponding to an intermediate situation.

2.4.2 The SIRI model with latency

We now consider the case for which there is a latency period. While in the latent class (L), individuals cannot transmit the disease and are protected against reinfection. The average latency period is $1/\nu$ yrs. Temporary immunity and reactivation are not taken into account ($\alpha = 0$ and $\omega = 0$)

$$\begin{cases} S' &= (1-v)\mu - \lambda S - \mu S \\ L' &= \lambda S + \sigma\lambda(R+V) - (\nu+\mu)L \\ I' &= \nu L - (\tau+\mu)I \\ R' &= \tau I - \sigma\lambda R - \mu R \\ V' &= v\mu - \sigma\lambda V - \mu V. \end{cases} \tag{2.14}$$

All infectious diseases have a latency period and, therefore, this model is widely applicable.

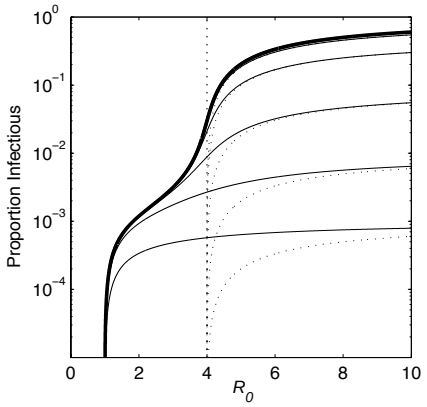


Figure 2.4: Equilibrium Proportion of Infectious for different latency periods. From top to bottom ν=120, 12, 1.2, 0.12 and 0.012. Heavy line represents the limit case SIRI with no latency period ($\nu \to \infty$). Dotted curves correspond to the equilibrium of the respective vaccination sub-model with $v = 1$ and vertical line marks the RT.

Figure 2.4 shows how the infectious equilibrium curve changes with the latency period. First, we observe that the reinfection threshold defined for the simple SIRI model, $R_0 = 1/\sigma$ (vertical line in Fgure (2.4)), generally corresponds to an increase in the disease level. However, its effects depend on parameters values. As we increase the latency period, the equilibrium is lower and the increase at the RT is not as pronounced. Note that for the lower curve we used $\nu = 0.012$ (and $\tau = 12$) which corresponds to a latency period of 83 yrs $(1/\nu)$ or, more realistically, it means that an infected individual has only 46% $(\nu/(\nu+\mu))$ chances of progressing to disease in its life span. Despite reinfection being possible it is not the most important mechanism in terms of the contribution to disease burden since for the majority of the population the infection is chronic. Nevertheless, for all cases the RT marks a critical transmission for the impact of vaccination, assuming that vaccine confers partial protection equivalent to natural immunity. The bifurcation in the reinfection sub-model ($v = 1$) marks the transmission intensity above which the disease remains endemic (dotted lines in Figure 2.4) revealing the RT for $R_0 = 1/\sigma$, independently of the duration of the latency period.

2.4.3 The SIRI model with temporary immunity

We now consider the case for which after recovering from infection, individuals have partial protection against reinfection but this protection wanes and is eventually lost. In the temporary immunity sub-model we do not consider the existence of a latency period ($\nu \to \infty$) nor the possibility of relapse after recovery ($\omega = 0$)

$$\begin{cases} S' &= (1-v)\mu + \alpha(R+V) - \lambda S - \mu S \\ I' &= \lambda S + \sigma\lambda(R+V) - (\tau+\mu)I \\ R' &= \tau I - \sigma\lambda R - (\mu+\alpha)R \\ V' &= v\mu - \sigma\lambda V - (\mu+\alpha)V. \end{cases} \tag{2.15}$$

There are numerous diseases for which immunity is thought to be partial and temporary such as pertussis, malaria or dengue.

Figure 2.5 shows the equilibrium proportion of infectious for different values of the rate of waning immunity, α. Disease level is higher for shorter protection periods, above limited by the SIS scenario when $\alpha \to \infty$ and below by SIRI scenario when $\alpha = 0$.

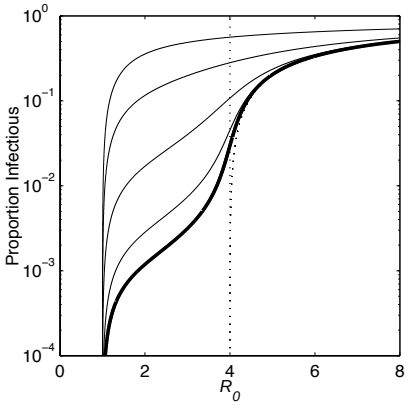


Figure 2.5: Equilibrium proportion of infectious for different rates of waning immunity in relation to R_0. From top to bottom α = 20, 2, 0.2 and 0.02. Heavy line represents the limit case SIRI with no waning immunity ($\alpha = 0$). Dotted curve corresponds to the equilibrium of the respective vaccination sub-model with $v = 1$ and vertical line marks the RT for the extreme case $\alpha = 0$.

The epidemic threshold $R_0 = \beta/(\tau+\mu) = 1$ does not depend on the α. And the RT of the SIRI model generally marks a behavioral change for the temporary immunity sub-model (vertical dotted line in Figure 2.5). The RT corresponds to the bifurcation in the reinfection sub-model $R_0 = 1/\sigma$, obtained by setting $v = 1$ but also $\alpha = 0$. Hence, the RT no longer coincides with the vaccination control limit, for a vaccine conferring the same protection as natural infection (i.e. partial and temporary). For a certain vaccination coverage v the region for which elimination is possible is limited by the epidemic threshold of system (2.15) corresponding to the disease free-equilibrium $(\frac{(1-v)\mu+\alpha}{\mu+\alpha}, 0, 0, \frac{v\mu}{\mu+\alpha})$ attained at

$$R_0 = \frac{\mu+\alpha}{(1-v)\mu+\alpha+\sigma v\mu}. \tag{2.16}$$

Hence, the control limit corresponds to $R_0 = (\mu+\alpha)/(\sigma\mu+\alpha)$ when $v = 1$, here defined as the vaccination threshold (VT).

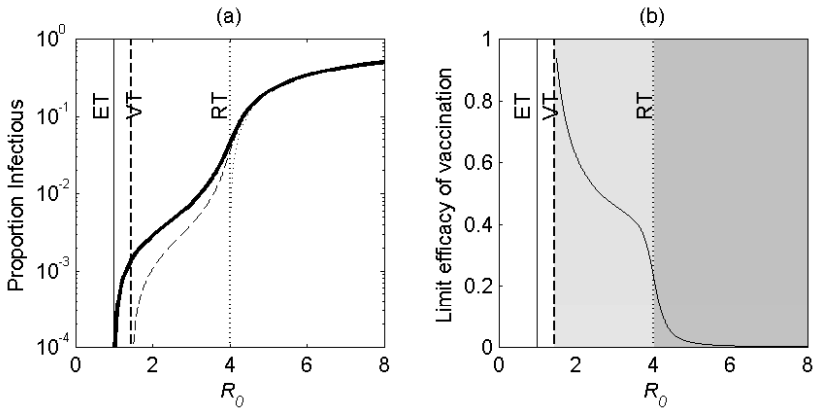


Figure 2.6: Vaccination versus reinfection threshold for the temporary immunity sub-model model (for $\alpha = 0.2$). **(a)** Heavy, light full and dotted curves correspond to the cases no vaccination ($v = 0$), limit vaccination coverage ($v = 1$) and reinfection sub-model ($v = 1$ and $\alpha = 0$). Vertical lines mark the vaccination and the reinfection thresholds (VT and RT). **(b)** Limit vaccination efficacy ($1 - I_{v=1}/I_{v=0}$) in relation to R_0.

Figure 2.6 illustrates the difference between the vaccination and the reinfection thresholds for a 5 year immunity duration ($\alpha = 0.2$). There is a transmission intensity range, $R_0 \in (VT, RT)$, for which vaccination is not able to eliminate the disease but it can reduce significantly its level (light grey region in Figure 2.6). As the RT is crossed, there is a steep decrease on vaccination impact. Therefore, in this context, RT marks the level of transmission above which partial immunity impairs control by vaccination, independently of temporary protection (dark grey region in Figure 2.6). In general, we note that the VT is always lower than the RT. Similarly, the effort needed to eliminate the disease by vaccination, for a population with a transmission intensity given by a certain R_0, depends on both temporary and partial immunity and it is higher than the vaccination coverage needed to eliminate the disease in the simple SIRI model. The effort to eliminate the disease is defined by the vaccination coverage $v(\alpha, \sigma)$ needed to reduce the epidemic threshold of the model with vaccination (2.16) below one. Hence,

$$v(\alpha,\sigma) = \Big[1 - \frac{1}{R_0}\Big]\frac{\mu+\alpha}{\mu(1-\sigma)} \;\geq v(0,\sigma) = \Big[1 - \frac{1}{R_0}\Big]\frac{1}{(1-\sigma)}. \tag{2.17}$$

For an improved vaccine that confers permanent immunity VT and RT coincide and represent the transmission potential below which elimination can be attained.

2.4.4 The SIRI model with endogenous reactivation/relapse

We now consider the possibility of endogenous reactivation or relapse after treatment, independently of infectious contact. Latency and temporary immunity are omitted ($\nu \to \infty$ and $\alpha = 0$)

$$\begin{cases} S' &= (1-v)\mu - \lambda S - \mu S \\ I' &= \lambda S + \sigma\lambda(R+V) + \omega R - (\tau+\mu)I \\ R' &= \tau I - \sigma\lambda R - (\mu+\omega)R \\ V' &= v\mu - \sigma\lambda V - \mu V. \end{cases} \tag{2.18}$$

Tuberculosis and varicella/zoster are examples of infections for which individuals can remain infected even after effective treatment of an active disease episode. These individuals cannot transmit, but infection can be endogenously reactivated back to the disease/infectious state independently of new infectious contacts. Reactivation is thought to depend mostly on immunosupressive factors such as age, co-infection with HIV or immunosupressive treatments. In the case of TB similar phenomena can also be associated with treatment failure or relapse. In what concerns transmission, the possibility of reactivation/relapse results in a new infectious period.

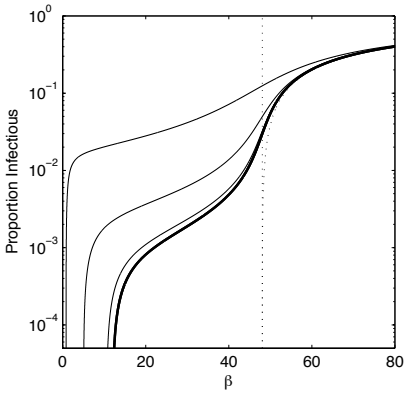


Figure 2.7: Equilibrium proportion of infectious for different relapse rates in relation to the transmission coefficient (β). From top to bottom ω= 0.2, 0.02 and 0.002. Heavy line represents the limit case SIRI with no relapse ($\omega = 0$).

Figure 2.7 shows the equilibrium proportion of infectious for different values of the reactivation/relapse rate ω in relation to the transmission coefficient β (instead of the basic reproduction number). The change in behavior typical of the RT is observed at a fixed value $\beta = (\tau + \mu)/\sigma$, independently of the reactivation rate. Conversely, the epidemic threshold varies with the reactivation rate, $R_0 = \beta(\mu + \omega)/\mu(\mu + \omega + \tau) = 1$. Hence, for the reactivation/relapse sub-model we no longer expect that the RT has the same relation with R_0 as before. In fact, the reinfection sub-model must be obtained by setting the reactivation/relapse rate to zero ($\omega = 0$ and $v = 1$).

$$\begin{cases} I' &= \sigma\lambda(R+V) - (\tau+\mu)I \\ (R+V)' &= \mu + \tau I - \sigma\lambda(R+V) - (\mu+\omega)(R+V). \end{cases} \tag{2.19}$$

The RT is then attained at

$$R_0 = \frac{1}{\sigma}\Big[1 + \frac{\tau\omega}{\mu(\tau+\mu+\omega)}\Big] \ (\geq \frac{1}{\sigma}) \tag{2.20}$$

meaning that endogenous reactivation/relapse shifts the impact of reinfection to higher transmission intensities (see also Figure 2.8 (a)).

Figure 2.8 illustrates the effects of reinfection in vaccination impact, where vaccinated individuals are assumed to have partial protection equivalent of that induced by natural infection. The average time until reactivation is assumed to be $1/\omega = 50\ yrs$. Disease can be eliminated for R_0 below the vaccination threshold $R_0 < VT = 1/\sigma$, and the effort to eliminate

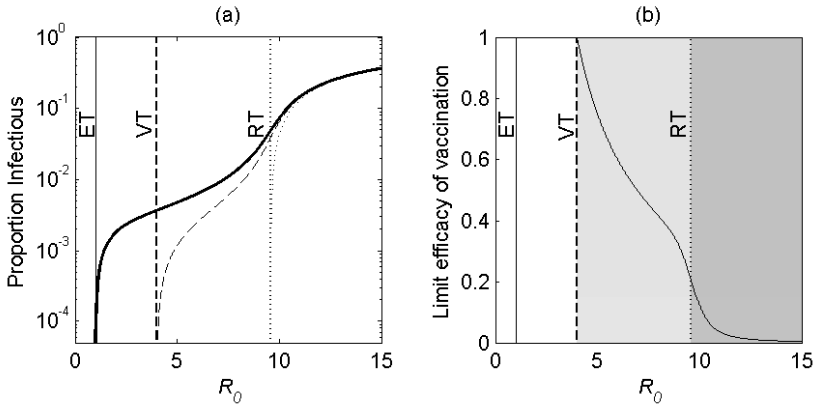


Figure 2.8: Vaccination versus reinfection threshold for the reactivation/relapse sub-model (for $\omega = 0.002$.) **(a)** Heavy, full and dashed curves correspond to the cases no vaccination, limit vaccination coverage $v = 1$ and reinfection sub-model ($v = 1$ and $\alpha = 0$). Vertical lines mark the vaccination and the reinfection thresholds. **(b)** Limit vaccination efficacy $(1 - I_{v=1}/I_{v=0})$ in relation to R_0.

the disease for a certain R_0 depends on both the reactivation rate and on the protection factor $v(\omega, \sigma) = \dfrac{1 - 1/R_0}{1 - \sigma}$. For $R_0 \in (VT, RT)$ (light grey region) the disease can be controlled but not eliminated, independently of the vaccination coverage used. Moreover, when RT is crossed the impact of vaccination is totally overcome by reinfection (dark grey region).

2.5 How to compute the reinfection threshold

In its original formulation RT is defined as the bifurcation in the transmission parameter for the reinfection sub-model, obtained by setting $v = 1$ for the vaccination model (Gomes *et al.*, 2004b, 2005). However, if the protection conferred by the vaccine is different from natural immunity or if there are other immunity processes involved such as temporary immunity or endogenous reactivation, the corresponding bifurcation for $v = 1$ (referred here as vaccination threshold (VT)), diverges from the RT. The reinfection sub-model must be defined as the sub-model for which all individuals are partially immunized. Since it should only reflect the reinfection impact, additional immune processes must be removed. Latency, for example, is considered since it is not a competing immunity process but a delay in the onset of infectiousness. For the generalized model (2.1) the reinfection sub-model is then defined as

$$\begin{cases} R' &= \mu + \tau I - \sigma\lambda R - \mu R \\ L' &= \sigma\lambda R - (\nu + \mu)L \\ I' &= \nu L - (\tau + \mu)I \end{cases} \tag{2.21}$$

The bifurcation of the reinfection sub-model is attained at $\beta = \dfrac{(\tau + \mu)(\nu + \mu)}{\nu\sigma}$. The RT corresponds the critical transmission level above which transmission can be sustained in a partially immune population. Since the RT is invariable for the different immunity processes other than reinfection, the relationship between RT and R_0 depends on how these processes

affect the epidemic threshold. Hence, the relation $R_0 = 1/\sigma$ is maintained for the case of temporary immunity but is changed for the reactivation/relapse model.

For systems where more than one partially immune class exists, multiple reinfection thresholds can be defined. These RTs are associated with bifurcations of the reinfection sub-models defined by each partially immune class, from which other competing immunity processes are set to zero. The SIRI model with heterogeneity in susceptibility to infection, in Section 5.5.2 and the tuberculosis model for post-exposure interventions, in Gomes *et al.* (2007), are examples of models with multiple reinfection thresholds. Overall, the importance of the multiple reinfection thresholds to the model behavior is determined by the relative size of the corresponding partially immune classes. Interventions that can alter the susceptibility profile of the population have the potential to create regions of bistability, as suported by the examples mentioned above. The study of the multiple RTs can reveal these regions of interest.

2.5.1 The reinfection sub-model

Generally, we can define the reinfection sub-model and corresponding RT for any given compartmental disease transmission model based on a system of ordinary differential equations. The model can accommodate heterogeneous populations in which the demographic and epidemiological parameters reflect a dependence on factors such as stage of the disease, social condition, age, behavior. We assume that the population can be divided into homogeneous sub-populations, or compartments, such that individuals in a given compartment are identical. We also assume that the parameters are independent of time. The model is based on a system of ordinary equations describing the evolution of the proportion of individuals in each compartment.

Let the disease transmission model consists of nonnegative initial conditions together with the following system of equations:

$$x_i' = f_i(x),\ i = 1, \ldots, n+m+k, \tag{2.22}$$

where $x = (x_1, ..., x_{n+m+k})^t$ with each $x_i \geq 0$, are the proportion of individuals in each compartment. Compartments are sorted in order that the first n compartments correspond to the partial immunized classes subject to reinfection, followed by the m remaining classes except for the k totally susceptible compartments, which are placed at the end. The distinction between infected, partially immunized and susceptible compartments must be determined from the epidemiological interpretation of the model and cannot be deduced from the structure of the equations alone.

We will construct the reinfection sub-model by first removing the totally susceptible classes and transferring the recruitment rate (source population) into the partial immunized compartments. Let, $y = (x_1, \ldots, x_{n+m})^t$ correspond to the proportion of individuals in each compartment for the sub-model. We define

$$\bar{g}_i(y) = \bar{f}_i(y, 0) + r_i(y, 0) \tag{2.23}$$

for $i = 1, \ldots, n+m$, where r_i is the recruitment rate verifying $r_i = 0$ for $i = n+1, \ldots, n+m$ and $\bar{f}_i$ is given by f_i with the rate of transfer of individuals into the last k compartments set to zero. In particular, this last condition implies that temporary immunity term is removed ($\alpha = 0$). Note that the recruitment functions r_i depend again on the epidemiological interpretation of

the model. If only one class is subject to reinfection, x_1, then r_1 is the sum of all recruitment rates of the model, corresponding to the total inflow of individuals into the model at each time step. If there is a correspondence (social condition, age or behavior) between the totally susceptible classes and partially immunized classes, then $n = k$ and r_i are the recruitment rates of the matching susceptible classes. Finally, we can have different compartments subject to reinfection with different protection or disease progression. In this case, for each of these classes alternative reinfection sub-models should be considered, giving rise to several RTs. For each sub-model, there should be only one or a matching number of reinfection classes of interest at a time, so the recruitment function will be defined as in the simpler cases.

The reinfection sub-model is then defined by the following set of $n+m$ differential equations

$$y_i' = g_i(y) = g_i^+(y) - g_i^-(y),\ i = 1, \ldots, n+m, \tag{2.24}$$

where $y = (y_1, \ldots, y_{n+m})^t$ and g_i^+ and g_i^- are the rates of transfer of individuals in and out of compartment i, respectively. Since each function represents a directed transfer of individuals proportion, they are all non-negative. Thus, we have

(i) if $y_i \geq 0$, then $g_i^+,\ g_i^- \geq 0$, for $i = 1, \ldots n+m$.

If a compartment is empty there can be no transfer out, thus

(ii) if $y_i = 0$, then $g_i^-(y_i) = 0$, for $i = 1, \ldots n+m$.

Consider the disease transmission model given by (2.24) with g_i satisfying conditions (i) and (ii), hence the non-negative cone ($y_i \geq 0$, $i = 1, \ldots, n+m$) is forward invariant. By Theorems 1.1.8 and 1.1.9 of Wiggins (1990) for each nonnegative initial condition there is a unique, non-negative solution of system (2.24).

Let us now define Y_0 to be the set of all disease free states, that is $Y_0 = \{y \geq 0 : y_i = 0,\ i = n+1 \ldots, n+m\}$. To ensure that the disease free subspace Y_0 is invariant, we assume that if the population is free of disease then the population will remain free of disease. No (density independent) immigration of infectives is allowed. This condition is stated as follows:

(iii) if $y \in Y_0$, then $g_i^+ = 0$, for $i = n+1, \ldots n+m$.

This condition implies, in particular, that some immune processes must be set to zero in the reinfection sub-model. As it is the case for endogenous reactivation/relapse rates.

2.5.2 The reinfection threshold

Let us construct the reinfection sub-model (2.24), satisfying conditions (i)-(iii). If the reinfection sub-model undergoes a bifurcation in the transmission parameter β, then this bifurcation will correspond to the RT for the original model. More rigorously, let us consider the linearized system

$$y' = Dg(y_0)(y - y_0), \tag{2.25}$$

where $Dg(y_0)$ is the Jacobian matrix evaluated at the disease free state $y_0 \in Y_0$. The bifurcation point for $y = y_0$ will be determined by setting to zero of the determinant of $Dg(y_0)$and solving the resulting equation for β. We denote the solution by β_{RT}, which will corresponds to a stability change of the disease free-equilibrium.

Finally, to obtain the relation between the RT and the basic reproduction number of the original model we must substitute β by the expression of β_{RT} in the R_0 formula.

2.6 Conclusions and outlook

Here we present a simple framework on how to include partial immunity in simple epidemiological models and how to separate the impact of reinfection from other disease processes such as latency, temporary immunity or endogenous reactivation, through the computation of the reinfection threshold. RT marks a critical transmission intensity level corresponding to a change from a SIR to an SIS transmission regime and has important consequences to disease endemic level and interventions effectiveness, even in the presence of other disease processes.

The manifestation of the characteristics associated with the reinfection threshold changes with disease factors such as the duration of the infectious period and life expectancy. The magnitude of this manifestation is affected by other immunity factors, as described in the temporary immunity and endogenous reactivation sub-models, reflecting the relative importance of reinfection in the context of each disease. Even if not evident, the two transmission regions defined by the RT have always considerably different dynamics that can be evidenced by massive interventions.

The basic reproduction number, R_0 has served as a reference quantity to define the control effort needed to eliminate a disease, in particular, by defining the vaccination coverage. However, for diseases where immunity is not fully protective, the control effort depends also on the factor of protection of the partial immunized population (σ). The RT defines the limit of the vaccination success, which corresponds to a steep decrease on the vaccine impact and the impossibility to eliminate the disease, even if it overcomes the temporary or reactivation/relapse effects of natural immunity.

When a population has different susceptibility groups with distinct protections there are multiple reinfection thresholds. These thresholds refer to the transmission intensity above which transmission can be sustained in each of the groups, independently. The impact on the overall behavior of the system can be imposed by just one or a part of these thresholds, depending on the relative size of the corresponding group. Interventions that can change either the protection or the size of these groups have the potential to dramatically change the disease landscape by changing the RT of interest. This point will be more clear after Section 5.5.2, where interventions on heterogeneous populations with different levels of susceptibility to infection, can display catastrophic behavior for a transmission region determined the associated reinfection thresholds. Another example can be found in Gomes *et al.* (2007).

Overall the RT marks a major change in the model behavior, which can help to explain unexpected behaviors for high transmission regions. The use of RT in combination with the classical epidemic threshold provide a better description and understanding of disease with partial immunity, especially in the evaluation of effectiveness of different control measures. Throughout this work we continue to emphasize the role of reinfection in different contexts, in particular in the study of tuberculosis transmission.

Chapter 3

Drug resistance in tuberculosis - a population perspective

3.1 Introduction

3.1.1 Motivation and aims

There is increasing recognition that partial immunity plays an important on TB transmission. It has been shown that reinfection has significant epidemiological consequences, particularly in what concerns disease prevalence and effectiveness of control measures. In this Chapter we explore the impact of partial immunity when the parasite population is heterogeneous with respect to drug sensitivity. We address the problem of drug-resistance as a competition between two types of strains of *Mycobacterium tuberculosis*: those that are sensitive to anti-tuberculosis drugs and those that are resistant. Our objective is to characterize how reinfection modifies the conditions for coexistence of sensitive and resistant strains, by giving an extra opportunity for resistant strains to spread. This sets the scene for discussing how strain prevalence is affected by different control strategies. It is shown that intervention effectiveness is highly sensitive to the baseline epidemiological setting. This chapter is adapted with minor changes from Rodrigues *et al.* (2007).

3.1.2 The epidemiology of drug-resistant tuberculosis

Despite intensive control efforts, recent data show that global TB incidence is increasing, largely associated to the increase in the prevalence of HIV (WHO, 2005) but also to the decrease in treatment efficacy, due to the emergence of multi-drug resistant strains (Dye *et al.*, 2002). According to a recent report of the World Health Organization (WHO/ IUATLD, 2004), the overall prevalence of drug-resistance ranges from 0% (Andorra, Iceland and Malta) to 63.9% (Karakalpakstan, Uzebekistan) with a median of 10.4%. The WHO distinguishes between two types of resistance: acquired resistance – resistance among previously treated patients; and primary resistance – resistance among new cases (WHO/ IUATLD, 1998). In all regions studied, prevalence of acquired resistance is higher than prevalence of primary resistance, but the size of this difference varies between regions (WHO/ IUATLD, 2004).

Treatment of tuberculosis consists of a combination of different drugs to avoid acquisition of resistance. Despite these precautions, drug resistance continues to emerge being favoured by the long duration of treatment and improper use of the antibiotics (Crofton *et al.*, 1997).

Drug resistant TB has higher rates of treatment failure and longer periods of infectiousness in part due to the time lapse between TB diagnosis and obtaining drug-sensitivity test results (Espinal *et al.*, 2000). Most worrisome is resistance to the two first line drugs, isoniazid and rifampicin, defined as multi-drug resistance (MDR). Geographical distribution of MDR is very heterogeneous: it is highly prevalent in several areas of the former Soviet Union and in Israel, Ecuador and some Provinces of China, but it is absent or present with very low prevalence in a significant number of countries. Prevalence of MDR TB ranges from 0% to 26.8%, with a median of 1.7% (WHO/ IUATLD, 2004). More recently extensively drug-resistant TB (XDR-TB) defined as TB resistant to multiple drugs as well as to any one of the fluoroquinolone drugs and to at least one of the three injectable second-line drugs (amikacin, capreomycin or kanamycin) (WHO, 2007), was reported in all regions of the world. It was rapidly classified by WHO as a serious emerging threat to global public health, especially, in countries with a high prevalence of HIV (WHO, 2007). Because XDR-TB is resistant to first- and second-line drugs, treatment options are seriously limited increasing the concern with the spread of these strains.

3.1.3 Transmission models of antibiotic resistance

Mathematical models have addressed the transmission dynamics of antibiotic resistance in general (Austin *et al.*, 1997; Bonhoeffer, 2002; Boni & Feldman, 2005). More specifically to tuberculosis, a number of mathematical models have also been proposed (Blower & Chou, 2004; Blower & Gerberding, 1998; Blower *et al.*, 1996; Castillo-Chavez & Feng, 1997; Cohen & Murray, 2004; Dye & Espinal, 2001; Dye & Williams, 2000). Overall these models assume that resistant strains are less transmissible, reflecting a trade-off between fitness and resistance. Combined results demonstrate that the relative fitness between resistant and sensitive strains is a crucial parameter: for some values it is predicted that second-line drugs would be needed to prevent future epidemics (Dye & Espinal, 2001), whereas for other values it appears as a local problem that can be managed through proper implementation of strategies currently recommended by the WHO (Dye & Williams, 2000). Moreover, Cohen & Murray (2004) find that even when resistant strains have, on average, a lower transmissibility a small subpopulation of a relatively fit MDR strain may outcompete both the drug-sensitive strains and the less fit MDR strains. The relation between resistance acquisition and fitness cost as well as its epidemiological consequences are, however, still under discussion (Cohen *et al.*, 2003; Gagneux *et al.*, 2006).

Although it is recognized that reinfection is an important component of TB transmission (Chiang *et al.*, 2005), few modellers take it into consideration. It has been shown that for infectious diseases where immunity acquired by individuals after exposure is not totally protective, allowing for reinfection to occur at a reduced rate, the equilibrium prevalence of infection is highly sensitive to a threshold other than the epidemic threshold. This has been named the 'reinfection threshold' and marks a critical transmission rate above which reinfection processes are dominant (Gomes *et al.*, 2004a,b, 2005; Breban & Blower, 2005). The reinfection threshold has strong implications on epidemiological reasoning, particularly in what respects the effectiveness of interventions.

For the case of resistant TB, a few models have considered reinfection (Blower & Chou, 2004; Castillo-Chavez & Feng, 1997; Cohen & Murray, 2004; Dye & Williams, 2000) but the implementations vary significantly. Blower & Chou (2004) and Dye & Williams (2000) incorporate reinfection at a reduced rate (partial immunity) applying to latent individuals

only. Blower & Chou (2004) assume that recovered individuals have either total protection against reinfection (if treated), or no protection at all (if self-cured). By contrast, Dye & Williams (2000) assume that self-cured individuals have a high relapse but cannot be reinfected. Castillo-Chavez & Feng (1997) neglect exogenous reinfection of latent individuals and assume super-infection but only by resistant strains. Cohen & Murray (2004) consider that latent and recovered individuals benefit from partial immunity and have identical susceptibilities to reinfection. Reinfection can happen with different strains and the new strain always replaces the previous one. The model characterises strains by both fitness and resistance status reaching a level of complexity that limits its analysis in what reinfection is concerned.

We extend previous work by devoting special care to the implementation of reinfection and to the analysis of its consequences to the spread of drug-resistant tuberculosis. The model is based on a reinfection framework for the transmission of tuberculosis (Gomes *et al.*, 2004a) and extended to describe the competition between two types of strains: sensitive and resistant to drugs. Model extension is made in steps permitting intermediate analysis in a systematic way.We describe how coexistence is shaped by reinfection dynamics and by the outcome of mixed infection. The model predicts that coexistence is common for highly endemic settings due to the greater relative importance of reinfection. Long term effectiveness of different control measures is considered, and shows important sensitivity to the baseline epidemiological setting.

3.2 Model construction

3.2.1 Exogenous reinfection and endogenous reactivation

The host population is divided into different categories based on the individual history of infection. Three classes characterize the host population: susceptible (S), who have never be exposed to the *mycobacterium*; latent (L), who are infected but not infectious; and infectious (I) with active disease (see the diagram in Figure 3.1). Population size is assumed constant over time. Susceptible individuals are infected at a rate proportional to the prevalence of

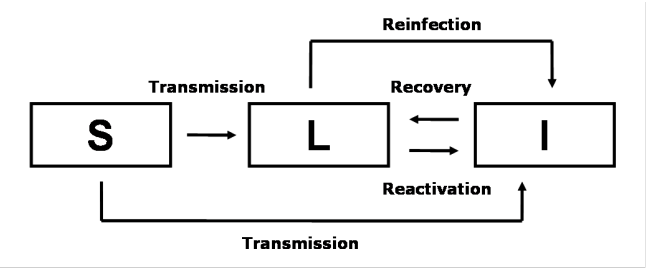


Figure 3.1: TB model. Individuals are classified according to infection state into susceptible (S), latently infected (L) and infectious (I).

active TB and may develop active disease (progress to I) or maintain a latent infection (enter L). Individuals who recover from active disease by treatment with antibiotics or self-cure are transferred from I back to L. Infected individuals acquire some immunity as a result of infection, which reduces the risk of subsequent infection but does not fully prevent it. Finally, latent individuals can progress to active TB due to endogenous reactivation or exogenous reinfection.

Figure 3.2 shows the equilibrium curve for the proportion of active infections and illustrates the reinfection threshold as defined in Chapter 2 and originally computed in Gomes *et al.* (2004a). Above this threshold most TB cases are due to reinfection. Dashed and full thinner lines in this Figure trace the equilibrium proportion of cases resulting from primary infection and reinfection, respectively.

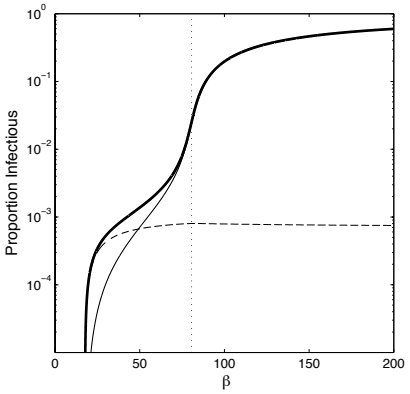


Figure 3.2: Equilibrium curve: heavy black line represents all TB cases. Thin dashed and full lines represent primary and reinfection cases, respectively. Vertical line marks the reinfection threshold.

3.2.2 Drug-resistance

The model is extended to include two strains with different sensitivities to antibiotics (see diagram in Figure 3.3). We specify drug-resistant and drug-sensitive strains by adding subscripts r and s to model variables and parameters.

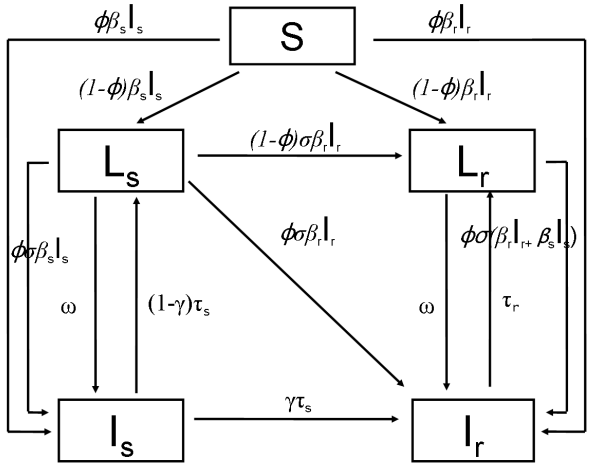


Figure 3.3: Two-strain TB model. Individuals are classified according to infection state into susceptible (S), latently infected (L) and infectious (I). Parameters are the transmission coefficient (β), the death and birth rate (μ), the proportion of individuals developing active TB (ϕ), the reinfection factor (σ), the rate of reactivation (ω), the rate of recovery under treatment (τ) and the proportion of resistance acquisition (γ). Subscripts s and r refer to sensitive and resistant strains, respectively.

Resistant cases may emerge when individuals are infected with a resistant strain (primary resistance) or as a result of treatment failure (acquired resistance). We assume that a fraction, γ, of infectious individuals with active sensitive TB (I_s) progresses into the infectious class of resistant strains (I_r) due to treatment failure. These correspond to cases of acquired resistance.

3.2.3 Strain interactions

Molecular epidemiological studies suggest that mixed infections (infections with more than one strain) are common (Warren *et al.*, 2004), and that once an individual is infected with both sensitive and resistant strains, a differential selection pressure will be imposed by treatment (van Rie *et al.*, 2004). Moreover, an individual infected with both resistant and sensitive strains may have two alternative progressions: (i) develop resistant TB if treated with the drugs to which one of the strains is resistant; or (ii) develop sensitive TB if untreated or if treated with a regimen set as to overcome the specific resistance pattern.

Initially we assume that when an individual is infected with both resistant and sensitive strains there will be a preferential activation (and transmission) of resistant strains – scenario (i) above. This corresponds to a worse case scenario where the treatment regimen available is not totally effective and selects for resistance. Later, in Section 3.5, we show that the results essentially extend to a more general implementation of mixed infection - scenario (ii) above.

The two-strain model can be represented as the system of differential equations (3.1).

$$\begin{cases} S' &= b-(\beta_s I_s+\beta_r I_r+\mu)S \\ L_s' &= (1-\phi)\beta_s I_s S-(\omega+\phi\sigma\beta_s I_s+\sigma\beta_r I_r+\mu)L_s+(1-\gamma)\tau_s I_s \\ L_r' &= (1-\phi)\beta_r I_r S+(1-\phi)\sigma\beta_r I_r L_s-(\omega+\phi\sigma\beta_s I_s+\phi\sigma\beta_r I_r+\mu)L_r+\tau_r I_r \\ I_s' &= \phi\beta_s I_s S+(\omega+\phi\sigma\beta_s I_s)L_s-(\tau_s+\mu+\delta)I_s \\ I_r' &= \phi\beta_r I_r S+\phi\sigma\beta_r I_r L_s+(\omega+\phi\sigma\beta_s I_s+\phi\sigma\beta_r I_r)L_r+\gamma\tau_s I_s-(\tau_r+\mu+\delta)I_r \end{cases} \tag{3.1}$$

Parameter values are given and described in Table 3.1. Parameters that refer to sensitive

Table 3.1: Two-strain model parameters

symbol	definition	value
β_s, β_r	transmission coefficient	variable
μ	death rate and birth rate	1/70 yr^{-1}
δ	death rate associated to TB	0.2 yr^{-1}
ϕ	proportion of individuals that develop active TB	0.1
	(the remaining $1-\phi$ have latent sensitive TB)	
σ	factor reducing the risk of infection as a result of acquired	0.25
	immunity to a previous infection with sensitive or resistant strains	
ω	rate of endogenous reactivation of latent TB	0.0002 yr^{-1}
τ_s, τ_r	rate of recovery under treatment of active sensitive and resistant TB	2, 1.5 yr^{-1}
γ	proportion of sensitive TB treatment failure acquiring resistance	0.003 (or $\gamma = 0$)

TB take values as in Gomes *et al.* (2004a). Reactivation rate is considered the same for sensitive and resistant infections. Individuals reactivate at a low rate so that a majority never progress to active disease (Gomes *et al.*, 2004a; Vynnycky and Fine, 1997). Different assumptions can be found in the literature that discriminate related mechanisms such as relapse of self-cured individuals or of treated patients, chronic infections and successive treatment failures (Blower & Chou (2004); Dye *et al.* (1998); Castillo-Chavez & Feng (1997); Dye & Williams (2000), respectively). We assume the rate of mortality associated to TB as in Dye & Espinal (2001). Birth rate b compensates for disease-induced and background mortality to keep the population size constant over time, so $b = \mu + \delta(I_s + I_r)$. The proportion acquiring resistance, γ, is on the lower bound of ranges considered in Cohen & Murray (2004) and Dye & Espinal (2001). We assume that the period of infectiousness of a resistant TB case is, on average, two months longer than that of a sensitive case. There is evidence that an individual infected with a resistant strain stays longer in the infectious state due to either improper regimen, late identification of the resistance phenotype, or lower efficacy of treatment (Espinal *et al.*, 2000). The factor reducing the risk of infection as a result of acquiring immunity, σ, is the same for both resistant and sensitive strains. Differences in transmission rates are explored by continuously varying the strain-specific transmission coefficients β_s and β_r.

3.3 Equilibria and stability

For system (3.1) the simplex

$$\mathbb{S} := \{(S, L_s, L_r, I_s, I_r) \in (\mathbb{R}_0^+)^5 : \ S + L_s + L_r + I_s + I_r = 1\}$$

is a positively invariant set, and thus we restrict the study of the solutions of the system to $\mathbb{S}$. By the fundamental theory of ODE's, we know that (3.1) defines a dynamical system on $\mathbb{S}$ as uniqueness, global existence and continuous dependence of solutions on initial data is guaranteed when initial values are in $\mathbb{S}$.

3.3.1 Basic reproduction number, R_0

We calculate the basic reproduction number, R_0, using the next generation approach, developed in van den Driessche & Watmough (2002). In order to compute the basic reproduction number it is important to distinguish new infections from all other class transitions in population. The infected classes are L_s, L_r, I_s and I_r, so we can write system (3.1) as

$$\dot{X} = f(X) \Leftrightarrow \dot{X} = \mathcal{F}(X) - \mathcal{V}(X), \tag{3.2}$$

where $X = (L_s, L_r, I_s, I_r, S)$, $\mathcal{F}$ is the rate of appearance of new infections in each class. Hence,

$$\mathcal{F} = (\ (1-\phi)\beta_s I_s S,\ (1-\phi)\beta_r I_r S,\ \phi\beta_s I_s S,\ \phi\beta_r I_r S, 0)^T,$$

and the disease-free equilibrium is $X_0 = (0, 0, 0, 0, 1)$.

Derivatives of $\mathcal{F}$ and $\mathcal{V}$ with respect to the infected classes at X_0 are

$$F = \begin{bmatrix} 0 & 0 & (1-\phi)\beta_s & 0 \\ 0 & 0 & 0 & (1-\phi)\beta_r \\ 0 & 0 & \phi\beta_s & 0 \\ 0 & 0 & 0 & \phi\beta_r \end{bmatrix}, \quad V = \begin{bmatrix} \mu+\omega & 0 & -(1-\gamma)\tau_s & 0 \\ 0 & \mu+\omega & 0 & \tau_r \\ -\omega & 0 & \mu+\delta+\tau_s & 0 \\ 0 & -\omega & \gamma\tau_s & \mu+\delta+\tau_r \end{bmatrix}.$$

The basic reproduction number is defined, following van den Driessche & Watmough (2002), as the spectral radius of the next generation matrix, FV^{-1}:

$$R_0 = \max\{R_{0s}, R_{0r}\}, \tag{3.3}$$

where R_{0s} and R_{0r} are the two eigenvalues:

$$\begin{aligned} R_{0s} &= \frac{\beta_s(\omega + \phi\mu)}{(\mu+\omega)(\mu+\delta+\tau_s) - (1-\gamma)\tau_s\omega} \\ R_{0r} &= \frac{\beta_r(\omega + \phi\mu)}{(\mu+\omega)(\mu+\delta+\tau_r) - \omega\tau_r}. \end{aligned} \tag{3.4}$$

We can also interpret R_{0s} and R_{0r} as the average number of secondary infectious cases that an infectious individual (with a sensitive or a resistant strain, respectively) would generate in a totally susceptible host population. A threshold condition for endemicity is given by $R_0 = 1$: the disease dies out if $R_0 < 1$, and becomes endemic if $R_0 > 1$.

3.3.2 Steady states

System (3.1) has one disease-free equilibrium, $E_0 = (1, 0, 0, 0, 0)$ and two endemic equilibria of the form: $E_r = (S^r, 0, L_r^r, 0, I_r^r)$ and $E_{rs} = (S^*, L_s^*, L_r^*, I_s^*, I_r^*)$, corresponding respectively to states where only resistant strains, or both types of strains are present.

The bifurcation diagram in Figure 3.4(a) divides the $(R_{0s}, R_{0r})-$space into three regions as characterised by the long-term epidemiological outcomes, each corresponding to a stable steady state of the system: disease eradication (**I**), persistence of only drug-resistant TB (**II**) or coexistence i.e. persistence of both drug-sensitive and drug-resistant TB (**III**).

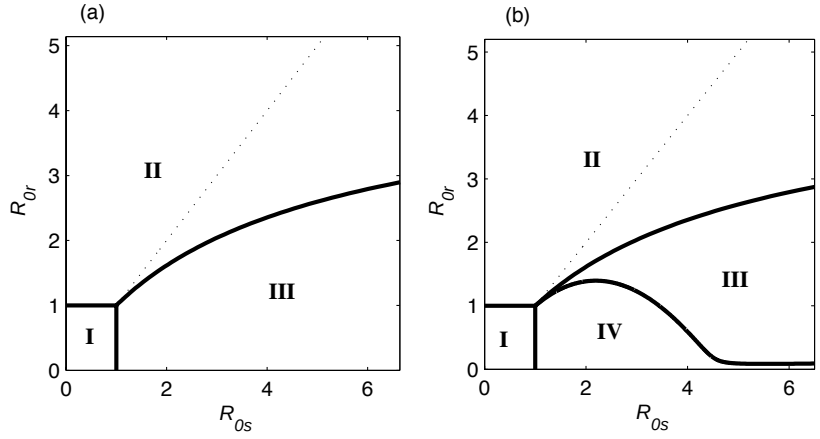


Figure 3.4: Long term epidemiological outcome: **(a)** $\gamma > 0$; **(b)** $\gamma = 0$. **I** - Disease eradication; **II**- Persistence drug-resistant TB only; **III** - Coexistence. **IV** - Persistence drug-sensitive TB only. The dotted line corresponds to the model without reinfection $\sigma = 0$.

Note that, infectious cases with sensitive strains give rise to new cases of resistant strains at a constant rate $\gamma > 0$, due the acquisition of resistance through treatment failure. It is therefore not possible to have an equilibrium where only sensitive strains are present. However, this equilibrium exists in the limit $\gamma = 0$, which corresponds to no acquired resistance.

The resulting equilibrium has the form $E_s = (S^s, L_s^s, 0, I_s^s, 0)$ and in Figure 3.4(b) we can see the corresponding stability region (marked as **IV**). We explore this limit case in more detail in Section 3.3.5, but otherwise we consider $\gamma > 0$.

3.3.3 Stability of the disease-free equilibrium

The stability properties of the disease-free equilibrium (trivial equilibrium) E_0, corresponding to the threshold condition for endemicity are given by the theorem 3.3.1, stated below.

Theorem 3.3.1. *The disease-free equilibrium E_0 of system (3.1) is locally asymptotically stable, if $R_0 < 1$, i.e. if $R_{0s} < 1$ and $R_{0r} < 1$, and it is unstable for $R_0 > 1$.*

Remark 3.3.1. *Numerical results suggest that the disease-free equilibrium is in fact globally asymptotically stable for $R_0 < 1$.*

Remark 3.3.2. *Numeric calculations and some analytical manipulations were obtained using* MATLAB 6.5®. *Equilibrium curves were computed with* MATCONT *continuation package of* MATLAB 6.5® *(Dhooge et al., 2003).*

Proof. By theorem 2 in van den Driessche & Watmough (2002) it is sufficient to prove conditions (A1)-(A5), with $\mathcal{F}$, $\mathcal{V}(X) = (\mathcal{V}^-(X) - \mathcal{V}^+(X))$ and X_0 as defined before and where $\mathcal{V}^+$ is the rate of transfer into each class by all other means and $\mathcal{V}^-$ is the rate of transfer out of each class. The verification of (A1)-(A4) is straightforward.

The Jacobian of f at X_0 with $\mathcal{F}$ set to zero, as

$$Df_{(\mathcal{F}=0)}(X_0) = \begin{bmatrix} -(\omega+\mu) & 0 & (1-\gamma)\tau_s & 0 & 0 \\ 0 & -(\omega+\mu) & 0 & \tau_r & 0 \\ \omega & 0 & -(\mu+\delta+\tau_s) & 0 & 0 \\ 0 & \omega & \gamma\tau_s & -(\mu+\delta+\tau_r) & 0 \\ 0 & 0 & \delta-\beta_s & \delta-\beta_r & -\mu \end{bmatrix}.$$

The eigenvalues are: $-\mu$ and the solutions of equation

$$p_1(\lambda)p_2(\lambda) = 0$$

where $p_1(\lambda) = \lambda^2 - a_1\lambda + a_0$ and $p_2(\lambda) = \lambda^2 - b_1\lambda + b_0$ and

$$\begin{aligned} -a_1 &= 2\mu + \delta + \tau_r + \omega, \\ a_0 &= \mu(\mu + \delta + \tau_r) + \omega(\mu + \delta), \\ -b_1 &= 2\mu + \delta + \tau_s + \omega, \\ b_0 &= \mu(\mu + \delta + \tau_s) + \omega(\mu + \delta + \gamma\tau_s). \end{aligned}$$

Since $-a_1, a_0$ and $-b_1, b_0$ are positive, all eigenvalues have negative real part and the result follows.

□

3.3.4 Stability of boundary and coexistence equilibria

The existence of an equilibrium for which only resistant strains persist is given by theorem 3.3.2, stated below.

Theorem 3.3.2. *System (3.1) has exactly one non-trivial boundary equilibrium, $E_r = (S^r, 0, L_r^r, 0, I_r^r)$, for $R_{0r} > 1$.*

Proof. From the first, second and third equations of system (3.1) at equilibrium, we get a relation between S, L_s, L_r and I_s, I_r:

$$\begin{aligned}
S &= \frac{\mu + \delta I_s + \delta I_r}{\mu + \beta_s I_s + \beta_r I_r} = F(I_s, I_r),\\
L_s &= I_s \frac{(1-\phi)\beta_s S + (1-\gamma)\tau_s}{\mu + \omega + \phi\sigma\beta_s I_s + \sigma\beta_r I_r} = I_s \frac{(1-\phi)\beta_s F(I_s, I_r) + (1-\gamma)\tau_s}{\mu + \omega + \phi\sigma\beta_s I_s + \sigma\beta_r I_r} =\\
&= G(I_s, I_r) I_s,\\
L_r &= I_r \frac{(1-\phi)\beta_r (S + \sigma L_s) + \tau_r}{\mu + \omega + \phi\sigma(\beta_s I_s + \beta_r I_r)} = I_r \frac{(1-\phi)\beta_r (F(I_s, I_r) + \sigma G(I_s, I_r) I_s) + \tau_r}{\mu + \omega + \phi\sigma(\beta_s I_s + \beta_r I_r)} =\\
&= H(I_s, I_r) I_r.
\end{aligned}$$

Suppose that $I_s = 0$ (and subsequently $L_s = 0$). If I_r is nonzero, from the fifth equation of the system (3.1) we get

$$\phi\beta_r F(0, I_r) + (\omega + \phi\sigma\beta_r I_r) H(0, I_r) - (\mu + \delta + \tau_r) = 0. \tag{3.5}$$

We can write this as follows

$$\frac{P(I_r)}{Q(I_r)} = 0,$$

where P and Q are polynomials of second degree such that:

$$\begin{aligned}
Q(I_r) &= (\mu + \beta_r I_r)(\mu + \omega + \phi\sigma\beta_r I_r) > 0,\\
P(I_r) &= \mu(p_2(\beta_r) I_r^2 + p_1(\beta_r) I_r + p_0(\beta_r)),
\end{aligned}$$

where

$$\begin{aligned}
p_2(\beta_r) &= -\phi\sigma\beta_r^2 < 0,\\
p_1(\beta_r) &= \phi\sigma\beta_r^2 - (\tau_r + \omega + \mu + (1-\phi)\delta + \phi\sigma(\mu + \delta))\beta_r,\\
p_0(\beta_r) &= \beta_r(\omega + \phi\mu) - (\mu(\mu + \tau_r + \delta) + \omega(\mu + \delta)).
\end{aligned}$$

If $\beta_r > \dfrac{\mu(\mu + \delta + \tau_r) + \omega(\mu + \delta)}{\phi\mu + \omega} \Leftrightarrow R_{0r} > 1$, then $p_0(\beta_r) > 0$ and we have exactly one positive solution of $P(I_r)$. If $\beta_r \leq \dfrac{\mu(\mu + \delta + \tau_r) + \omega(\mu + \delta)}{\phi\mu + \omega} \Leftrightarrow R_{0r} \leq 1$, then $p_0(\beta_r) \leq 0$ but also $p_1(\beta_r) \leq 0$, since $0 < \phi, \sigma < 1$. So there are no positive solutions of $P(I_r)$.

□

In order to derive an expression for the region of stability of the boundary equilibrium we measure the capacity of sensitive TB strains to invade and persist in a population where resistant TB is at equilibrium. In this context, $E_r = (S^r, 0, L_r^r, 0, I_r^r)$ corresponds to an equilibrium free of sensitive TB. Applying the methods in van den Driessche & Watmough (2002) once again we find the basic reproduction number of the sensitive strains in a population where resistant strains are fixed. Hence, consider the case when only the only sensitive TB is transmissible, in a population where resistant TB is at equilibrium. The infected compartments are L_s and I_s. Following van den Driessche & Watmough (2002), we write system (3.1) as in (6.2) where $X = (L_s, I_s, S, L_r, I_r)$ and $\mathcal{F} = (\ (1-\phi)\beta_s I_s S,\ \phi\beta_s I_s S, 0, 0, 0)^T$. The disease (sensitive-TB)-free equilibrium is $(0, 0, S^r, L_r^r, I_r^r)$.

We can compute F and V that correspond to the derivatives at X_0 with respect to the infected classes of $\mathcal{F}$ and $\mathcal{V}$, respectively:

$$F = \begin{bmatrix} 0 & (1-\phi)\beta_s \\ 0 & \beta_s \end{bmatrix}, \quad V = \begin{bmatrix} \mu+\omega+\sigma\beta_r I_r^r & -(1-\gamma)\tau_s \\ -\omega & \mu+\delta+\tau_s \end{bmatrix}.$$

The basic reproduction number of the sensitive strains in a population where resistant strains are fixed is then the spectral radius of the next generation matrix, FV^{-1}:

$$R_{0s}(E_r) = \frac{S^r\beta_s(\phi(\mu+\sigma\beta_r I_r^r)+\omega)}{(\mu+\sigma\beta_r I_r^r+)(\mu+\delta+\tau_s)-(1-\gamma)\tau_s\omega}. \tag{3.6}$$

Remark 3.3.3. *Note that this is still valid for $R_{0r} < 1$. In this case the disease-free equilibrium is $E_0 = (1,0,0,0,0)$ and we restore the endemicity threshold.*

This formalism permits the derivation of a threshold condition for coexistence, now equivalent to a threshold condition for sensitive TB endemicity in a population where resistant strains are at equilibrium, $R_{0s}(E_r) = 1$: only resistant TB persists for $R_{0s}(E_r) < 1$, while for $R_{0s}(E_r) > 1$ sensitive strains can invade a population where resistant strains are fixed, that is to say coexistence is possible.

Theorem 3.3.3 below expresses this result in terms of stability for the equilibrium E_r.

Theorem 3.3.3. *If $R_{0r} > 1$ the equilibrium E_r of system (3.1) is stable for $R_{0s}(E_r) < 1$ and unstable for $R_{0s}(E_r) > 1$.*

Proof. By theorem 2 in van den Driessche & Watmough (2002) it is sufficient to prove conditions (A1)-(A5). Once more, conditions (A1)-(A4) are of trivial verification. To prove the remaining condition (A5) we write the Jacobian of f at X_0, with $\mathcal{F}$ set to zero, ordering coordinates as (S, L_r, I_r, L_s, I_s). Then, the Jacobian has the form

$$Df_{(\mathcal{F}=0)}(S^r, L_r^r, I_r^r, 0, 0) = \begin{bmatrix} G_1 & G_2 \\ 0 & G_4 \end{bmatrix}.$$

where

$$G_1 = \begin{bmatrix} -(\mu+\beta_r I_r^r) & 0 & \delta-\beta_r S^r \\ (1-\phi)\beta_r I_r^r & -(\mu+\omega+\phi\sigma\beta_r I_r^r) & (1-\phi)\beta_r S^r - \phi\sigma\beta_r L_r^r + \tau_r \\ \phi\beta_r I_r^r & \omega+\phi\sigma\beta_r I_r^r & \phi\beta_r(S^r+\sigma L_r^r)-(\mu+\delta+\tau_r) \end{bmatrix}$$

and

$$G_4 = \begin{bmatrix} -(\mu+\omega+\sigma\beta_r I_r^r) & (1-\gamma)\tau_s \\ \omega & -(\tau_s+\mu+\delta) \end{bmatrix}.$$

Therefore, the eigenvalues of the Jacobian are given by the eigenvalues of G_1 and G_4.

For G_1 the eigenvalues are $-\mu$ and the roots of the polynomial

$$p_1(\lambda) = (\lambda^2 - a_1\lambda + a_0)$$

where

$$\begin{aligned} -a_1 &= -(\phi S^r + \phi\sigma L_r^r)\beta_r + (1+\phi\sigma)I_r^r\beta_r + (2\mu+\delta+\tau_r+\omega), \\ a_0 &= \phi\sigma\beta_r^2 I_r^{r2} + \\ &\quad +[-\phi\sigma\beta_r^2(S^r+L_r^r) + \beta_r(\tau_r+\omega+\mu+(1-\phi)\delta+\phi\sigma(\mu+\delta))]I_r^r + \\ &\quad +\mu(\mu+\delta+\tau_r)+\omega(\mu+\delta)-\beta_r((\omega+\phi\mu)S^r-\phi\sigma\mu L_r^r). \end{aligned}$$

From equation five of the system (3.1) at the equilibrium E_r we get:

$$(\phi S^r + \phi\sigma L_r^r)\beta_r I_r^r = (\mu + \delta + \tau_r) - \omega L_r^r$$

so $-a_1 I_r^r = \omega L_r^r + (\mu + \omega) I_r^r + (1 + \phi\sigma) {I_r^r}^2 \beta_r > 0$. Since $I_r^r > 0$, $-a_1 > 0$. From the proof of result 3.3.2 we know that I_r^r is the only positive solution of $P(I_r) = \mu(p_2(\beta_r) I_r^2 + p_1(\beta_r) I_r + p_0(\beta_r))$. We can write a_0 as

$$\begin{aligned} a_0 &= -p_2(\beta_r) {I_r^r}^2 - p1(\beta_r) I_r^r + \phi\sigma\beta_r^2 I_r^r - \phi\sigma\beta_r^2 (S^r + L_r^r) I_r^r \\ &\quad -p_0(\beta_r) + \beta_r(\omega + \phi\mu) - \beta_r((\omega + \phi\mu) S^r - \phi\sigma\mu L_r^r), \end{aligned}$$

Now using the fact that $1 = S^r + L_r^r + I_r^r$ we get

$$\begin{aligned} a_0 &= \phi\sigma\beta_r^2 I_r^r (1 - S^r - L_r^r) + \beta_r(\omega + \phi\mu)(1 - S^r - L_r^r) + \beta_r(\omega + \phi\mu) L_r^r - \beta_r\phi\sigma\mu L_r^r \\ &= \phi\sigma\beta_r^2 {I_r^r}^2 + \beta_r(\omega + \phi\mu) I_r^r + \beta_r(\omega + \phi\mu(1 - \sigma)) L_r^r > 0 \end{aligned}$$

Since $-a_1$ and a_0 are positive for all possible values of $\beta_r > \dfrac{\mu(\mu + \delta + \tau_r) + \omega(\mu + \delta)}{\phi\mu + \omega}$ all eigenvalues of $G1$ have negative real part.

For G_4 the characteristic polynomial is

$$p_2(\lambda) = \lambda^2 - b_1\lambda + b_0$$

where

$$\begin{aligned} b_0 &= (\mu + \sigma\beta_r I_r^r)(\mu + \delta + \tau_s) + \omega(\mu + \delta + \gamma\tau_s), \\ -b1 &= 2\mu + \delta + \tau_s + \omega + \sigma\beta_r I_r^r. \end{aligned}$$

Since $b_0 > 0$ and $-b_1 > 0$ are both positive we conclude that all eigenvalues of G_4 have negative real part.

□

Remark 3.3.4. *From the proof of this result we conclude that stability of E_r is equivalent to stability of the endemic equilibrium of the sub-system with only resistant strains and simultaneously stability of the sensitive TB-free equilibrium.*

Remark 3.3.5. *The curve that defines the coexistence region is given by the following relation (see Figure 3.4):*

$$R_{0s}(E_r) = 1 \Longleftrightarrow \beta_s = f(\beta_r) = \frac{(\mu + \sigma\beta_r I_r^r)(\mu + \delta + \tau_s) + \omega(\mu + \delta + \gamma\tau_s)}{S^r(\phi(\mu + \sigma\beta_r I_r^r) + \omega)}. \tag{3.7}$$

Remark 3.3.6. *Numerical results support that below the curve defined by f in the (R_{0s}, R_{0r})-space both types of strains will persist.*

Relation (3.7) reveals that persistence of sensitive strains depends on the reinfection process. The expression of $R_{0s}(E_r)$ is similar to that for R_{0s} in (3.4) with an additional term, $\sigma\beta_r I_r^r$. This term corresponds to reinfection by resistant strains of latent individuals infected with sensitive TB. Contrasting with the case where reinfection is not considered, $\sigma = 0$ (dotted line in Figure 3.4), reveals that persistence of only resistant strains is now possible even when these have lower transmissibility $R_{0r} < R_{0s}$. Coexistence is no longer governed solely by the invasion capacities of each strain (R_{0s} and R_{0r}) but also by the ability of sensitive

strains to overcome the reinfection pressure exerted by resistant strains. In particular, our results can be compared to the analysis of Blower & Gerberding (1998) (see Figure 2 and table 1 within), which does not consider reinfection. The model developed by these authors has the same possible outcomes (**I**,**II**,**III**) but these are fully determined by a linear relation between pathogen fitness as measured by the respective R_0: disease eradication (**I**) if $R_{0s} < 1$ and $R_{0r} < 1$; persistence of only resistant tuberculosis (**II**) if $R_{0r} > 1$ and $R_{0r} > R_{0s}$; of both drug sensitive and drug-resistant tuberculosis (**III**) if $R_{0s} > 1$ and $R_{0s} > R_{0r}$.

3.3.5 Limit case: $\gamma = 0$

The limit case $\gamma = 0$ is equivalent to assuming that there is no acquisition of drug resistance through treatment failure. Analysis of this limit case reveals regions where the elimination of drug-resistant strains may result from prevention of acquired resistance alone.

For $\gamma = 0$, the system has three non-trivial equilibria corresponding to the presence of each type of strains alone and coexistence (Figure 3.4(b)). The existence of the first two is given by theorem 3.3.4 stated below.

Theorem 3.3.4. *For $\gamma = 0$, system (3.1) has exactly two non-trivial boundary equilibria: $E_r = (S^r, 0, L_r^r, 0, I_r^r)$ for $R_{0r} > 1$ and $E_s = (S^s, L_s^s, 0, I_s^s, 0)$ for $R_{0s} > 1$.*

Proof. To show the existence of E_r we just have to repeat the calculations in proof of result 3.3.2 with $\gamma = 0$.

Suppose now that $I_r = 0$ (and subsequently $L_r = 0$). If I_r is nonzero, from the fourth equation of the system (3.1) we get

$$\phi\beta_s F(I_s, 0) + (\omega + \phi\sigma\beta_s I_s)G(I_s, 0) - (\mu + \delta + \tau_s) = 0 \tag{3.8}$$

where F and G are the same functions as in proof of result 3.3.2. Note that $F(I_s, 0), G(I_s, 0)$ have the same expression as $F(0, I_r), H(0, I_r)$ respectively if we just change the subscripts s, r. Moreover, equation (3.8) will be the same as equation (3.5) if we just change the subscripts s, r. Therefore we conclude that for $R_{0s} > 1$ we have exactly one positive solution of $P(I_s)$, that corresponds to E_s.

□

Two coexistence thresholds must be calculated: the first separates the region where only sensitive TB persists from the region of coexistence; the second marks the shift from coexistence to persistence of resistant TB alone.

Regarding the second threshold, it can be verified that the threshold condition is the same as when $\gamma > 0$, i.e., $R_{0s}(E_r) = 1$. Moreover, the stability results pertaining the equilibrium E_{sr} (Theorem 3.3.3) can be extended to the case $\gamma = 0$.

To compute the first threshold we use the same reasoning as before. We consider resistant TB as the phenotype invading a population where sensitive TB is already endemic. Then, $E_s = (S^s, L_s^s, 0, I_s^s, 0)$ corresponds to the equilibrium free of resistant TB. Hence, let us assume that only resistant TB is considered disease,then the infected compartments are L_r and I_r and following (van den Driessche & Watmough, 2002), we can write system (3.1) as in 6.2 with $X = (L_r, I_r, S, L_s, I_s)$ and $\mathcal{F} = (\ (1-\phi)\beta_s I_s S,\ \phi\beta_s I_s S, 0, 0, 0)^T$.

The disease (resistant-TB)-free equilibrium is now $X_0 = (0, 0, S^r, L_r^r, I_r^r)$. Let us compute F and V corresponding to the derivatives at X_0, with respect to the infected classes, of $\mathcal{F}$ and $\mathcal{V}$, respectively:

$$F = \begin{bmatrix} 0 & (1-\phi)\beta_r(S^s + \sigma L_s^s) \\ 0 & \phi\beta_r(S^s + \sigma L_s^s) \end{bmatrix}, \quad V = \begin{bmatrix} \mu + \omega + \phi\sigma\beta_s I_s^s & -\tau_r \\ -(\omega + \phi\sigma\beta_s I_s^s) & \mu + \delta + \tau_r \end{bmatrix}.$$

The basic reproduction number of the resistant strains, in a population where the sensitive strains are fixed, is the spectral radius of the next generation matrix, FV^{-1}:

$$R_{0r}(E_s) = \frac{(S^s + \sigma L_s^s)\beta_r(\phi\mu + \omega + \phi\sigma\beta_s I_s^s)}{(\mu + \omega + \phi\sigma\beta_s I_s^s)(\mu + \delta + \tau_r) - (\omega + \phi\sigma\beta_s I_s^s)\tau_r}. \tag{3.9}$$

Resistant strains can invade a population where sensitive strains are fixed when $R_{0r}(E_s) > 1$.

The corresponding result for the stability of the boundary equilibrium is expressed by theorem 3.3.5 stated below.

Theorem 3.3.5. *Consider system (3.1) with $\gamma = 0$. When $R_{0r} > 1$, the equilibrium E_r is stable if $R_{0s}(E_r) < 1$ and unstable if $R_{0s}(E_r) > 1$. When $R_{0s} > 1$, the equilibrium E_s is stable for $R_{0r}(E_s) < 1$ and unstable for $R_{0r}(E_s) > 1$.*

Proof. In what matters the stability of E_r we can repeat the calculations in the proof of result 3.3.3 with $\gamma = 0$.

For the case of equilibrium $E_s = (S^s, L_s^s, 0, I_s^s, 0)$ by the theorem 2 in (van den Driessche & Watmough, 2002) is sufficient to prove conditions (A1)-(A5) for the system as we described above. It is straightforward to check (A1)-(A4).

Let us prove the condition (A5). For simplicity of calculations let us write the Jacobian of f, with $\mathcal{F}$ set to zero, at X_0 with the following order in the coordinates (S, L_s, I_s, L_r, I_r). Then the Jacobian can be written in the following way

$$Df_{(\mathcal{F}=0)}(S, L_s^s, I_s^s, 0, 0) = \begin{bmatrix} H_1 & H_2 \\ 0 & H_4 \end{bmatrix}.$$

where

$$H_1 = \begin{bmatrix} -(\beta_r I_s^s + \mu) & 0 & \delta - \beta_s S^s \\ (1-\phi)\beta_s I_s^s & -(\mu + \omega + \phi\sigma\beta_s I_s^s) & (1-\phi)\beta_s S^s - \phi\sigma\beta_s L_s^s + \tau_s \\ \phi\beta_s I_s^s & \omega + \phi\sigma\beta_s I_s^s & \phi\beta_s(S^s + \sigma L_s^s) - (\mu + \delta + \tau_s) \end{bmatrix}$$

and

$$H_4 = \begin{bmatrix} -(\mu + \omega + \phi\sigma\beta_s I_s^s) & \tau_r \\ \omega + \phi\sigma\beta_s I_s & -(\tau_r + \mu + \delta) \end{bmatrix}.$$

Therefore, the eigenvalues of the Jacobian are given by the eigenvalues of H_1 and H_4. Note that H_1 is similar to G_1 in the proof of *result* 3.3.3 if we just replace the subscript r by s. So we conclude that all eigenvalues of H_1 have negative real part. For H_4 the characteristic polynomial is

$$p_2(\lambda) = \lambda^2 - b_1\lambda + b_0$$

where

$$\begin{aligned} b_0 &= (\mu + \sigma\beta_s I_s^s)(\mu + \delta + \tau_r) + \omega(\mu + \delta), \\ -b1 &= (2\mu + \delta + \tau_r + \omega + \sigma\beta_s I_s^s). \end{aligned}$$

Since both $b_0 > 0$ and $-b_1 > 0$ all eigenvalues of H_4 have negative real parts.

□

Again we emphasize the dependence of the coexistence threshold on reinfection. Susceptible and latent individuals infected with sensitive strains are susceptible to (re)infection with resistant strains at rates $\beta_r I_r$ (infection) and $\sigma\beta_r I_r$ (superinfection) respectively. The result is the nonlinear curve in Figure 3.4(b).

3.4 Fitness impact on the coexistence region

Drug resistance among *Mtb* isolates is caused by point mutations in the bacterial genome that affect anti-mycobacterial drug activity. If a mutation that confers drug resistance can exert a cost to the parasite we may expect these strains to be less transmissible than the drug sensitive. To explore the epidemiological consequences of resistance cost we fix the relative transmission coefficient, $\alpha = \beta_r/\beta_s$, and explore the system behaviour by varying a parameter β such that

$$\beta_s := \beta \qquad \beta_r := \alpha\beta.$$

As such, $\alpha < 1$ means that the resistant strains have lower transmissibility than the sensitive. Despite being less likely, the possibility $\alpha > 1$ is also considered since this topic is still open to discussion (Cohen *et al.*, 2003; Gagneux *et al.*, 2006). Figure 3.5 shows the bifurcation diagrams obtained for two values of α. When $\alpha = 0.5$ (full line) low values of β_s lead to coexistence, but only resistant strains persist for high rates of transmission, where reinfection prevails. In this scenario it is possible to induce coexistence of sensitive and resistant strains by reducing the disease transmission rate. In turn, coexistence improves the chance of controlling drug-resistance prevalence. For $\alpha = 1.1$ (dashed line) β_s and β_r lie in regions **I** and **II** thus, only resistant strains may persist.

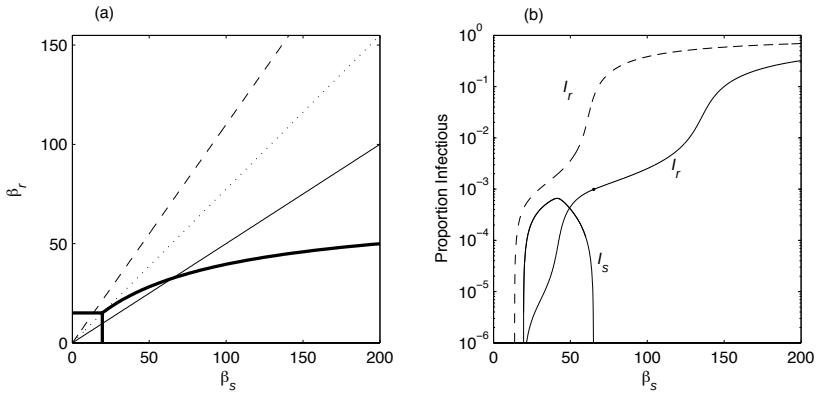


Figure 3.5: Decreased transmission: **(a)** Bifurcation diagram: Straight lines correspond to $\beta_r = \alpha\beta_s$ for different values of α: $\alpha = 1.1$ dashed line, $\alpha = 0.5$ full line, and $\alpha = \alpha_C$ dotted line. **(b)** Corresponding equilibrium curves: $\alpha = 1.1$ dashed line, $\alpha = 0.5$ full lines (only stable equilibria represented).

We derive a critical value for α below which a reduction in the overall transmission can open the possibility for coexistence:

$$\alpha_C = \frac{\mu(\mu + \delta + \tau_r) + \omega(\mu + \delta)}{\mu(\mu + \delta + \tau_s) + \omega(\mu + \delta + \gamma\tau_s)}. \tag{3.10}$$

Note that, for the choice of parameters as in Table 1, $\alpha_C \approx 0.7745 < 1$ (dotted line in Figure 3.5(a)). The critical value α_C will be later used to compare the impact of different control measures on the coexistence region.

In the case illustrated by $\alpha = 0.5$, as the transmission coefficient, β, increases, the system evolves from dominance of the sensitive strain to dominance of the resistant. This can be interpreted as follow. The minimal transmissibility above which resistant strains can be sustained in the population where sensitive strains are endemic, without the contribution of acquired resistance ($\gamma = 0$), is given by the condition $R_{0r}(E_s) = 1$. This marks a threshold in trasmission above which superinfection of sensitive by resistant strains occurs. Below the threshold, resistant strains are outcompeted by the sensitive due to the higher transmission coefficient of the latter (recall that $\alpha < 1$). In this regime, resistant cases can only be maintained due to acquired resistance ($\gamma > 0$). This superinfection threshold is marked in Figure 3.6 (a). Despite being a threshold imposed by reinfection, it is formally obtained as an invading threshold using an adaptation of the methods for the computation of the basic reproduction number (see deduction of equation (3.9)). This is possible since we were considering independent transmission parameters for each strain.

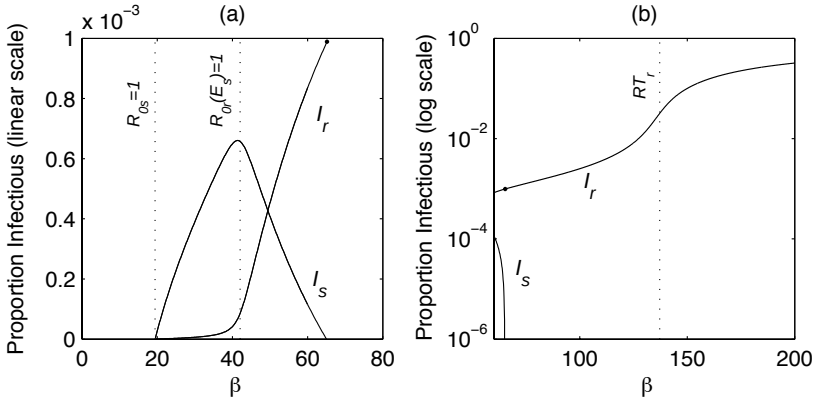


Figure 3.6: Transmission thresholds for $\alpha = 0.5$. **(a)** Equilibrium curves on the coexistence region. Vertical dotted lines mark the epidemic threshold of sensitive strains and the superinfection threshold of resistant strains. **(b)** Equilibrium curves on the region of persistence of drug-resistant TB only. Vertical dotted line marks the reinfection threshold for the resistant strains, RT_r.

Disease prevalence exhibits a new steep increase, for sufficiently high transmission rates, marked by a threshold in reinfection of the resistant strains, RT_r, marked in Figure 3.6 (b). Since sensitive strains are no longer circulating in the population this threshold is now a reinfection threshold as defined in Chapter 2 and it corresponds to a shift in dominance from primary infections to reinfections. Following the method defined in Section 2.5 we can construct the reinfection sub-model from system (3.1) with $L_s = I_s = 0$ and $\beta_r = \alpha\beta_s$ as

$$\begin{cases} L_r' &= \mu + \delta I_r + \tau_r I_r - \phi\sigma\beta_r I_r L_r - \mu L_r \\ I_r' &= \phi\sigma\beta_r I_r L_r - (\tau_r + \mu + \delta) I_r. \end{cases} \tag{3.11}$$

The system undergoes a bifurcation for the disease-free equilibrium $(1, 0)$ at

$$\beta_r = \frac{\tau_r + \mu + \delta}{\phi\sigma}, \tag{3.12}$$

when the determinant of the Jacobian matrix evaluated at the disease-free equilibrium is zero.

3.5 Model extensions - mixed infections

In the model presented in Section 3.2 we assumed that active TB resulting from a mixed infection would always express the resistant phenotype. Now we relax this assumption by also allowing individuals with a mixed infection to progress to sensitive TB (scenario (ii) in Section 3.2.3). Molecular studies suggest several possible outcomes for mixed infections (van Rie *et al.*, 2004): sensitive TB may develop in untreated individuals carrying mixed infections due to the faster replication of sensitive strains; sensitive strains may prevail when treatment matches drug regimen to the resistance pattern specific to each case; resistant strains may emerge when treating with first line anti-tuberculosis drugs. Moreover, fitness trade-offs may favour sensitive strains when competition takes place during the latent stage but, this will only have impact on transmission once individuals progress to the disease stage. Although the possible outcomes we describe here are intuitive and expected they are the product of different and complex mechanisms. These mechanisms are still, quantitatively and qualitatively, unclear from the molecular point of view.
We extend the two-strain model by introducing a mixed latent class, L_m, representing the proportion of individuals with a latent infection that combines both resistant and sensitive strains - mixed infection. When individuals with mixed infections progress to active TB, either by endogenous reactivation or exogenous reinfection, a fraction θ will manifest resistant TB entering I_{mr} while the remainder will develop sensitive TB progressing into I_{ms}. The model

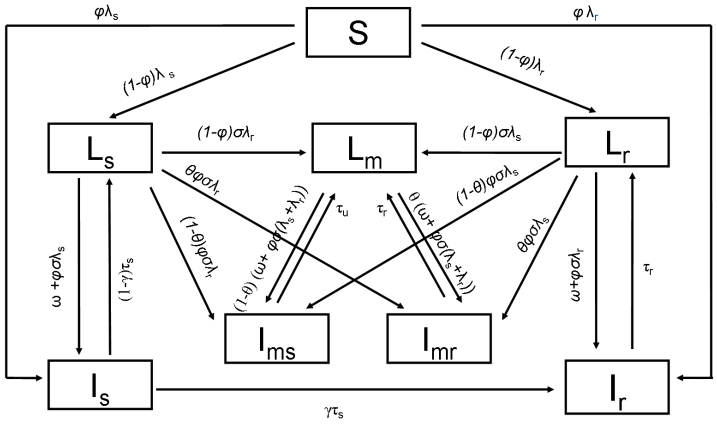


Figure 3.7: Mixed infections model. Individuals are classified according to infection state into susceptible (S), latently infected (L) and infectious (I). Parameters are the transmission coefficient (β), the death and birth rate (μ), the proportion of individuals developing active TB (ϕ), the reinfection factor (σ), the rate of reactivation (ω), the rate of recovery under treatment (τ), the proportion of resistance acquisition (γ) and proportion of mixed infection contributing to transmission of resistant strains (θ). Subscripts s, r and m relate to sensitive, resistant or mixed infection, respectively.

is represented diagrammatically by Figure 3.7 and corresponds to the system of equations

$$\begin{cases} S' &= b-(\lambda_s+\lambda_r+\mu)S \\ L'_s &= (1-\phi)\lambda_s S-(\omega+\phi\sigma\lambda_s+\sigma\lambda_r+\mu)L_s+(1-\gamma)\tau_s I_s \\ I'_s &= \phi\lambda_s S+(\omega+\phi\sigma\lambda_s)L_s-(\tau_s+\mu+\delta)I_s \\ L'_m &= (1-\phi)\sigma\lambda_r L_s-(\omega+\phi\sigma(\lambda_s+\lambda_r)+\mu)L_m+\tau_s I_{ms}+\tau_r I_{mr}+(1-\phi)\sigma\lambda_s L_r \\ I'_{ms} &= (1-\theta)\phi\sigma\lambda_r L_s+(1-\theta)(\omega+\phi\sigma(\lambda_s+\lambda_r))L_m-(\tau_s+\mu+\delta)I_{ms}+(1-\theta)\phi\sigma\lambda_s L_r \\ I'_{mr} &= \theta\phi\sigma\lambda_r L_s+\theta(\omega+\phi\sigma(\lambda_s+\lambda_r))L_m-(\tau_r+\mu+\delta)I_{mr}+\theta\phi\sigma\lambda_s L_r \\ L'_r &= (1-\phi)\lambda_r S-(\omega+\sigma\lambda_s+\phi\sigma\lambda_r+\mu)L_r+\tau_r I_r \\ I'_r &= \phi\lambda_r S+\gamma\tau_s I_s+(\omega+\phi\sigma\lambda_r)L_r-(\tau_r+\mu+\delta)I_r \end{cases} \tag{3.13}$$

where $\lambda_s = \beta_s(I_s + I_{ms})$ and $\lambda_r = \beta_r(I_r + I_{mr})$ represent the force of infection of the two types of TB. The parameters are the same as before with exception of θ and the birth rate, b, that we consider in such way that the population size is constant over time, so $b = \mu + \delta(I_s + I_{ms} + I_r + I_{mr})$. Parameter θ summarizes all mechanisms that determine the prevailing strain in a mixed infection. It can be varied to explore different scenarios, depending on the relative contribution of each mechanism to the overall situation. Note that with $\theta = 1$ we recover the two-strain model presented in Section 3.2.

Figure 3.8 shows the long-term behavior of the mixed infection model when we change parameter θ. Notably, the coexistence region increases as the percentage of mixed infections that progress to sensitive active-TB increases. The limit case ($\theta = 1$) is, in fact, the worst case scenario. Moreover, coexistence again depends on the transmission coefficients of both types of strains in a nonlinear manner. A more subtle result is that coexistence is possible

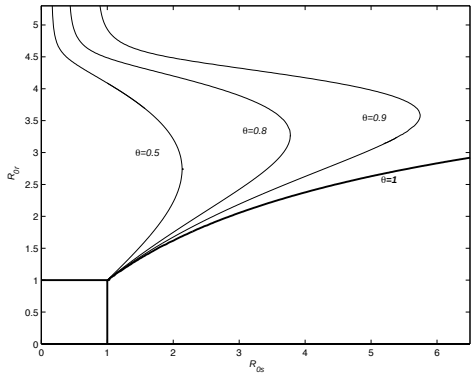


Figure 3.8: Long-term epidemiological outcome: Bifurcation diagram on R_{0s} and R_{0r}. Curves separate coexistence region from persistence of only resistant strains for different values of parameter θ. For $\theta = 1$ we have the same curve as in Figure 3.4.

for high transmission levels of drug-resistant strains even when sensitive strains have low transmissibility. This is related to the assumption that individuals never succeed in fully clearing tuberculosis bacteria and therefore, mixed infections are very frequent when either or both strains are highly transmissible. Under the current assumption, a fraction $\theta(< 1)$ of these infections will progress to resistant TB and the remaining will progress to sensitive TB,

thus forcing coexistence. In contrast, all mixed infections will develop into resistant TB when $\theta = 1$.

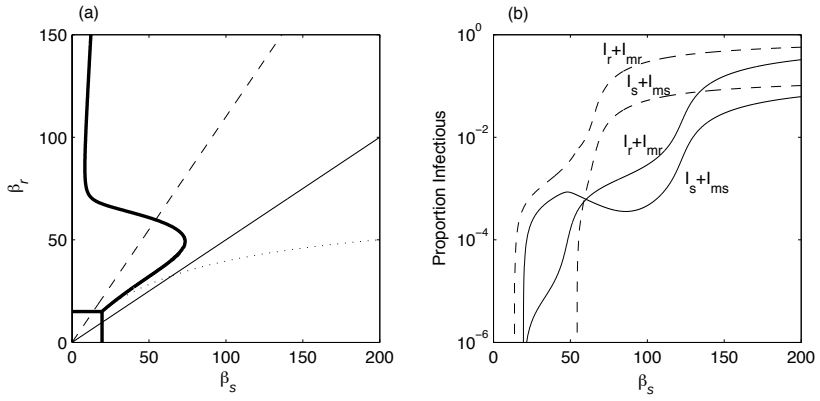


Figure 3.9: Mixed infections case $\theta = 0.8$: **(a)** Bifurcation diagram: Straight lines correspond to $\beta_r = \alpha\beta_s$ for different values of α: $\alpha = 1.1$ dashed line, $\alpha = 0.5$ full line. Dotted line corresponds to $\theta = 1$. **(b)** Corresponding equilibrium curves: $\alpha = 1.1$ dashed line, $\alpha = 0.5$ full lines (only stable equilibria represented).

Let us again explore what happens when the transmission rate of resistant and sensitive strains have a linear association: $\beta_r = \alpha\beta_s$. Parameter α thus expresses the impact of resistance on pathogen fitness. In Figure 3.9(a) straight lines exemplify two contrasting cases: drug-resistance has an associated cost ($\alpha = 0.5$, full lines) or resistant strains have a higher transmission rate ($\alpha = 1.1$, dashed lines). When $\alpha = 0.5$, resistant and sensitive strains coexist for all possible values of β_s. If transmission (β_s) increases, resistant strains start to dominate. But inversely to the case $\theta = 1$ (two-strain model) this does not drive sensitive strains to extinction because some mixed infections develop sensitive cases (compare Figure 3.9(b) with Figure 3.5(b), full lines). Above a certain transmission level, mixed infections represent almost the totality of TB infections, and the proportion of resistant TB in the total TB burden is then driven by θ.

3.6 Control strategies

The World Health Organization has two major control programs for TB: DOTS, Directly Observed Treatment Short-course, consisting of standardized short-course treatment of TB cases given under direct observation to ensure treatment adequacy and compliance; and DOTS-Plus, an extension of DOTS specifically designed for controlling multi-drug resistant TB. DOTS-Plus uses more effective, but also more expensive and toxic drugs. It is not always clear what should be the strategy of choice to manage resistant TB in a given setting (Dye *et al.*, 2002; Pablo-Mendez *et al.*, 2002): is DOTS enough or should it be extended to DOTS-Plus?

Knowing that reinfection can have strong consequences on the effectiveness of interventions (Gomes *et al.*, 2004a) we explore how our model behaves under these two strategies. These control measures are designed to fight different processes: DOTS prevents the acquisition of resistance due to treatment failure by ensuring compliance; whereas DOTS-Plus reduces transmission of resistant strains by adapting the treatment regimen to better suit resistant

cases. Therefore, we model DOTS by reducing the proportion of failed treatments that leads to acquired resistance i.e., lowering γ. DOTS-Plus is modelled by reducing the time during which individuals infected with resistant strains are infectious i.e., increasing the rate of recovery from active disease with resistant strains, τ_r.

We will focus on the case $\theta = 1$ which corresponds to the two-strain TB model (3.1). However, the mixed-infection model has similar results as we will discuss.

3.6.1 Coexistence region

In Section 3.4 we fixed $\alpha = \beta_r/\beta_s$ and described a trend of strain coexistence at low transmission and dominance of the resistant strain at high transmission. This trend is verified when α is below a critical value, α_C. Above this critical value, resistance is always dominant irrespective of the transmission intensity. Therefore, the impact of control strategies on α_C gives an indication of its effect on the extent of the coexistence region. We evaluate the sensitivity and elasticity of α_C to the two parameters, γ and τ_r, manipulated by DOTS and DOTS-Plus, respectively. Using the terminology from mathematical demography in (Caswell, 2001), we introduce the partial derivatives

$$s_p = \frac{\partial \alpha_C}{\partial p} \quad \text{and} \quad e_p = \frac{p}{\alpha_C}\frac{\partial \alpha_C}{\partial p} = \frac{\partial \ln \alpha_C}{\partial \ln p}$$

to define, respectively, the sensitivity and elasticity of α_C to a parameter p, where p is γ or τ_r. Note that, since equal increments on a logarithmic scale correspond to equal proportions on an arithmetic scale, we can say that elasticity measures proportional sensitivity.

Table 3.2: Sensitivity and elasticity of α_C to γ and τ_r

p	Initial value (1)	Change of $\frac{1}{3}$ (2)	Sensitivity (3)	Elasticity (4)	Abs. variation in α_C (5) $\approx$ (2).(3)	New α_C $\alpha_C + (5)$	% varation in α_C $\frac{(2)}{(1)}.(4).100$
γ	0.003	-0.001	-0.0098	-3.7883×10^{-5}	9.7797×10^{-6}	0.7745	0.0013
τ_r	1.5	0.5	0.4510	0.8735	0.2255	1.0000	29.1157

Table 3.2 shows the sensitivities and elasticities of α_C to changes in γ and τ_r for the case of 1/3 of change in each parameter. Both changes increase α_C which implies an improvement on conditions to coexistence. Elasticity is approximately $-3.7883. \times 10^{-5}$ for γ and 0.8735 for τ_r, corresponding to a variation of approximately 0.001% and 29%, respectively. Thus, for the case of γ the improvement is almost undetectable.

More generally, we can compare the elasticity of α_C to the two parameters γ and τ_r, by looking to the quotient between absolute value of the elasticities:

$$\left|\frac{e_{\tau_r}}{e_\gamma}\right| = \frac{\mu\tau_r}{\omega\tau_s\gamma}\frac{1}{\alpha_C}. \tag{3.14}$$

Since the rate of endogenous reactivation of latent TB, ω is several orders of magnitude smaller than the death rate, μ, the rates of recovery under treatment, τ_r and τ_s are of the same order of magnitude and $\gamma\alpha_C$ is small, we conclude that the quotient is greater than one.

These results show that α_C is more sensitive to changes in the infectious period than in the proportion of sensitive TB treatment failure acquiring resistance. Therefore, the impact on the coexistence region is greatest for the DOTS-plus strategy.

3.6.2 Prevalence of infection

A complementary way to assess the effectiveness of the two control measures is to compare the equilibrium prevalence of resistant TB before and after the intervention. Interventions affect both the prevalence of resistant active TB cases in the population and the percentage of active TB cases that carry the resistant phenotype (Figure 3.10(a) and (b), respectively).

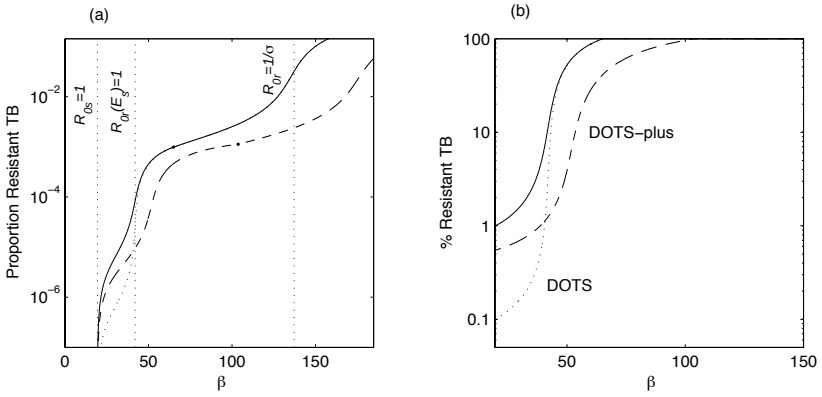


Figure 3.10: Impact of different control measures on resistant TB (case with $\alpha = 0.5$): **(a)** Proportion of resistant TB in total population; **(b)** Percentage of resistant phenotype in total TB cases. Full line corresponds to baseline proportion (no intervention), dotted line represents a DOTS like intervention ($\gamma = 0.0003$) and dashed line represents a DOTS-plus like intervention ($\tau_r = 2$).

DOTS-plus like interventions decrease not only the percentage of resistant TB in the coexistence region but also the overall prevalence of drug-resistant strains at all transmission potentials. As the results of the sensitivity analysis suggest, DOTS-plus can significantly increase the coexistence region which, by itself, inhibits the transmission of resistance due to strain competition. Moreover, this control strategy, shifts to the right the superinfection and reinfection thresholds of resistant strains ($R_{0r}(E_s) = 1$ and $R_{0r} = 1/\sigma$) delaying the predominance of drug resistance (see Figure 3.10(b)).

We can also observe that a DOTS like intervention has impact at low transmissibility. In fact, Figure 3.10(a) shows that DOTS is not effective above the superinfection threshold of resistant strains, $R_{0r}(E_s) = 1$. As we have stressed before, above this threshold the sensitive strains start to decline and the resistant strains become dominant. Therefore, any intervention that depends on the incidence of sensitive TB, I_s, has negligible impact. Indeed, above the superinfection threshold, the contribution of acquired drug resistance through treatment failure ($\gamma\tau_s I_s$) is minimum compared to cases caused by transmission of resistant strains. When the transmission potential is below this threshold, on the contrary, DOTS is the most effective strategy, both in relative and absolute terms. Moreover, in the limit case $\gamma = 0$, system (3.1) has another equilibrium, E_s, corresponding to the presence of only sensitive TB. Below the superinfection threshold of resistant strains i.e., for $R_{0r}(E_s) < 1$, this equilibrium is stable (region **IV** in Figure 3.4 (b)). This means that if acquired drug-resistance could be completely blocked ($\gamma = 0$) drug-resistant strains would be eradicated.

The control strategies modeled here have the same qualitative outcome in the mixed infection model as in the particular case $\theta = 1$. DOTS causes a decrease in resistant TB prevalence only below the superinfection threshold of resistant strains whereas DOTS-plus

forces a decrease in resistant TB prevalence for all endemic scenarios (results not shown).

Consequently, DOTS-plus may benefit regions of high endemic prevalence where infection with resistant strains wipes out the impact of DOTS. By contrast, DOTS is only effective for low endemic settings and in such scenarios it is, in fact, more suitable than DOTS-plus.

3.7 Discussion

By using simple models with reinfection we describe how thresholds in transmission shape the conditions for coexistence of resistant and sensitive TB strains and how this affects resistant TB prevalence and control.

First, we assumed that individuals carrying at least one resistant strain always manifest and transmit resistant TB. This simplification is justifiable by the fact that standard regimens confer a selection advantage to resistant strains, while the availability of treatment regimens that are recommended to combat resistance is limited. However, other possibilities can and should be considered. In van Rie *et al.* (2004), the authors conclude that treatment and adherence determine which strains are dominant in a mixed infection with sensitive and resistant strains. They find that treatment with second-line drugs leads to re-emergence of drug-sensitive strains. Furthermore, within-host competition may also favour drug-sensitive strains during latency.

We extended the first model by implementing two alternative progressions of mixed infections into active disease: a proportion θ activates resistant TB; while the remaining $(1-\theta)$ activates sensitive TB. When $\theta = 1$ (original model) coexistence is only observed at low transmissibility. By contrast, when $\theta < 1$ (mixed infection model) coexistence extends to higher transmissibility. A reinfection threshold marks the endemic level above which the majority of individuals harbor mixed infections. The fact that mixed infections can result in sensitive or resistant active infections, favours coexistence.

The results obtained are significantly different from those found in models where reinfection is not considered (Blower & Gerberding, 1998; Dye *et al.*, 2002). For R_0 near 1, the system is governed by primary transmission and coexistence is only possible when resistant strains are comparatively less transmissible (Austin *et al.*, 1997; Boni & Feldman, 2005). However, as we move away from $R_0 = 1$ reinfection starts to play a greater role. When the majority of individuals harbor mixed infections, the outcome of within-host competition shapes the frequency of resistance in the population and may sustain coexistence in the community.

The mechanisms that determine which phenotype prevails in mixed infections (during latency or active disease) are still poorly understood. And even if different pathways have been described (van Rie *et al.*, 2004), little is known about their frequency in the population. More epidemiological studies are needed to clarify this issue so that explicit, detailed models can be constructed and used to explore different interventions.

Reinfection also has implications on the effectiveness of different control strategies. A DOTS like intervention is ineffective against resistance in regions where primary resistance is common – above the superinfection threshold by resistant strains. It is precisely in those populations that a switch from DOTS to DOTS-plus can have the greatest impact. However, DOTS should continue to be the strategy of choice in populations where superinfection is rare. Even though DOTS and DOTS-plus interventions are much more complex than considered here, our work already highlights fundamental differences in outcome between the two strategies. Although coexistence results for $\theta = 1$ differ from those obtained with $\theta < 1$,

results concerning intervention efficacy are qualitatively the same.

In conclusion, primary resistance plays a fundamental role on the outcome of competition between sensitive and resistant strains in the host population. The strategy of choice to counteract the spread of resistance depends critically on the superinfection threshold of resistant strains.

Chapter 4

Multi-scale models in tuberculosis

4.1 Introduction

New molecular methods have challenged the widely accepted idea that TB is caused by a monoclonal bacterial infection. Several recent studies have shown that mixed infections (infections with more than one strain) are common, especially in highly endemic scenarios where reinfection is frequent (van Rie *et al.*, 2004; Warren *et al.*, 2004). In the context of drug resistant TB, mixed infections can have particular importance in determining treatment success. Reinfection by drug-resistant strains can be a powerful source of resistant TB as compared to resistance acquisition. While reinfection depends on the transmission rate, resistance aquisition depends on spontaneous mutation rate, which is constant throughout the different endemic scenarios. In Section 3.5, we extended the initial drug resistance tuberculosis model to accommodate alternative progressions of mixed infections into active disease – a proportion θ activates resistant TB, while the remaining $(1-\theta)$ activates sensitive TB. Parameter θ summarizes the mechanisms that determine the competition outcome between strains resulting in transmission. However, nothing was said about these mechanisms. This motivated a change in model scale from host-population models into within-host models in order to try to reveal some of the mechanisms underling strain competition in mixed infections. Our final goal is to link the within-host model results with the epidemiological model as a way to infer how competition at individual level combines with transmission at the population level. Although this work is in progress we describe some of our initial results in this Chapter.

4.2 Within-host models for *Mtb* mixed infections

Molecular epidemiological studies suggest that mixed infections are common (Warren *et al.*, 2004), and that once an individual is infected with both sensitive and resistant strains, a differential selection pressure will be imposed by treatment (van Rie *et al.*, 2004). Molecular analysis of mixed infections reveal several possible observed outcomes (van Rie *et al.*, 2004): (i) resistant strains may emerge when treating with first line anti-tuberculosis drugs, independently of a drug resistance acquisition event; (ii) sensitive strains may reemerge after matching drug regimen to the resistance pattern; (iii) sensitive TB may prevail due to the faster replication of sensitive strains. The possible outcomes here described are intuitive and expected and serve the purpose of model validation. The established formalism can be used

for further research.

We set as our first objective to build a model of within-host competition between sensitive and resistant strains, during an episode of active TB, that is able to reproduce the three different patterns described. To avoid the unknown mechanisms that lead to active TB upon reinfection or reactivation, we focus on the period of active TB, assuming that both strains are present at the moment of resuscitation of the granulomas, at some proportion. We allow the sensitive and resistant strains to have different relative fitnesses and consider the cases where mixed active disease is induced by primary infection (primary resistance) or endogenous reactivation of both strains. We do not consider the case of acquisition of resistance during treatment (acquired resistance). The model should contain four main mechanisms: bacterial growth, immune response, chemotherapy, and drug-resistance. Additional complexity can be included, such as bacterial heterogeneity, the impact of the immune response and treatment specificities. We start by a preliminary model where all the assumptions are reduced to its simplest form and we will then increase complexity.

There are several mathematical models addressing tuberculosis infection dynamics within a host. For some of them the main objective has been to describe the mechanisms that distinguish progression to active TB or maintenance in a latent state upon a first infection (Antia *et al.* (1996); Gammack *et al.* (2005) and references therein). A balance between the immune response and the bacteria ability to avoid that response is needed to achieve persistent infections. Other within-host models refer to the acquisition of resistance during antibiotic treatment in bacterial infections (Alavez-Ramirez *et al.*, 2007; Austin & Anderson, 1999; Austin *et al.*, 1998; Lipsitch & Levin, 1997, 1998; Nikolaou & Tam, 2005; Webb *et al.*, 2005). Treatment is considered for active TB and for latent TB and in all cases there is only one strain present at the beginning of infection. Resistant strains develop by mutation and treatment selection. The goal is to understand which is the appropriated treatment regimen to minimize the acquisition of drug resistance. It is also discussed the impact of noncompliance and heterogeneous bacterial populations on the time to develop drug resistant phenotype. To our best knowledge there are no models that address reinfection with a different strain, even out of the drug resistance context.

4.2.1 Ground Zero Model

We start by the simplest model that will set the basis for more complex ones. Both populations of bacteria grow exponentially and they are killed by the immune system at a constant rate γ. We consider two interventions, initially a first line treatment (T1) which is ineffective to the resistant population. Later the treatment is changed to match the pattern of resistance to a second line treatment (T2). Hence, the first intervention does not affect the resistant population and kills the sensitive bacteria at a rate μ_1. For the second intervention, both sensitive and resistant populations are killed at rates μ_2 and μ_3, respectively. The model can be written as the following system of differential equations (in time)

$$\left\{ \begin{array}{rcl} B_s' & = & (\nu - \gamma - (1-g)\mu_1 - g\mu_2)B_s \\ B_r' & = & (f\nu - \gamma - g\mu_3)B_r \end{array} \right. , \tag{4.1}$$

where B_s and B_r represent the bacterial load of sensitive and resistant strains, respectively.

Time is measured in days of infection. Parameter ν is the rate of growth of sensitive bacteria and f corresponds to the relative fitness between strain populations, measured as ability to grow within a host. Parameter g assumes the values 1 or 0 for treatments T_1 and T_2, respectively. The two equations are independent and can be easily solved for a given set of initial conditions $B_s(0) = B_{s0}$ and $B_r(0) = B_{r0}$:

$$\begin{cases} B_s(t) &= B_{s0}e^{(\nu-\gamma-(1-g)\mu_1-g\mu_2)t} \\ B_r(t) &= B_{r0}e^{(f\nu-\gamma-g\mu_3)t} \end{cases} . \tag{4.2}$$

We fix parameters ν and γ according to Lipsitch & Levin (1998). Table 4.1 summarizes parameter definition and values. Since $\nu > \gamma$, in the absence of treatment sensitive and

Table 4.1: Ground Zero model parameters

symbol	definition	value
ν	bacterial growth	0.4 day^{-1}
γ	killing rate by the immune system	0.3 day^{-1}
μ_i	killing rate by drugs	variable
f	relative fitness	variable
g	drugs targeting resistant population	0 or 1

resistant populations with sufficiently high relative fitness ($f > \gamma/\nu$) grow exponentially. When we introduce treatment T1 with $\mu_1 > \nu - \gamma$ then sensitive strains are contained. For the case of treatment T2, resistant strains can be controlled for sufficiently high treatment rate ($\mu_3 > f\nu - \gamma$).

Let us try to define and reproduce the behavior of a typical episode of active TB for different values of the relative fitness parameter and for a particular treatment schedule. For the first 135 days we let both populations grow in the presence of a immune response. At that point, we introduce treatment $T1$. The time of first treatment introduction for these simulations is chosen such that a monoclonal infection with a sensitive strain remains infectious for a period close to six months to match what was assumed for the epidemiological model in the previous chapter ($\tau_s = 2$). After 45 days, if resistant strains bacterial load is above its initial value, treatment is changed to a new set of drugs to match the resistant pattern. Figure 4.1 shows the dynamic behavior of the model for three scenarios: relative fitness $f = 1.1$, 0.9 and 0.5 that correspond to the three cases (i)–(iii) described before by the molecular studies. Two situations can happen: either relative fitness is too small and resistant strains are killed by the immune system ($f = 0.5$) or, first treatment gives an advantage to the resistant strains, that grow above the sensitives and take over the infection ($f = 1.1$ and $f = 0.9$). In this case, it is necessary to introduce a second set of drugs to kill the resistant bacteria. For intermediate values of f, this can give again advantage to sensitive strains ($f = 0.9$).

From this simple model we can already have an idea of which are the transmission patterns that can be generated by mixed infections depending on the relative fitness, f. However, it is evident that it cannot reproduce other simple features that are important to this biological system. Namely, here bacteria can grow indefinitely in the absence of the immune response and the immune response is constant over time. There is no true competition between the strains since they grow independently of each other. The fact that both populations never saturate makes the time of intervention to have a false impact on the intervention outcome.

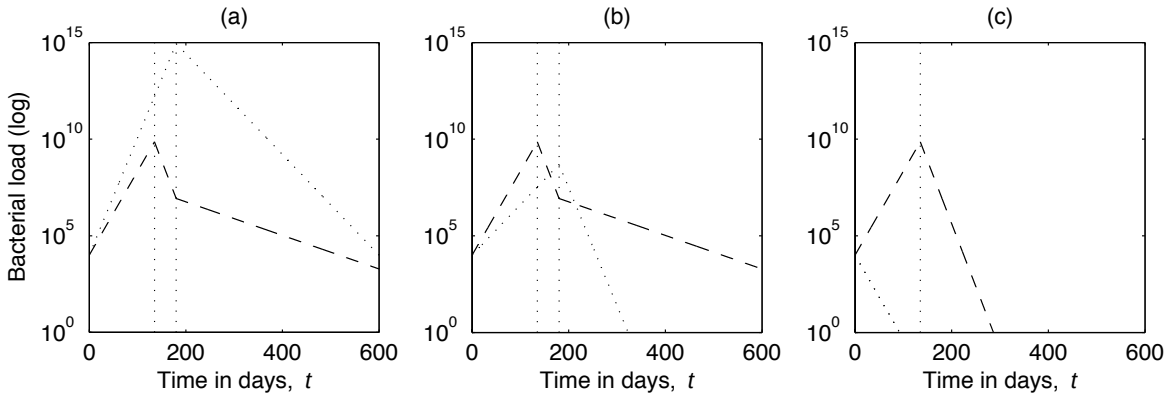


Figure 4.1: Episode of active TB: dynamical behaviour of the Ground Zero model for three different relative fitness values: $f = 1.1$, $f = 0.9$ and $f = 0.5$. Dashed, dotted and full curves represent sensitive, resistant and total bacteria load, respectively. For all scenarios initial conditions are $B_{s0} = B_{r0} = 10^4$ that is approximately the carrying capacity of a granuloma. Treatment T1 was introduced at $t = 135$ *days* with $\mu = 0.25\ days^{-1}$ and changed by treatment T2 at $t = 180$ *days* with $\mu_2 = 0.12$ and $\mu_3 = 0.2\ day^{-1}$ and $g = 1$, indicated by the vertical lines.

Moreover, the treatment or the immune response can only have two alternatives: either it eliminates completely the bacteria population or it keeps growing. It is not possible that bacterial population settles at an intermediate level, alone or in coexistence.

4.2.2 Model with non-constant immune response

The first mechanism of interaction between the tow strains is competition via the immune system. After that it is more important which drugs are used and the rates at which they kill each bacterial population. A model for persistent infections is adapted from Antia *et al.* (1996) to account for two distinct populations. We let X represent the intensity of the immune response to bacteria and B_s, B_r represent the bacterial load of sensitive and resistant strains, respectively. In the absence of immune response we assume that the parasite grows exponentially at rate ν. The intensity of the immune response is assumed to be proportional to the density of the T-cells specific to the bacteria. These cells migrate from the thymus at a constant rate a, and die at a rate d. Bacteria stimulate the proliferation of T-cell at a rate that is proportional to the density of bacteria at low bacteria densities and that saturates at high bacteria densities. Sensitive and resistant bacteria compete indirectly through the immune response. Equations are scaled so that in the absence of parasite the density of immune cells equals unity ($a = d$). Relative magnitudes of the various parameters for biological reasonable cases are $\gamma < a = d < \nu,\ s \approx 1 << k$ (Antia *et al.*, 1996).

We consider a protected class P for each strain, mimicking the dormant stage (Antia *et al.*, 1996; Lipsitch & Levin, 1997). During this metabolic state bacteria do not stimulate the immune response and they are not affected by drugs action. In fact, it has been proposed that bacteria can undergo four intermediate stages according to its metabolic state (revision on the subject by Urlichs & Kaufmann (2002)). Parameters m and n are the rates of transition between the two metabolic states. The equations describing the dynamics of bacteria and the

immune response are:

$$\begin{cases} B_s' &= \nu B_s - \gamma X B_s - ((1-g)\mu_1 + g\mu_2 + m)B_s + nP_s \\ B_r' &= f\nu B_r - \gamma X B_r - (g\mu_3 + m)B_r + nP_r \\ P_s' &= mB_s - nP_s \\ P_r' &= mB_r - nP_r \\ X' &= a + sX(\dfrac{B_s + B_r}{k + B_s + B_r}) - dX \end{cases} \tag{4.3}$$

The system has a trivial equilibrium where the bacteria density is zero and two steady states corresponding to the presence of each strain alone. A coexistence equilibrium is also possible but for a particular relation of the parameters, which has no biological meaning.

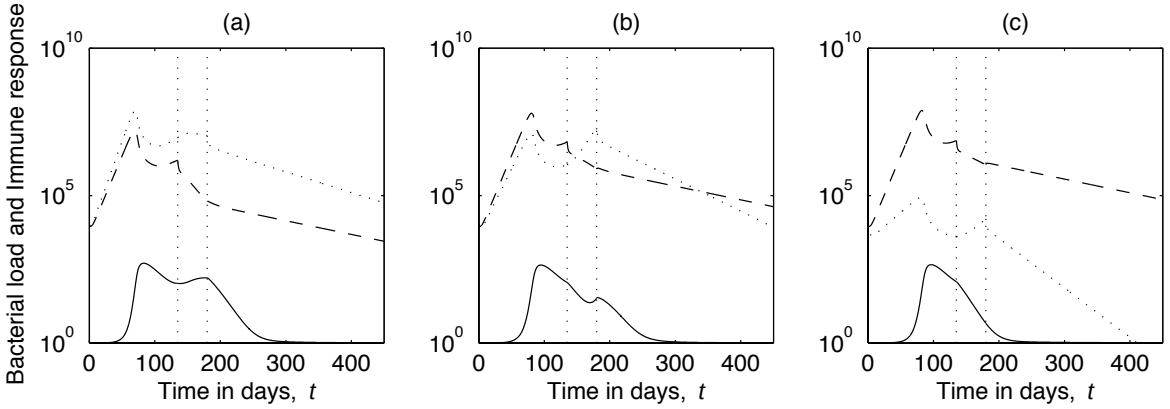


Figure 4.2: Episode of active TB: dynamical behaviour of the model with proportional immune response for different values of the relative fitness: $f = 1.1$, $f = 0.9$ and $f = 0.5$. Dashed and dotted curves represent sensitive and resistant active bacteria. Full curve represents the immune response. For all scenarios initial conditions are $B_{s0} = B_{r0} = 10^4$ that is approximately the carrying capacity of a granuloma. Parameter values are: $\nu = 0.4$, $\gamma = 0.003$, $a = d = 0.1$, $s = 1$, $k = 10^8$, $m = 0.5$ and $n = 0.1$. Treatment T1 was introduced at $t = 135$ *days* with $\mu_1 = 0.6$ $days^{-1}$ and changed by treatment T2 at $t = 180$ *days* with $\mu_2 = 0.47$ and $\mu_3 = 0.55$ day^{-1} and $g = 1$, indicated by the vertical lines.

The dynamical behaviour of the model is illustrated in Figure 4.2, for the cases $f = 1.1$, 0.9 and 0.5. Patterns of strain dominance observed are closely the same as before. Since the immune response is proportional to the bacterial load, it reaches a lower bacterial load, it takes longer to clear infection and it is not possible to eliminate the resistant population without introducing the treatment T2, even for the case of very low relative fitness ($f < \gamma/\nu$).

4.3 From a within- to a between-host models

Multi-scale models, integrating within and between-host models, have been proposed in different contexts. In the context of evolution of virulence, for the study of chronic diseases (Gilchrist & Coombs, 2006) or the existence of evolutionary stable coexistence of pathogens driven by superinfection (Boldin & Diekmann, 2008). The goal is to link traits that affect within-host dynamics, such as virulence, to effects or other traits acting on the between-host level, such as transmission. In a different context, multi-scale models were used to investigate drug-resistant bacterial epidemics in hospitals (Austin & Anderson, 1999; Webb *et al.*, 2005), by integrating the selection mechanism on resistant aquisition into the transmission process.

The linkage between the two levels can be done directly through model variables or by using common parameters.

We have shown that the simple models explored here can already reproduce the patterns of infections described in molecular studies. As second objective, we articulate the within-host model results with the epidemiological model from Section 3.5. Through the within-host model, we define a 'typical' active TB episode with a mixed infection depending on the relative fitness (f) of the strains and assuming a given treatment schedule. We estimate the length of the infectious period (p) for mixed infections and the proportion at which each strain would be transmitted (θ), for a certain 'typical' episode. Then, we use these as parameters for the epidemiological model.

To illustrate our purpose, we use the results of the Ground Zero model. First, we define a transmission threshold as the bacterial load above which we consider that at least one of the strains can be transmitted. We set its value arbitrarily at 10^5, represented in Figure 4.3 (a) by the horizontal line. Hence, the transmission threshold determines the infectious period duration. Secondly, we define the average proportion at which resistant strains are present

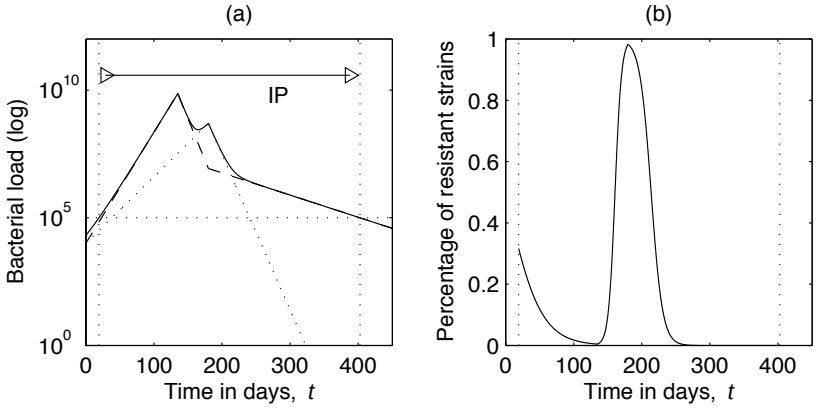


Figure 4.3: **(a)** Episode of active TB for $f = 0.9$. Dashed, dotted and full curves represent sensitive, resistant and total bacteria load, respectively. Vertical lines correspond to the beginning and end of the infectious period. Horizontal line marks the transmission threshold. **(b)** Percentage of drug resistant bacteria during the infectious period.

and therefore can be transmitted during the infectious period by

$$\theta = \frac{1}{p}\int_0^p \frac{B_r(t)}{B_s(t) + B_r(t)}\mathrm{d}t. \tag{4.4}$$

Illustrated in panel (b) of Figure 4.3. Finally, for the particular case of the Ground Zero model with relative fitness $f = 0.9$, we would get mixed infections lasting for a period of approximately $p = 384$ days, during which resistant strains would have a 16% ($\theta = 0.16$) chance of being transmitted.

A more systematic way to link the within- and between-host model is to derive not only the proportion of resistant strains transmitted θ but also the rate of recovery from a mixed infection τ_m, defined by $\tau_m = 365/p$ yr^{-1}, as functions of the relative fitness f and relate

the long-term behavior of the epidemiological model with this within-host parameter. To do so, we choose to use a slightly different version from the epidemiological model presented in Section 3.5, for which we can better integrate parameters θ and τ_m resulting from the within-host model. The refined version of the between-host model is schematically represented in Figure 4.4. We define a mixed infectious class (I_m), where individuals can have different

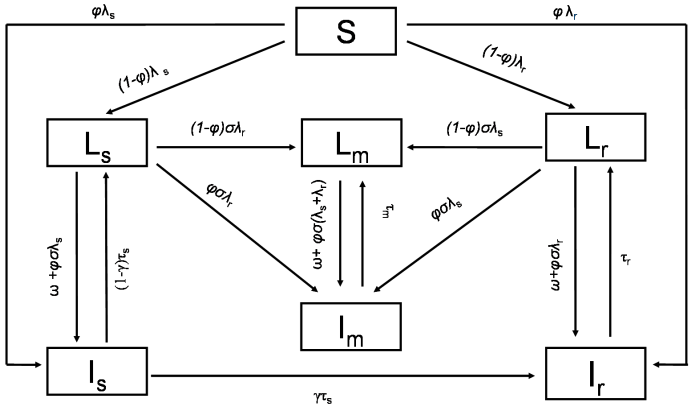


Figure 4.4: Epidemiological model for transmission of drug-sensitive and -resistant strains. Individuals are classified according to infection state into susceptible (S), latently infected (L) and infectious (I). Parameters are the force of infection (λ), the death and birth rate (μ), the proportion of individuals developing active TB (ϕ), the reinfection factor (σ), the rate of reactivation (ω), the rate of recovery under treatment (τ) and the proportion of resistance acquisition (γ). Subscripts s, r and m relate to sensitive, resistant or mixed infection, respectively.

drug resistance patterns depending on the relative fitness parameter f. The frequency at which resistant strains are transmitted by mixed infectious individuals, θ, determines the contribution of the mixed infections to the overall force of infection $\lambda_r = \beta_r(I_r + \theta I_m)$ and $\lambda_s = \beta_s(I_s + (1-\theta)I_m)$. The rate at which these individuals recover is defined by the infectious period duration, $\tau_m = 365/p$ yr^{-1}. As before we assume that in an event of resistance acquisition, resistant strains replace the sensitive ones, giving rise to a resistant infection. The new model equations are given by

$$\begin{cases} S' &= b - (\lambda_s + \lambda_r + \mu)S \\ L_s' &= (1-\phi)\lambda_s S - (\omega + \phi\sigma\lambda_s + \sigma\lambda_r + \mu)L_s + (1-\gamma)\tau_s I_s \\ I_s' &= \phi\lambda_s S + (\omega + \phi\sigma\lambda_s)L_s - (\tau_s + \mu + \delta)I_s \\ L_m' &= (1-\phi)\sigma(\lambda_s L_r + \lambda_r L_s) - (\omega + \phi\sigma(\lambda_s + \lambda_r) + \mu)L_m + \tau_m I_m \\ I_m' &= \phi\sigma(\lambda_s L_r + \lambda_r L_s) + (\omega + \phi\sigma(\lambda_s + \lambda_r))L_m - (\tau_m + \mu + \delta)I_m \\ L_r' &= (1-\phi)\lambda_r S - (\omega + \sigma\lambda_s + \phi\sigma\lambda_r + \mu)L_r + \tau_r I_r \\ I_r' &= \phi\lambda_r S + \gamma\tau_s I_s + (\omega + \phi\sigma\lambda_r)L_r - (\tau_r + \mu + \delta)I_r \end{cases} \tag{4.5}$$

Figure 4.5 (a) and (b) show the curves for of θ and τ_m when varying the relative fitness f. For very low relative fitness, mixed infections resemble a sensitive infection. But for sufficient high relative fitness, a second-line treatment must be introduced to control resistant bacteria.

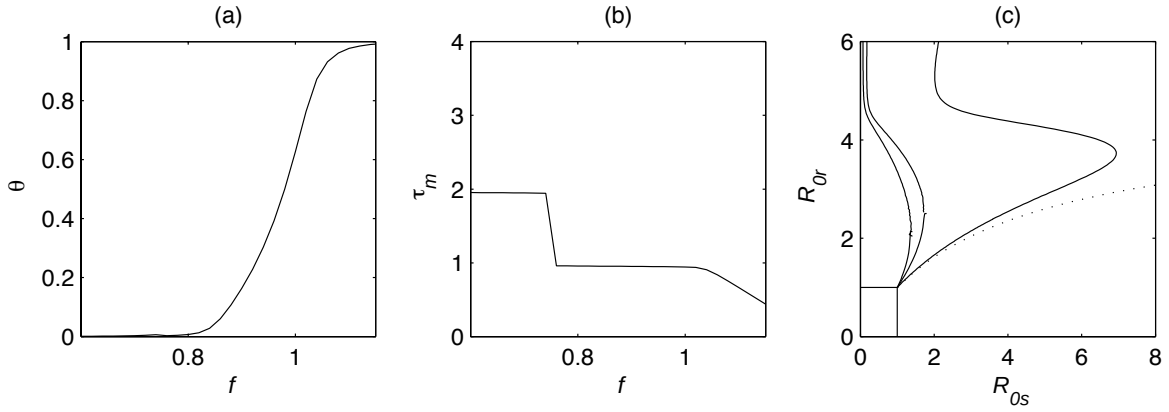


Figure 4.5: **(a)** and **(b)** Curves θ and τ_m when varying the relative fitness f. Results using the Ground Zero model for the fixed parameters from table 4.1 and for the treatment T1 and T2 implemented as before. **(c)** Long-term epidemiological outcome of the between-host model: bifurcation diagram on R_{0s} and R_{0s}. Curves separate coexistence region from persistence of only resistant strains for different values of f: 0.7, 1 and 1.1, from left to right. These values correspond to (τ_m, θ) of (1.9, 0.003), (0.94, 0.63) and (0.66, 0.97), respectively.

This explains the sudden increase in the infectious period (discontinuity) and in the possibility of resistant strains transmission.

Finally, the resulting long-term epidemiological outcome depends on within-host relative fitness instead on θ. This has important consequences since the first can be measured directly whereas the second was a theoretical construction. Results are represented in panel (c) of Figure 4.5. As expected, coexistence region (region to the right of the curves) decreases with the relative within-host fitness f, as the contribution of resistant strains for transmission θ, increases (compare with Figure 3.8 in Section 3.5).

Note that in both within- and between-host models we can have a relative fitness parameter: at the epidemiological level, α is the relative transmission coefficient, corresponding to the relative ability of strains to be transmitted; and at the individual level, f is the relative growth rate of strains. Surely, these two strain traits are not independent, since the ability to be transmitted depends on the bacterial load. However, the relation between them is still an open question.

4.4 Final remarks

The simple models presented can already reproduce the different patterns described by the molecular studies. And the linkage to the epidemiological model provides a comprehensive relation between pathogen-specific growth rates and general between- host transmission ability. The lack of more detailed information on some crucial processes involved is compensated here by the imposition of arbitrarily threshold values for which we do not have direct evidences, such as the initial frequency of bacteria, the time of treatment introduction or the transmission threshold. Consequently, the model dynamics which is sensitive to these quantities still requires validation.

At this point we decided to give a step back and investigate the within-host dynamics of mixed infection more deeply. There are challenging questions about the reinfection process that are still far from being understood and mathematical models are suitable to explore the different alternative mechanisms and to formulate suitable hypotheses. Existing models for the course of a primary TB infection (Gammack *et al.*, 2005; Murphy *et al.*, 2003, 2002) can be adapted to incorporate reinfection by a different strain. Increasing data available on the reinfection process and subsequently progression to disease can then be used to calibrate and validate the models (Cosma *et al.*, 2004, 2008).

Chapter 5

Including host heterogeneity in an SIRI model

5.1 Introduction

Heterogeneity in susceptibility and infectivity is an important feature of many infectious diseases and has been considered to improve the accuracy of epidemiological models. In the analysis of these models, focus has been on the impact of heterogeneity in the final size of epidemics (Ball, 1985; Miller, 2007) and on its consequences to disease control (Anderson & Britton, 1998; Britton, 1998) and data interpretation (Gart, 1968; Anderson & May, 1991). In the context of SIR epidemic models, it has been shown that the final size of the epidemic is reduced when the risk of infection is heterogeneously distributed in the population, both for the deterministic and the stochastic formulations (Gart, 1968; Ball, 1985; Anderson & Britton, 1998). More recently, results were extended to the investigation of epidemic spread on a random network (Miller, 2007).

In this work we explore the consequences of host heterogeneity in the susceptibility to infection for endemic models for which immunity conferred by infection is not fully protective – SIRI model, introduced in Chapter 2. Here, the model is expanded to accommodate multiple risk groups classified accordingly to risk of infection. We are concerned not only with the impact on disease prevalence but also on how transmission changes the risk profile of the population groups that are subject to reinfection.

The SIRI model exhibits two important thresholds in transmission: the epidemic threshold that marks the transmission intensity necessary to maintain disease endemic in a population; and the reinfection threshold that indicates whether self-sustained transmission occurs in a population which has developed a degree of partial immunity. The reinfection threshold separates two fundamentally distinct model behaviors. Low endemic levels with SIR-like transmission are maintained below threshold, while high endemic levels with SIS-like transmission characterise the regime above threshold. Therefore, first we consider the case of SIR and SIS models, exploring their simplicity and mathematical tractability to extract general trends. We describe how disease prevalence, risk profiles for specific population compartments, and contribution of the high-risk group to overall incidence, change with the parameters describing heterogeneity. Second, the same framework is used to explore the SIRI model. Of particular interest is the interplay between reinfection and the risk profile for the uninfected compartments, S and R.

The results offer a plausible explanation for observations of higher than expected reinfection rates. In particular, rates of reinfection that surpass rates of first infection have been reported for tuberculosis in a high transmission setting in South Africa (Verver *et al.*, 2005). One could attribute this effect to some form of immunologically dependent enhancement whereby immunological memory would render individuals more susceptible to subsequent infections. An alternative hypothesis, suggested by the analysis presented here is that relatively high rates of reinfection can result from the presence of a high-risk group that, being at higher frequency in the recovered compartment due to selection imposed by the first infection, can sustain rates of reinfection that are, on average, higher than the rates of first infection even in the presence of partially protective immunity.

Heterogeneity has many implications for public health policy. In particular, we characterise how the impact of vaccination strategies varies with transmission intensity and quantify the benefit of targeting high-risk groups. Moreover we analyze how interventions that affect susceptibility to infection can improve condition for disease control.

The results presented were recently published Rodrigues *et al.* (2009).

5.2 The model

To incorporate heterogeneity in the infection risk in an SIRI transmission model we use a formulation analogous to those presented by Ball (1985) and (Coutinho *et al.*, 1999) for SIR epidemic models. We assume that the population is divided in n different subgroups according to the susceptibility to infection, α_i. Within each risk group, individuals are classified according to their disease history into susceptible, infectious or recovered. A schematic version of the model is shown in Figure 5.1.

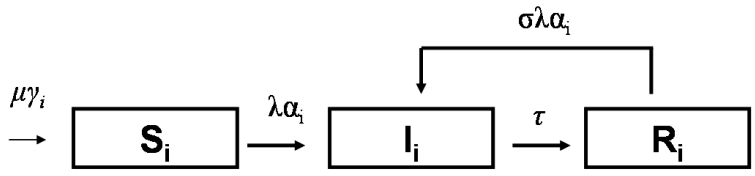


Figure 5.1: SIRI model with heterogeneous susceptibility to infection. The population is divided into Susceptible (S_i), Infectious (I_i) and Recovered (R_i) classes, where the index i refers to the risk group to which the individuals belong. Individuals are born at rate μ and enter the susceptible compartments in proportions γ_i. susceptible individuals are infected at a rate $\alpha_i\lambda = \alpha_i\beta I$, where α_i denotes the risk factor, β is the transmission coefficient and I is the proportion of infectious individuals. Infectious individuals recover at a rate τ and recovered individuals have a reduced rate of reinfection according to the factor σ.

It is assumed that the n risk groups have constant size over time and represent different proportions of the total population, γ_i, such that $\sum_{i=1}^{n}\gamma_i = 1$. Individuals are born into each group at the rate $\mu\gamma_i$. We use S_i, I_i and R_i as the proportion of the total population that are susceptible, infected or recovered, respectively, and belong to the i risk group. Hence we have $\sum_{i=1}^{n}(S_i + I_i + R_i) = 1$. In the following we denote by I the proportion of infectious individuals in the population, that is $I = \sum_{i=1}^{n} I_i$.

For concreteness, we fix the parameters as described in table 5.1. The table describes an

Table 5.1: Parameters of the SIRI model with heterogeneous susceptibility to infection

symbol	definition	value
β	transmission coefficient	variable
σ	factor reducing the risk of infection as a result of acquired	0, 1 or
	immunity to a previous infection	0.25
μ	death and birth rate	1/70 yr^{-1}
τ	rate of recovery	52 yr^{-1}
γ_i	relative size of each risk group	variable
α_i	relative risk of infection of each risk group	variable

average life expectancy of 70 years (that is $\mu = 1/70$) and an average infectious period of one week (that is $\tau = 52$). The factor reducing the risk of infection as a result of acquired immunity is $\sigma = 0.25$. For the limiting cases of the SIR and SIS models, parameter σ is 0 or 1, respectively. Parameters β, γ_i and α_i are varied to explore different scenarios for transmission intensity and host heterogeneity. Each risk group has an average risk of infection that differs from the population average by a factor α_i, which we refer to as the relative risk of infection (Gart, 1968; Ball, 1985). We assume that this factor controls the rate of infection and reinfection in the ith risk group. In general, the parameters are chosen to resemble an acute respiratory infection in a developed country. However, we stress that the results are valid for a wider set of parameters. Differences reside more on the quantitative than on the qualitative behavior. The model can be written as a system of $3n$ differential equations

$$\begin{cases} S_i' &= \mu\gamma_i - \lambda\alpha_i S_i - \mu S_i \\ I_i' &= \lambda\alpha_i S_i + \sigma\lambda\alpha_i R_i - (\tau+\mu) I_i \\ R_i' &= \tau I_i - \sigma\lambda\alpha_i R_i - \mu R_i, \quad i = 1, \ldots, n, \end{cases} \tag{5.1}$$

where $\lambda = \beta I$. To ensure comparison between different assumptions on risk distribution, including the comparison with the homogeneous version of the model, we impose the normalization $\bar{\alpha} = \sum \alpha_i \gamma_i = 1$.

Throughout this Chapter we analyze the case $n = 2$. We denote by γ the proportion of individuals belonging to the low-risk group (that is, $\gamma_1 = \gamma$ and $\gamma_2 = 1 - \gamma$). For a given population structure (γ) we vary the infection risk distribution by changing α_1, obtaining α_2 through the normalization $\alpha_1\gamma + \alpha_2(1-\gamma) = 1$.

We use the variance as a summary measure of variations in α_1,

$$var_\alpha = (\bar{\alpha} - \alpha_1)^2\gamma + (\bar{\alpha} - \alpha_2)^2(1-\gamma) = \frac{(1-\alpha_1)^2\gamma}{1-\gamma}. \tag{5.2}$$

Note that for a given population structure the variance is a decreasing function of α_1. The homogeneous model is obtained for $\alpha_1 = \bar{\alpha} = 1$ which, consistently, corresponds to zero variance.

5.2.1 Basic reproduction number

The basic reproduction number is an important concept in the study of epidemiological models. We recall from Chapter 2 (Section 2.4.1), that in the case of the corresponding model for

homogeneous populations ($\alpha_1 = 1$) the basic reproduction number is given by

$$R_0 = \frac{\beta}{\tau + \mu}. \tag{5.3}$$

Considering the heterogeneous model, the basic reproduction number is not altered. In fact,

$$R_0^{HET} = \frac{\beta}{\tau + \mu} \sum_{i=1}^{2} \alpha_i \gamma_i = \frac{\beta \bar{\alpha}}{\tau + \mu} = \frac{\beta}{\tau + \mu} = R_0.$$

For a more detailed discussion on the calculation of the basic reproduction number in heterogeneous populations see Hyman & Li (2000). A threshold condition for endemicity is given by $R_0 = 1$ (the disease dies out if $R_0 < 1$ and becomes endemic if $R_0 > 1$).

Note that the basic reproduction number for the entire population is a weighted average of the basic reproduction number within each independent risk group, R_{0i}, given by

$$R_0^{HET} = \sum_i \frac{\alpha_i \beta}{\tau + \mu} \gamma_i = \sum_i R_{0i} \gamma_i.$$

Therefore, if the basic reproduction number for each group is greater than one, then the disease is also endemic in the entire population. On the other hand, it is not necessary to have all reproductive numbers greater than one to have endemicity.

5.3 The limit cases, SIR ($\sigma = 0$) and SIS ($\sigma = 1$)

Before studying the SIRI model, we analyze the impact of host heterogeneity in the case of SIR and SIS models, corresponding to $\sigma = 0$ and $\sigma = 1$, respectively. The identification between the SIS model and our model with $\sigma = 1$ is made in a natural way, by collapsing the classes S and R of this last model into a class $S + R$, which we identify with the susceptible class of the SIS model. However, in order to make possible the comparison between the limit case with $\sigma = 1$ and the intermediate SIRI model, in what follows we keep distinct the S and R classes even for $\sigma = 1$. We will consider the class $S + R$ in Remark 5.3.4, where we examine the effect of heterogeneity on the prevalence in the SIS framework.

5.3.1 Endemic equilibrium

For $\sigma = 0$ or 1, system (5.1) has one disease-free equilibrium of the form $E_0^\sigma = (\gamma, 1 - \gamma, 0, 0, 0, 0)$. Above $R_0 = 1$, the system has also an endemic equilibrium, E_1^σ. Stability results for these equilibria are stated in the two theorems below. We use the superscript σ to denote the correspondence with the SIR ($\sigma = 0$) or the SIS ($\sigma = 1$) models.

Theorem 5.3.1. *For $\sigma = 0$ or 1, the disease-free equilibrium, E_0^σ, of system* (5.1) *is globally asymptotically stable if $R_0 < 1$ and it is unstable for $R_0 > 1$.*

Proof. First lets rewrite the model equations. For $\sigma = 0$ (SIR model), from system (5.1) we obtain

$$\begin{cases} S_1' &= \mu\gamma - \beta I \alpha_1 S_1 - \mu S_1 \\ S_2' &= \mu(1-\gamma) - \beta I \alpha_2 S_2 - \mu S_2 \\ I' &= \beta I(\alpha_1 S_1 + \alpha_2 S_2) - (\tau + \mu) I. \end{cases} \tag{5.4}$$

When $\sigma = 1$ (SIS model), we can collapse the recovered classes into the susceptible ones in system (5.1). Then, if we denote for simplicity by S_i the classes $S_i + R_i$, $i = 1, 2$, and use the fact that $\gamma_i = S_i + I_i$, we obtain the following system

$$\left\{ \begin{array}{rcl} S_1' & = & (\tau+\mu)\gamma - \beta I \alpha_1 S_1 - (\tau+\mu) S_1 \\ S_2' & = & (\tau+\mu)(1-\gamma) - \beta I \alpha_2 S_2 - (\tau+\mu) S_2 \\ I' & = & \beta I(\alpha_1 S_1 + \alpha_2 S_2) - (\tau+\mu) I. \end{array} \right. \tag{5.5}$$

Note that system (5.5) is equivalent to an SIR model where the birth and death rate are equal to $\tilde{\mu} = \tau + \mu$ and the recovery rate is $\tilde{\tau} = 0$.

Let us first consider $\sigma = 0$. The Jacobian of system (5.4) at the disease-free equilibrium is

$$J(E_0) = \begin{bmatrix} -\mu & 0 & -\beta\alpha_1\gamma \\ 0 & -\mu & -\beta\alpha_2(1-\gamma) \\ 0 & 0 & \beta - (\tau+\mu) \end{bmatrix}.$$

The eigenvalues of this matrix are $-\mu$ and $\beta - (\tau + \mu)$. So we conclude that E_0 is locally asymptotically stable for $R_0 < 1$ and unstable for $R_0 > 1$. Moreover, system (5.4) is equivalent to system (3.1) in Hyman & Li (2005) for $n = 2$. In Theorem 3.1 of that paper, the authors prove the global stability for the disease-free equilibrium for $R_0 < 1$.

For the case $\sigma = 1$, calculations can be repeated using $\tilde{\mu} = \tau + \mu$ as the new birth and death rates and $\tilde{\tau} = 0$ as the new rate of recovery. □

Theorem 5.3.2. *Let $\sigma = 0$ or 1 and assume that $R_0 > 1$. Then system (5.1) has exactly one endemic equilibrium, E_1^σ, that is globally asymptotically stable.*

Proof. The second member of system (5.4) vanishes at the equilibria. From the two first equations we get a relation between S_i and I: $S_1 = \dfrac{\gamma\mu}{\mu + \beta I \alpha_1}$ and $S_2 = \dfrac{(1-\gamma)\mu}{\mu + \beta I \alpha_2}$. Substituting in the third one we get $(\tau+\mu)\dfrac{P(I)}{Q(I)} I = 0$, where $P(I) = a_2 I^2 + a_1 I + a_0$ with $a_2 = -\alpha_1\alpha_2 R_0^2(\tau+\mu)^2$, $a_1 = R_0(\tau+\mu)\mu(\alpha_1\alpha_2 R_0 - (\alpha_1+\alpha_2))$ and $a_0 = \mu^2(R_0 - 1)$ and $Q(I) = (\mu + \beta I\alpha_1)(\mu + \beta I \alpha_2)$. Note that for $I \geq 0$ we have $Q(I) > 0$. We conclude that the I coordinate of the nontrivial equilibria of system (5.4) will correspond to a positive solution of $P(I) = 0$. Since $a_2 < 0$ and $a_0 > 0$ for $R_0 > 1$ we conclude that the polynomial P has exactly one positive solution of the form:

$$I^0(R_0) = \frac{-a_1 - \sqrt{a_1^2 - 4a_2 a_0}}{2a_2} \tag{5.6}$$

and this proves the first part of the theorem.

In what concerns stability, system (5.4) is equivalent to system (3.1) in Hyman & Li (2005) for $n = 2$. In Theorem 3.2 of that paper, the authors prove the stability for the endemic equilibrium for $R_0 > 1$ via Lyapunov stability theory.

As in the previous proof, for the case $\sigma = 1$, calculations can be repeated using $\tilde{\mu} = \tau + \mu$ as the new birth and death rates and $\tilde{\tau} = 0$ as the new rate of recovery. □

Remark 5.3.1. *From the Proof of theorem 5.3.2 it is possible to establish a relation between the disease prevalence of both models. In fact, for every $\gamma \in (0,1)$ and $\alpha_1 \in (0,1]$ we have*

$$I^1 = \frac{\tau + \mu}{\mu} I^0, \tag{5.7}$$

where I^0 and I^1 represent the disease prevalence at equilibrium for the SIR and SIS models, respectively. This relation is systematically used to extend the proofs from the case $\sigma = 0$ to the case $\sigma = 1$.

We can also conclude that for all $\gamma \in (0,1)$ and $\alpha_1 \in (0,1]$

$$\lim_{R_0 \to +\infty} I^0 = \frac{\mu}{\tau + \mu} \text{ and } \lim_{R_0 \to +\infty} I^1 = 1. \tag{5.8}$$

Remark 5.3.2. *From the Proof of Theorem 5.3.2, taking into account Remark 5.3.1, we recover the expression of the endemic equilibrium for the homogeneous model, both for the $\sigma = 0$ and the $\sigma = 1$ cases, by using $\alpha_1 = \alpha_2 = 1$ (or $\alpha_1 = 1$):*

$$I^0_{Hom}(R_0) = \frac{\mu}{\tau + \mu}\Big(1 - \frac{1}{R_0}\Big) \text{ and } I^1_{Hom}(R_0) = 1 - \frac{1}{R_0}. \tag{5.9}$$

We analyze the impact of heterogeneity on disease prevalence at equilibrium. Figure 5.2 illustrates how disease prevalence changes for different assumptions on population structure and distribution of infection risk for the SIR and SIS models. We observe that for a fixed R_0, the equilibrium disease prevalence is lower when assuming heterogeneous populations. From each plot, it is evident that for fixed γ, the prevalence curve goes down as variance increases. Comparing the three plots it is also apparent that for fixed α_1, the prevalence curve goes down as the proportion of the population at low risk (γ) increases. Moreover, the disease prevalence appears to increase monotonically with the transmission potential, R_0. The following theorem summarizes these results.

Theorem 5.3.3. *Let $\sigma = 0$ or $\sigma = 1$ and let I^σ, $\sigma = 0, 1$, designate the disease prevalence at equilibrium, for the corresponding system (5.1) with $R_0 > 1$. Then, for γ and $\alpha_1 \in (0,1)$*

$$\frac{\partial I^\sigma}{\partial \gamma} \leq 0, \quad \frac{\partial I^\sigma}{\partial \alpha_1} \geq 0 \tag{5.10}$$

$$\text{and} \quad \frac{\partial I^\sigma}{\partial R_0} > 0, \quad \sigma = 0, 1. \tag{5.11}$$

Proof. First, let $\sigma = 0$ and denote by I^0 and I^{0^*}, respectively, the unique positive and negative roots of the polynomial P defined in the Proof of Theorem 5.3.2. Differentiating $P(I^0) = 0$ with respect to a parameter ϵ, we get

$$\frac{\partial I^0}{\partial \epsilon} = \frac{-\frac{\partial a_2}{\partial \epsilon} I^{0^2} - \frac{\partial a_1}{\partial \epsilon} I^0 - \frac{\partial a_0}{\partial \epsilon}}{2a_2 I^0 + a_1}. \tag{5.12}$$

Note that we have $I^0 + I^{0*} = -a_1/a_2$ and that $a_2 < 0$. Hence we conclude that the denominator of (5.12) verifies $2a_2 I^0 + a_1 = a_2(2I^0 + a_1/a_2) = a_2(2I^0 - I^0 - I^{0*}) = a_2(I^0 - I^{0*}) < 0$. As a consequence of this fact, the sign of (5.12) will be the opposite of the one of the numerator.

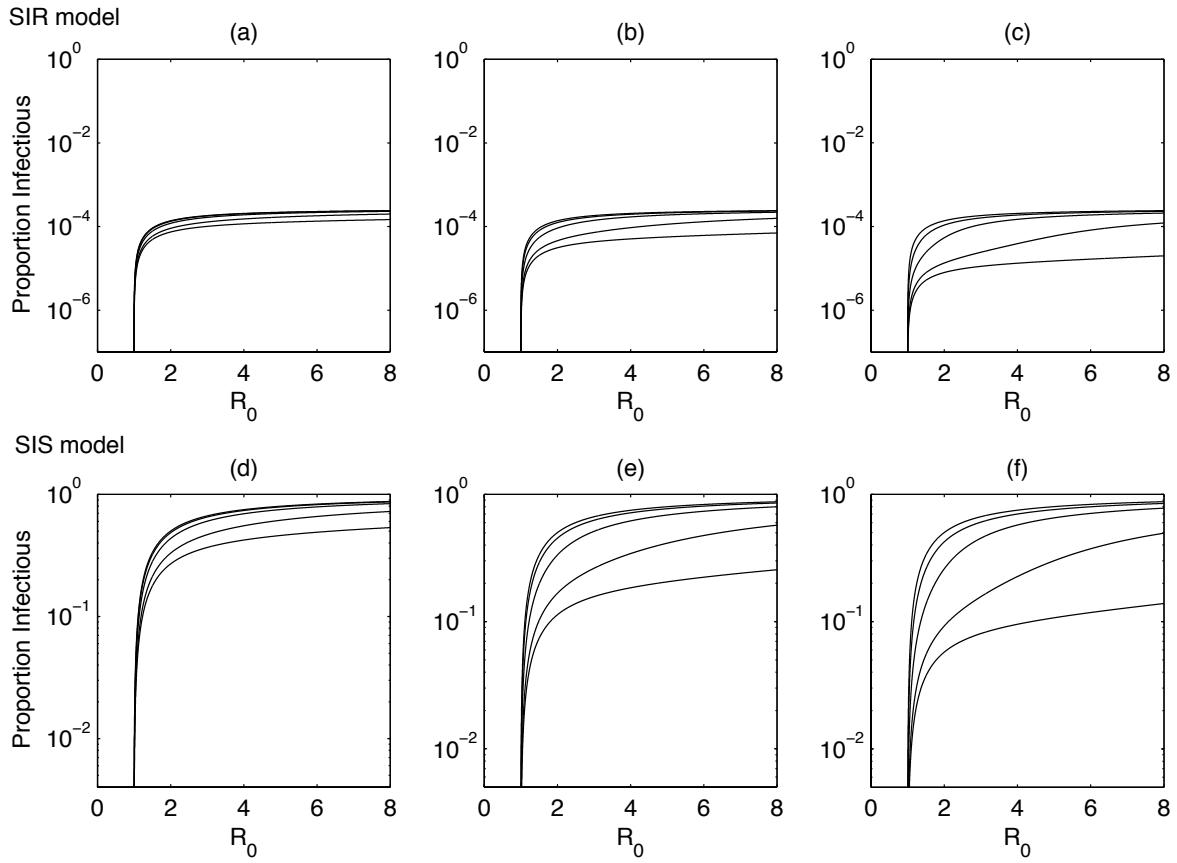


Figure 5.2: Prevalence of infection for the SIR and SIS models under different implementations of two risk groups (low and high). Top and bottom panels correspond to the SIR and SIS models, respectively: (a)–(c) $\sigma = 0$; (d)–(f) $\sigma = 1$. The three columns of panels correspond to different proportions of population at low risk: (a),(d) $\gamma = 0.5$; (b),(e) $\gamma = 0.8$; (c),(f) $\gamma = 0.95$. In each plot, different curves indicate the equilibrium prevalence of infection under different susceptibility ratios between the low-risk group and the average: $\alpha_1 = 1$, 0.75, 0.5, 0.2, 0.05, from the higher to the lower curves.

Let $\epsilon = \gamma$. In this case, $\dfrac{\partial a_0}{\partial \gamma} = 0$. So, $\dfrac{\partial I^0}{\partial \gamma} \leq 0$ iff $\dfrac{\partial a_2}{\partial \gamma} I^0 + \dfrac{\partial a_1}{\partial \gamma} < 0$. Now we replace I^0 by its expression in (5.6). Since $a_2 < 0$ and $\partial a_2/\partial \gamma < 0$ we get the following equivalent condition

$$-a_1 + 2a_2 \frac{\partial a_1}{\partial \gamma} \Big/ \frac{\partial a_2}{\partial \gamma} < \sqrt{a_1^2 - 4a_0 a_2}.$$

If the left-hand side is negative, then the condition is true and the result is proved. Otherwise, we can square both sides. In this case we get

$$\Big(-a_1 + 2a_2 \frac{\partial a_1}{\partial \gamma} \Big/ \frac{\partial a_2}{\partial \gamma} \Big)^2 - a_1^2 + 4a_0 a_2 = -4\alpha_1 R_0 (1-\gamma)(1-\alpha)(1-\gamma\alpha_1) < 0,$$

which ends this part of the proof.

Let $\epsilon = \alpha_1$. Now we have $\partial I^0/\partial\alpha_1 \geq 0$ iff

$$\frac{\partial a_2}{\partial \alpha_1} I^0 + \frac{\partial a_1}{\partial \alpha_1} > 0. \tag{5.13}$$

Again, we substitute I^0 by its expression in (5.6). Depending on the sign of $\partial a_2/\partial\alpha_1 (\neq 0)$ we obtain two different cases:

1. If $\dfrac{\partial a_2}{\partial \alpha_1} < 0$ then (5.13) is equivalent to $-a_1 + 2a_2 \dfrac{\partial a_1}{\partial \alpha_1} / \dfrac{\partial a_2}{\partial \alpha_1} > \sqrt{a_1^2 - 4a_0a_2}$ (a);

2. If $\dfrac{\partial a_2}{\partial \alpha_1} > 0$ then (5.13) is equivalent to $-a_1 + 2a_2 \dfrac{\partial a_1}{\partial \alpha_1} / \dfrac{\partial a_2}{\partial \alpha_1} < \sqrt{a_1^2 - 4a_0a_2}$ (b).

Note that if $\partial a_2/\partial\alpha_1 = 0$ then we must see if $\partial a_1/\partial\alpha_1 > 0$, which is true for $R_0 > 1$.

The sign of $\partial a_2/\partial\alpha_1$ is the same as the sign of $2\alpha_1\gamma - 1$. So, for case 1, let us assume $2\alpha_1\gamma < 1$. The left-hand side of (a) is

$$\frac{R_0(\tau+\mu)\mu[R_0\alpha_1((1-\alpha_1\gamma)(1-2\alpha_1\gamma)) + 1 - \alpha_1]}{(1-\gamma)(1-2\alpha_1\gamma)}$$

which is positive for $2\alpha_1\gamma > 1$. Hence, we can square both sides of (a) and we get

$$\Big(-a_1 + 2a_2 \frac{\partial a_1}{\partial \gamma} / \frac{\partial a_2}{\partial \gamma} \Big)^2 - a_1^2 + 4a_0a_2 = \frac{R_0^2(\tau+\mu)^2\mu^2 4\gamma\alpha_1}{(1-\gamma)^2(1-2\alpha_1\gamma)^2}(1-\alpha_1)(1-\alpha_1\gamma)$$
$$\times[R_0((1-2\alpha_1\gamma)^2 + \alpha_1(1-2\alpha_1\gamma)) + 1 - \alpha_1] > 0$$

which ends the proof of case 1.

In case 2 the left-hand side can change sign. Let us denote the left-hand side by B. So, to verify (b) we have to show that if B is positive then $B^2 - a_1^2 + 4a_0a_2 < 0$. By the calculations for the previous case we get that, for $2\alpha_1\gamma > 1$, $B > 0$ iff $C = R_0\alpha_1((1-\alpha_1\gamma)(1-2\alpha_1\gamma)) + 1 - \alpha_1 < 0$. Again, from the previous case we know that the sign of $B^2 - a_1^2 + 4a_0a_2$ is the same as the one of $R_0((1-2\alpha_1\gamma)^2 + \alpha_1(1-2\alpha_1\gamma)) + 1 - \alpha_1 = C/\alpha_1 + R_0(1-2\alpha_1\gamma)\alpha_1(1-\gamma) - (1-\alpha_1)^2/\alpha_1 < 0$, since $C < 0$. This ends the proof of case (b).

Finally, let $\epsilon = R_0$. In this case $\partial a_0/\partial R_0 \neq 0$. So, from (5.12) and since the denominator is non-positive, to prove that $\partial I^0/\partial R_0 \geq 0$ we need to prove that

$$\frac{\partial a_2}{\partial R_0} I^{0^2} + \frac{\partial a_1}{\partial R_0} I^0 + \frac{\partial a_0}{\partial R_0} > 0. \tag{5.14}$$

Now, taking into account that $I^{0^2} = -(a_1 I^0 + a_0)/a_2$ and using the expression of I_0 given in (5.6) we obtain

$$-Aa_1 + 2a_2B > A\sqrt{a_1^2 - 4a_0a_2},$$

where $A = \dfrac{\partial a_1}{\partial R_0}a_2 - \dfrac{\partial a_2}{\partial R_0}a_1$ and $B = a_2\dfrac{\partial a_0}{\partial R_0} - a_0\dfrac{\partial a_2}{\partial R_0}$. By substituting a_i and its derivatives in A we conclude that

$$A = -\alpha_1(1-\alpha_1\gamma)R_0^2(\tau+\mu)^3\mu(\alpha_1(1-\gamma) + (1-\alpha_1\gamma))/(1-\gamma)^2 < 0.$$

So, we can divide both sides by A and get

$$-a_1 + 2a_2B/A > \sqrt{a_1^2 - 4a_0a_2}. \tag{5.15}$$

If the left-hand side is negative, then the condition is true and the result is proved. Otherwise, we can square both sides of (5.15). Finally, we need to prove that the following expression $-4C(1-\alpha_1\gamma)\alpha_1\mu^2(\tau+\mu)^2R_0^2/(1-\gamma)(\alpha_1(1-\gamma)+(1-\alpha_1\gamma))^2$ is negative, where C is a polynomial in R_0 of degree 2 with coefficients $c_2 = \alpha_1(1-\alpha_1\gamma)(\alpha_1(1-\gamma)+\gamma(1-\alpha_1))$, $c_1 = 2\alpha_1(1-\alpha_1)(1-\alpha_1\gamma)(1-2\gamma)$ and $c_0 = (1-\alpha_1)^2$. We note that, the minimum value of C is attained at $R_0^m = 1-(1-\gamma)/(\alpha_1(1-\gamma)+\gamma(1-\alpha_1)) < 1$ and C is positive at R_0^m. As a consequence, since $c_2 > 0$ we conclude that C is positive for $R_0 > 1$. This concludes this part of the proof.

For $\sigma = 1$ the proofs follow easily from Remark 5.3.1. □

Previous studies based on the SIR framework have shown that heterogeneity in susceptibility to infection gives rise to smaller epidemics (Gart, 1968; Ball, 1985; Anderson & Britton, 1998). Here we find that disease prevalence at equilibrium is also lower in the presence of heterogeneity, and this is true for both SIR and SIS models. This effect is more pronounced the higher the variance in risk distribution.

5.3.2 Infection risk profiles

The profiles of the infection risk, within the susceptible and the recovered classes at endemic equilibrium, depend on assumptions on population heterogeneity and transmission intensity. For $\sigma = 0, 1$, we define the average risk factor among susceptible and recovered individuals as

$$\bar{\alpha}_S^\sigma = \frac{\alpha_1 S_1^* + \alpha_2 S_2^*}{S_1^* + S_2^*}, \quad \bar{\alpha}_R^\sigma = \frac{\alpha_1 R_1^* + \alpha_2 R_2^*}{R_1^* + R_2^*}, \tag{5.16}$$

where S_i^* and R_i^* are the susceptible and recovered individuals in each risk group, represented as proportions of the total population at endemic equilibrium. Figure 5.3 shows contour plots for the average risk factor among individuals never infected (S) and those infected and recovered at least once (R). Note, however, that these factors are further multiplied by λ and $\sigma\lambda$ to produce the average per capita rates of infection in S and R, respectively. This figure reflects how selection imposed by infection acts on the risk profiles.

In the SIR model, the average risk decreases as R_0 increases both for never-infected individuals and previously-infected individuals (Figure 5.3(a) and Figure 5.3(b), respectively). This selection mechanism underlies counter-intuitive trends that will emerge with the exploration of $\sigma \in (0,1)$ in Section 5.4, such as rates of reinfection decreasing with increasing R_0 and rates of reinfection appearing higher than rates of first infection even in the presence of partially protective immunity.

In the SIS model, selection maintains a large proportion of the high-risk group in the infected class and the mechanism is not entirely visible in the uninfected sub-population. Note that the average risk among never-infected individuals is roughly constant with R_0 (Figure 5.3(c)) while among previously infected individuals (Figure 5.3(d)) the average risk decreases with increasing R_0 as in the susceptible class of the SIR model (Figure 5.3(a)).

The properties observed for $\bar{\alpha}_S^0$ are summarized in the following theorem.

Theorem 5.3.4. *Let $R_0 > 1$. Then, for γ and $\alpha_1 \in (0,1)$*

$$\frac{\partial \bar{\alpha}_S^0}{\partial \gamma} \leq 0, \quad \frac{\partial \bar{\alpha}_S^0}{\partial \alpha_1} \geq 0 \tag{5.17}$$

$$\textit{and} \quad \frac{\partial \bar{\alpha}_S^0}{\partial R_0} < 0. \tag{5.18}$$

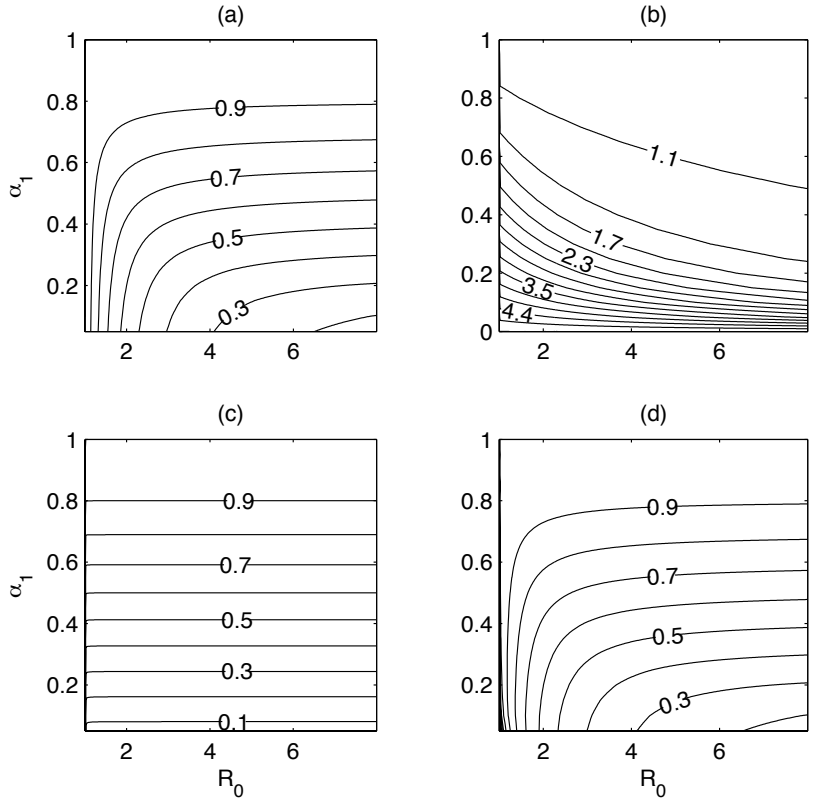


Figure 5.3: Average risk factor before and after infection. (a),(b) Contour plots for $\bar{\alpha}_S^0$, $\bar{\alpha}_R^0$ in the SIR model; (c),(d) contour plots for $\bar{\alpha}_S^1$, $\bar{\alpha}_R^1$ in the SIS model. Contours are represented in terms of the basic reproduction number, R_0, and the relative susceptibility of the low risk group, α_1. The proportion at low risk is $\gamma = 0.8$ in both cases.

Proof. First we derive an expression which relates the disease prevalence with the relative risk of the susceptible class in the case of the SIR model and the average risk of infection of the $S + R$ class in the case of the SIS system. Then we can use the results from the previous section to prove the theorem.

In the SIR case, from (5.4) letting $S = S_1 + S_2$, we obtain

$$\begin{cases} S' &= \mu - \beta I \bar{\alpha}_S^0 S - \mu S \\ I' &= \beta I \bar{\alpha}_S^0 S - (\tau + \mu) I. \end{cases} \tag{5.19}$$

Hence, we get an implicit expression for the disease prevalence at equilibrium in the case $\sigma = 0$

$$I^0 = \frac{\mu}{\tau + \mu}\big(1 - \frac{1}{\bar{\alpha}_S^0 R_0}\big). \tag{5.20}$$

Similarly, from system (5.5), we obtain the relation between the disease prevalence I^1 and the average risk of infection

$$I^1 = 1 - \frac{1}{\bar{\alpha}_{S+R}^1 R_0}. \tag{5.21}$$

Let $\sigma = 0$. From (5.20) we obtain the following expression for $\bar{\alpha}_S^0$

$$\bar{\alpha}_S^0 = \frac{\mu}{R_0(\mu - (\tau + \mu)I^0)}. \tag{5.22}$$

Thus, for $\epsilon = \gamma$ or α_1 we get

$$\frac{\partial \bar{\alpha}_S^0}{\partial \epsilon} = \frac{\mu(\tau + \mu)}{R_0[\mu - (\tau + \mu)I^0]^2} \frac{\partial I^0}{\partial \epsilon}, \tag{5.23}$$

which has the same sign as $\partial I^0/\partial \epsilon$.

For the derivative of $\bar{\alpha}_S^0$ with respect to R_0 we get

$$\frac{\partial \bar{\alpha}_S^0}{\partial R_0} = \frac{\mu}{[R_0(\mu - (\tau + \mu)I^0)]^2} \Big[\mu - (\tau + \mu)I^0 - R_0(\tau + \mu)\frac{\partial I^0}{\partial R_0}\Big]. \tag{5.24}$$

Hence, to prove that the derivative is negative is equivalent to prove that $I^0 + \partial I^0/\partial R_0 < \mu/(\tau + \mu)$. Now we substitute $(I^0)^2$ by $-(a_1 I^0 + a_0)/a_2$ and $\partial I^0/\partial R_0$ by the expression from (5.12). Furthermore, we replace I^0 by its expression in (5.6). So, taking into account that $a_2 < 0$ and

$$A = R_0 \frac{\partial a_2}{\partial R_0} a_1 - a_2 a_1 - \frac{\partial a_1}{\partial R_0} a_2 R_0 - 2a_2^2 \mu/(\tau + \mu) = -\mu(\tau + \mu)^3 \alpha_1^2 (1 - \alpha_1\gamma)^2 R_0^4/(1 - \gamma)^2 < 0$$

we get to the equivalent condition

$$-a_1 + 2a_2 B/A < \sqrt{a_1^2 - 4a_0 a_2}. \tag{5.25}$$

If the left-hand side is negative then the condition is true. Otherwise we can square both sides of (5.25). Hence $\big(-a_1 + 2a_2 B/A\big)^2 - a_1^2 + 4a_0 a_2 = -4R_0^2 \gamma \mu^2 (\tau + \mu)^2 (1 - \alpha_1)^2/(1 - \gamma) < 0$. which ends this part of the proof.

□

Remark 5.3.3. *In particular, from Remark 5.3.1 and by equality* (5.21) *we conclude that* $\bar{\alpha}_S^0 = \bar{\alpha}_{S+R}^1$.

The decrease on the average risk of infection of the susceptible class explains how prevalence decreases with population heterogeneity. In fact, the average force of infection, $\bar{\lambda}^0$, depends on the transmission intensity and on the average infection risk of the population subject to infection,

$$\bar{\lambda}^0 = \frac{\lambda\alpha_1 S_1^* + \lambda\alpha_2 S_2^*}{S_1^* + S_2^*} = \lambda^0 \bar{\alpha}_S^0 = \beta I^0 \bar{\alpha}_S^0 = \mu(R_0 \bar{\alpha}_S^0 - 1), \tag{5.26}$$

where we expressed I^0 as a function of $\bar{\alpha}_S^0$ according to formula (5.20). Directly from Theorem 5.3.4 it follows that heterogeneity decreases the force of infection, since

$$\frac{\partial \bar{\lambda}^0}{\partial \gamma} \leq 0, \quad \frac{\partial \bar{\lambda}^0}{\partial \alpha_1} \geq 0. \tag{5.27}$$

Remark 5.3.4. *As mentioned above, when $\sigma = 1$ we can identify our model with a SIS model through the identification of the class $S + R$ with the susceptible class of the SIS. It is then natural to investigate the effect of heterogeneity on prevalence by considering the dependence on the parameters of the average risk of infection of the $S + R$ class, $\bar{\alpha}^1_{S+R}$, defined as*

$$\bar{\alpha}^1_{S+R} := \frac{\alpha_1(S_1^* + R_1^*) + \alpha_2(S_2^* + R_2^*)}{(S_1^* + R_1^*) + (S_2^* + R_2^*)}$$

and of the corresponding force of infection, $\bar{\lambda}^1$, defined as $\bar{\lambda}^1 = \lambda^1 \bar{\alpha}^1_{S+R} = \beta I^1 \bar{\alpha}^1_{S+R}$. This is easily done, since from Remark 5.3.3 *and equation* (5.7) *we have, respectively, that $\bar{\alpha}^1_{S+R} = \bar{\alpha}^0_S$ and that $\bar{\lambda}^1 = \frac{\tau+\mu}{\mu}\bar{\lambda}^0$. As a consequence, $\bar{\lambda}^1$ satisfies the inequalities* (5.27)*, and we conclude that the decrease on the average risk of infection of the susceptible plus recovered class explains how prevalence decreases with population heterogeneity in the SIS model.*

Finally, as a side remark, we would like to note that with respect to the quantities defined in (5.16)*, it is $\bar{\alpha}^1_{S+R} = \dfrac{\bar{\alpha}^1_S S^* + \bar{\alpha}^1_R R^*}{S^* + R^*}$.*

Despite having the same infectivity, the risk groups contribute differently to the force of infection. Disease is more easily spread on the high-risk group due to its increased susceptibility, so the relative size of class I_2 is also greater. To further explore how the contribution of the high-risk group to the total proportion of infections changes with transmission intensity and heterogeneity, we define the quotient $Q^\sigma = I_2^\sigma / I^\sigma$ at equilibrium. For $\sigma = 0$ or $\sigma = 1$, the contribution of the high-risk group, Q^σ, decreases as transmissibility increases and it is greater when the high-risk group is larger (γ close to 0) or when its relative risk of infection is further from the population average (α_1 close to 0). The following theorem summarizes these results.

Theorem 5.3.5. *Let $R_0 > 1$. Then, for γ and $\alpha_1 \in (0, 1)$*

$$\frac{\partial Q^\sigma}{\partial \gamma} \leq 0, \quad \frac{\partial Q^\sigma}{\partial \alpha_1} \leq 0 \tag{5.28}$$

$$\text{and} \quad \frac{\partial Q^\sigma}{\partial R_0} < 0 \quad \sigma = 0, 1. \tag{5.29}$$

Proof. As for the previous proofs, we start by studying the case $\sigma = 0$ and then the case $\sigma = 1$ follows directly from Remark 5.3.1. In fact, in this case we have $Q^0 = Q^1$.

For simplicity we write $Q^0 = I_2^0/I^0$ as $1 - I_1^0/I^0 = 1 - R_0\alpha_1\mu\gamma/(R_0(\tau+\mu)I^0\alpha_1 + \mu)$. The derivative of Q^0 with respect to γ is

$$\frac{\partial Q^0}{\partial \gamma} = -\frac{R_0\alpha_1\mu[\mu + R_0\alpha_1(\tau+\mu)\big(I^0 - \gamma\frac{\partial I^0}{\partial \gamma}\big)]}{(R_0(\tau+\mu)I^0\alpha_1 + \mu)^2}. \tag{5.30}$$

But $I^0 - \gamma\dfrac{\partial I^0}{\partial \gamma} = I^0\Big(1 + \gamma\dfrac{\frac{\partial a_2}{\partial \gamma}I^0 + \frac{\partial a_1}{\partial \gamma}}{2a_2I^0 + a_1}\Big) \geq 0$ from what was seen in the Proof of Theorem 5.3.3. Thus we conclude that $\frac{\partial Q^0}{\partial \gamma} \leq 0$.

The derivative of Q^0 with respect to α_1 is

$$\frac{\partial Q^0}{\partial \alpha_1} = -\frac{R_0\gamma\mu[\mu - R_0\alpha_1^2(\tau+\mu)\frac{\partial I^0}{\partial \alpha_1}]}{(R_0(\tau+\mu)I^0\alpha_1 + \mu)^2}. \tag{5.31}$$

This expression has the opposite sign of $C = \mu - R_0\alpha_1^2(\tau+\mu)\frac{\partial I^0}{\partial \alpha_1}$. Again, we replace ${I^0}^2$ by $-(a_1 I^0 + a_0)/a_2$ and then I^0 by its expression in (5.6). Finally, we conclude that $C \geq 0$ iff $-a_1 A + 2a_2 B \leq A\sqrt{a_1^2 + 4a_2 a_0}$, where $A = 2\mu a_2^2 - R_0\alpha_1^2(\tau+\mu)\frac{\partial a_2}{\partial \alpha_1}a_1 + R_0\alpha_1^2(\tau+\mu)\frac{\partial a_1}{\partial \alpha_1}a_2$ and $B = a_1 a_2 \mu - R_0\alpha_1^2(\tau+\mu)\frac{\partial a_2}{\partial \alpha_1}a_0$. Note that by substituting a_i and its derivatives in A we can easily conclude that A is positive. Therefore, we can divide both sides by A, obtaining that $C \geq 0$ iff

$$-a_1 + 2a_2 B/A \leq \sqrt{a_1^2 + 4a_2 a_0}. \tag{5.32}$$

If the left-hand side is negative, then the condition is verified. Otherwise we can square both sides of (5.32). Hence, we get $-\frac{4(1-\alpha_1\gamma)^2\gamma\alpha_1^2\mu^2(\tau+\mu)^2R_0^4}{(1-\gamma)} \leq 0$. This implies that $C \geq 0$ or, equivalently, that $\partial Q^0/\partial\alpha_1 \leq 0$, which ends this part of the proof.

The derivative of Q^0 with respect to R_0 is

$$\frac{\partial Q^0}{\partial R_0} = \frac{\alpha_1\gamma\mu[\mu - R_0^2\alpha_1(\tau+\mu)\frac{\partial I^0}{\partial R_0}]}{(R_0(\tau+\mu)I^0\alpha_1 + \mu)^2}. \tag{5.33}$$

This expression has the same sign of $C = b - R_0^2\frac{\partial I^0}{\partial R_0}$, where $b = \mu/(\alpha_1(\tau+\mu))$. We conclude that $C > 0$ iff $-a_1 A + 2a_2 B > A\sqrt{a_1^2 + 4a_2 a_0}$, where $A = \frac{\partial a_2}{\partial R_0}a_1 - \frac{\partial a_1}{\partial R_0}a_2 - 2a_2^2 b'$, $B = a_0\frac{\partial a_2}{\partial R_0} - \frac{\partial a_0}{\partial R_0}a_2 - b' a_1 a_2$ and $b' = b a_2/R_0^2$. Note that by substituting a_i and its derivatives in A we can easily conclude that A is negative. So, we can divide both sides by A, obtaining that $C > 0$ iff

$$-a_1 + 2a_2 B/A < \sqrt{a_1^2 + 4a_2 a_0}. \tag{5.34}$$

If the left-hand side is negative the condition is verified. Otherwise we can square both sides of (5.34). Hence, we get $-\frac{4(1-\alpha_1\gamma)^2\gamma\alpha_1^2\mu^2(\tau+\mu)^2R_0^4}{(1-\gamma)} < 0$. This ends the proof. □

Overall, the contribution of the high-risk group can vary from α_2 times its relative size, near the epidemic threshold, to its relative size, for sufficiently high transmission. This can have important consequences for the effectiveness of interventions, specially in low endemic regions where the groups with increased risk have more impact. We will focus more on this aspect when studying the SIRI model.

5.4 The SIRI model

5.4.1 Thresholds in Transmission

Here we consider the effect of heterogeneity in the intermediate scenario where infection induces partial immunity. It is assumed that individuals are protected while infected but regain some susceptibility upon recovery. Susceptibility to reinfection is reduced by a factor $\sigma \in [0,1]$, compared to susceptibility to first infection. Endemic equilibria and infection risk profiles have been analyzed for the limiting cases $\sigma = 0, 1$ (corresponding to SIR, SIS models) in Section 5.3. In both cases, disease persistence is determined by the threshold condition, $R_0 > 1$, irrespective of population structure, sustaining levels of infection that are generally much higher in the SIS scenario due to reinfection. For the SIRI model, R_0

remains a threshold parameter and it corresponds to a shift in stability from the disease-free equilibrium to the endemic one. In the following theorem we state the result relative to the stability of the disease-free equilibrium.

Theorem 5.4.1. *For $\sigma \in [0,1]$, the disease-free equilibrium, E_0^σ, of system* (5.1) *is globally asymptotically stable if $R_0 < 1$ and it is unstable for $R_0 > 1$.*

Proof. In theorem 5.3.1 we proved this result in the cases $\sigma = 0$ and $\sigma = 1$. So, assume $\sigma \in (0,1)$ and $n = 2$. Following van den Driessche & Watmough (2002), first we must distinguish new infections from all other class transitions in population. The infected classes are I_1 and I_2, so we can write system (3.1) as

$$\dot{X} = f(X) \Leftrightarrow \dot{X} = \mathcal{F}(X) - \mathcal{V}(X) = \mathcal{F}(X) - (\mathcal{V}^-(X) - \mathcal{V}^+(X)), \tag{5.35}$$

where $X = (I_1, I_2, S_1, R_1, S_2, R_2)$, $\mathcal{F}$ is the rate of appearance of new infections in each class; $\mathcal{V}^+$ is the rate of transfer into each class by all other means and $\mathcal{V}^-$ is the rate of transfer out of each class. Hence, $\mathcal{F} = (\beta(I_1+I_2)\alpha_1 S_1, \beta(I_1+I_2)\alpha_1 S_2, 0, 0, 0, 0)^T$, and the disease-free equilibrium is $X_0 = (0, 0, \gamma, 0, 1-\gamma, 0)$.

By theorem 2 in van den Driessche & Watmough (2002) it is sufficient to prove conditions (A1)-(A5). The verification of (A1)-(A4) is straightforward. To show that condition (A5) is verified, we write the Jacobian of f at X_0 with $\mathcal{F}$ set to zero, as

$$Df_{(\mathcal{F}=0)}(X_0) = \begin{bmatrix} -(\tau+\mu) & 0 & 0 & 0 & 0 & 0 \\ 0 & -(\tau+\mu) & 0 & 0 & 0 & 0 \\ -\alpha_1\beta\gamma & -\alpha_1\beta\gamma & -\mu & 0 & 0 & 0 \\ \tau & 0 & 0 & -\mu & 0 & 0 \\ -\alpha_2\beta(1-\gamma) & -\alpha_2\beta(1-\gamma) & 0 & 0 & -\mu & 0 \\ 0 & \tau & 0 & 0 & 0 & -\mu \end{bmatrix}.$$

Since the eigenvalues, $-\mu$ and $-(\tau+\mu)$, are real and negative the result follows. □

Remark 5.4.1. *The endemic equilibrium is numerically computed as a root of a fourth order polynomial on $I = I_1 + I_2$. Analytical expression of the polynomial can be obtained but, due to its complexity, we were unable to show the existence of a positive real root.*

In the intermediate case, another threshold is identified, similar to what happens in the homogeneous case for $R_0 = 1/\sigma$, which marks a transition from SIR- to SIS-like behavior (Section 2.4.1). We show that the same expression holds for the reinfection threshold in the presence of heterogeneity in susceptibility to infection. Following the method developed in Section 2.5, we first have to define the reinfection sub-model that corresponds to the limit situation where all individuals that enter in the system are partially immunized and only subjected to reinfection. Hence, the reinfection sub-model for system (5.1) with $n = 2$, has only four classes and can be represented by the following system of differential equations:

$$\begin{cases} R_1' &= \mu\gamma + \tau I_1 - \sigma\lambda\alpha_1 R_1 - \mu R_1, \\ R_2' &= \mu(1-\gamma) + \tau I_2 - \sigma\lambda\alpha_2 R_2 - \mu R_2, \\ I_1' &= \sigma\lambda\alpha_1 R_1 - (\tau+\mu) I_1 \\ I_2' &= \sigma\lambda\alpha_2 R_2 - (\tau+\mu) I_2 \end{cases} \tag{5.36}$$

The reinfection sub-model has a unique disease-free equilibrium $E_0 = (\gamma, 1-\gamma, 0, 0)$. Analyzing the Jacobian at E_0, we conclude that a bifurcation on the transmission parameter β occurs at $\beta = (\tau + \mu)/\sigma$, when the disease-free equilibrium changes its stability. In terms of the basic reproduction number the bifurcation is attained at $R_0 = 1/\sigma$.

5.4.2 Endemic equilibrium

Quantitative discrepancies between epidemiological data and model results have been reported previously and generally attributed to case sub-notification or population heterogeneity not captured by simple models (Fine & Clarkson (1982), Anderson & May (1985), van Boven *at al.* (2001)). Systematic investigations of these factors are expected to provide valuable insights with wide application in infectious disease epidemiology. In Section 5.3 we have shown that heterogeneity in susceptibility to infection reduces prevalence of infection in SIR and SIS models and here we extend this conclusion to the general SIRI framework. Figure 5.4

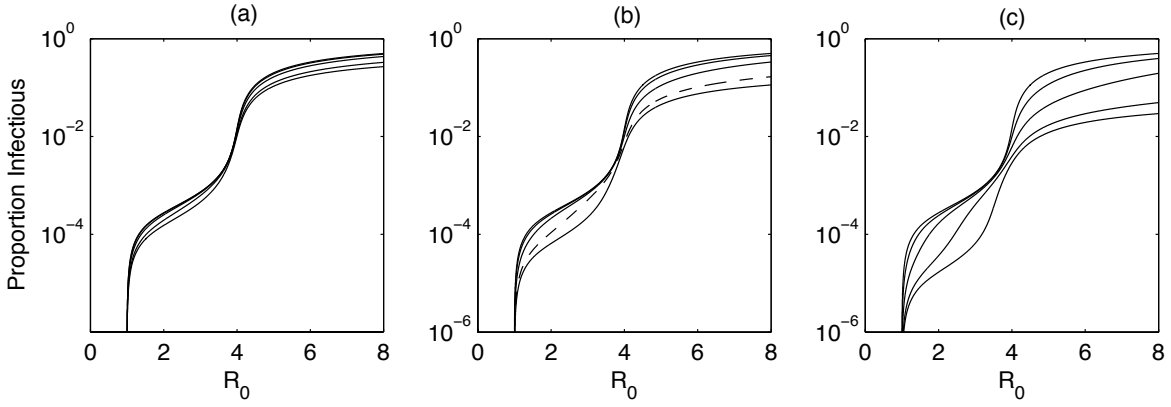


Figure 5.4: Prevalence of infection for the SIRI model under different implementations of two risk groups (low and high). The three panels correspond to different proportions of the population at risk: (a),(d) $\gamma = 0.5$; (b),(e) $\gamma = 0.8$; (c),(f) $\gamma = 0.95$. In each plot, different curves indicate the equilibrium prevalence of infection under different susceptibility ratios between the low-risk group and the average: $\alpha_1 = 1$, 0.75, 0.5, 0.2, 0.05, from the higher to the lower curves.

shows the endemic equilibrium for different infection risk profiles of the population. When heterogeneity is considered the disease prevalence is lower than in the homogeneous case, and this effect is more pronounced when the variance, var_α, is higher (high γ and low α_1). These trends are observed for $0 \leq \sigma \leq 1$, including the particular cases $\sigma = 0, 1$, analyzed previously.

Remark 5.4.2. *For the SIRI model, it was not possible to obtain detailed proofs of the main results. In the next sections, we present numerical results as illustrations of a more exhaustive numerical investigation, performed using* MATLAB 6.5® *software. We were unable to adapt the argument used to determine the sign of the denominator of formula* (5.12) *for the SIRI model, which was central to proof the remanning results.*

5.4.3 Infection risk profiles

The reduction in disease prevalence is associated with the changes in the infection risk profile imposed by transmission on both susceptible and recovered classes. In this section we analyze how the average infection risk of susceptible and recovered individuals change with R_0 and heterogeneity (represented by the proportion of the population with low risk, γ, and risk of these individuals relative to the average, α_1). We remark that when $\sigma \in (0, 1)$ the average risk factors in the susceptible and recovered classes, $\bar{\alpha}_S$ and $\bar{\alpha}_R$ respectively, are defined as in (5.16).

Figure 5.5 illustrates the average risk factor for susceptible and recovered classes for $\gamma = 0.8$ and $\sigma = 0.25$, by means of contour plots in the parameter space of transmissibility, R_0, and heterogeneity, α_1. Generally, the average risk among susceptible and recovered individuals decreases as R_0 increases (Figure 5.5(a) and (b) respectively). The reinfection threshold, indicated by vertical dotted lines, marks the shift from SIR to SIS regime. It is associated with a saturation of the trend observed for the susceptible ($\bar{\alpha}_S$ appears constant for R_0 above threshold as for $\sigma = 1$) and an average risk equal to one in the recovered class ($\bar{\alpha}_R = 1$). Compare with Figure 5.3. Overall, we have two equilibrium regimes. Below the

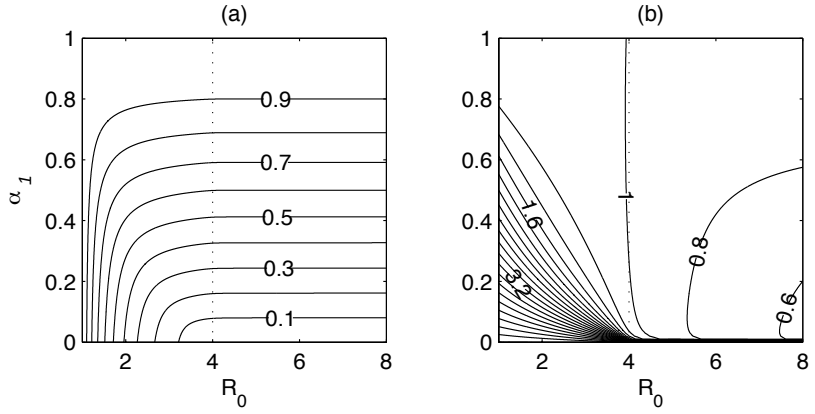


Figure 5.5: Average risk factor for susceptible and recovered classes in the SIRI model. Contour plots for: (a) the susceptible class, $\bar{\alpha}_S^\sigma$; and (b) the recovered class, $\bar{\alpha}_R^\sigma$. Contours are represented in terms of the basic reproduction number, R_0, and the relative susceptibility of the low risk group, α_1. The proportion at low risk is $\gamma = 0.8$ and susceptibility reduction due to partial immunity is $\sigma = 0.25$.

reinfection threshold, the uninfected population is composed of many susceptible individuals with an average risk factor below one, and few recovered individuals with high risk due to selection imposed by infection. Above the reinfection threshold, most individuals have already experienced at least one infection and are still susceptible to reinfection but have an average risk factor below one. In the latter case, selection maintains a large proportion of the population in the infected class.

The patterns described for susceptible and recovered risk profiles have strong implications for the interpretation of disease dynamics, notably the contribution of reinfection to the overall disease incidence. We define the incidence of first infection and the incidence of reinfection,

in the respective populations at risk, as

$$Y_1 = \frac{\lambda(\alpha_1 S_1^* + \alpha_2 S_2^*)}{S^*} = \lambda \bar{\alpha}_S, \tag{5.37}$$

$$Y_2 = \frac{\sigma\lambda(\alpha_1 R_1^* + \alpha_2 R_2^*)}{R^*} = \sigma\lambda \bar{\alpha}_R. \tag{5.38}$$

The total incidence in the entire uninfected population is then calculated as

$$Y_{total} = \frac{Y_1 S^* + Y_2 R^*}{S^* + R^*}.$$

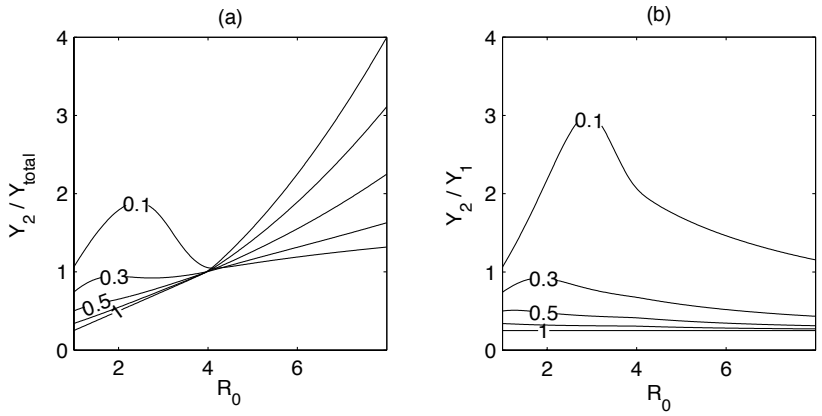


Figure 5.6: Intensity of reinfection. (a) Rate of reinfection among recovered individuals over the total incidence. (b) Rate of reinfection among recovered individuals over the incidence of first infection. Different values are considered for the relative risk of the low risk group, α_1, including the homogeneous case, $\alpha_1 = 1$. The proportion of the population at low risk, γ, is fixed at 0.8.

Figure 5.6 shows that despite reinfection being hindered by heterogeneity, the rate of reinfection among recovered individuals, Y_2, can be higher than overall rate of infection in the entire uninfected population, Y_{total}. We see that, for the homogeneous case ($\alpha_1 = 1$), the quotient, Y_2/Y_{total}, increases monotonically with R_0, and for $R_0 > 1/\sigma$ it is above one. For the heterogeneous case, reinfection among the recovered class can be higher than disease incidence also below the reinfection threshold. Even for low endemic populations, where the contribution of reinfection is low, it is possible that recovered individuals, as a group, show a higher risk of reinfection than expected when assuming partial immunity. This can have major implications for the interpretation of epidemiological data. In particular, overlooking host heterogeneity may lead to misleading expectations for the effectiveness of control measures.

5.4.4 Contribution of the high-risk group

As we have observed before, for the SIR and SIS models, the contribution of the high-risk group decreases when transmission increases (Theorem 5.3.5), due to the accumulation of high-risk individuals in the recovered or infectious class, respectively. For the SIRI model the contribution of the high-risk group is expected to be greater than in the previous cases, due

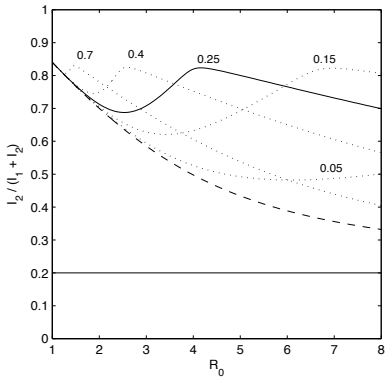


Figure 5.7: Contribution of the high-risk group to the total disease prevalence. Infection risk distribution is given by $\gamma = 0.8$ and $\alpha_1 = 0.2$. The dashed line corresponds to both $\sigma = 0$ and $\sigma = 1$, while other lines correspond to values of σ as indicated. The special case $\sigma = 0.25$ is distinguished as a full line as this parameter values is used for illustration in other figures. The horizontal line represents the case for which the infection risk distribution is homogeneous ($\alpha_1 = \alpha_2 = 1$).

to the additional contribution of reinfection, for intermediate levels of transmission, where the high-risk group is still overrepresented in the recovered class (α_R close to one).

Figure 5.7 shows the contribution of the high-risk group to the total disease prevalence for the particular case $\gamma = 0.8$ and $\alpha_1 = 0.2$. This corresponds to a risk group of 20% of the total population with an increased risk of infection α_2 equal to 4.2 times that of the total population, and 21 times that of the low-risk group. Moreover, for this choice of parameters and for $\sigma = 0.25$, disease prevalence corresponds to about 30% of the homogeneous model prediction as represented by the dashed line in Figure 5.4(b). Here a sub-population of 20% accounts from 70% to 85% of the infection, depending on the intensity of transmission. The contribution of the high-risk group is stronger near the endemic and reinfection thresholds. Near the thresholds the classes that are susceptible to infection and reinfection, S and R, respectively, reach their maximum capacity, accounting for almost the entire population. Therefore, the average risk on these classes and the selection pressure on the high-risk group are maximum.

When considering heterogeneous infectivity, theoretical work and different field studies have suggested that roughly 20% of the infectious individuals can be responsible for 80% of transmission (Galvani & May, 2005; Woolhouse *et al.*, 1997). This 20/80 rule has important consequences for disease control (Woolhouse *et al.*, 1997). Here we obtain similar effects by assuming heterogeneity in susceptibility to infection as previously estimated for the case of malaria transmission, where 20% of people receive 80% of all infections (Smith *at al.*, 2005) due to heterogeneity in mosquito biting or in susceptibility to infection.

5.5 Interventions

5.5.1 Targeted vaccination

The greater impact of the high-risk group on transmission should be taken into account when planning interventions for disease control. In this section, we compare uniform and

targeted vaccination strategies. Comparison is made on the basis of the vaccination coverage required, under different strategies, to obtain the same impact. We implement vaccination at birth assuming that the protection conferred by the vaccine is equivalent to that of natural immunity. Vaccination reduces the risk of infection but the relative susceptibility of the two risk groups is maintained. This is formalised by the following system

$$\begin{cases} S_i' &= (1-v_i)\mu\gamma_i - \lambda\alpha_i S_i - \mu S_i \\ I_i' &= \lambda\alpha_i S_i + \sigma\lambda\alpha_i R_i - (\tau+\mu) I_i \\ R_i' &= v_i\mu\gamma_i + \tau I_i - \sigma\lambda\alpha_i R_i - \mu R_i, \quad i=1,2, \end{cases} \tag{5.39}$$

where we assume that vaccinated individuals born directly into the R class. The vaccination model has a disease-free equilibrium $E_0 = (v_1\gamma_1, 0, (1-v_1)\gamma_1, v_2\gamma_2, 0, (1-v_2)\gamma_2$ and it undergoes a bifurcation at $\beta = \dfrac{\tau+\mu}{(1-v_1+\sigma v_1)\gamma\alpha_1 + ((1-v_2)+\sigma v_2)(1-\gamma)}$, when the determinant of the Jacobian matrix

$$\begin{bmatrix} -\mu & -\alpha_1\beta(1-v_1)\gamma_1 & 0 & 0 & -\alpha_1\beta(1-v_1)\gamma_1 & 0 \\ 0 & \alpha_1\beta\gamma_1((1-v_1)+\sigma v_1) - (\tau+\mu) & 0 & 0 & \alpha_1\beta\gamma_1((1-v_1)+\sigma v_1) & 0 \\ 0 & \tau - \sigma\alpha_1\beta v_1\gamma_1 & -\mu & 0 & -\sigma\alpha_1\beta v_1\gamma_1 & 0 \\ 0 & -\alpha_2\beta(1-v_2)\gamma_2 & 0 & -\mu & -\alpha_2\beta(1-v_2)\gamma_2 & 0 \\ 0 & \alpha_2\beta\gamma_2((1-v_2)+\sigma v_2) & 0 & 0 & \alpha_2\beta\gamma_2((1-v_2)+\sigma v_2) - (\tau+\mu) & 0 \\ 0 & -\sigma\alpha_2\beta v_2\gamma_2 & 0 & 0 & \tau - \sigma\alpha_2\beta v_2\gamma_2 & -\mu \end{bmatrix}$$

evaluated at E_0 is zero, corresponding to a change in stability of E_0. This marks a transmission threshold below which disease is eliminated by vaccination, that we refer to as the vaccination threshold. With respect to the basic reproduction number (5.3) of the original system we can say that the vaccination threshold is achieved at

$$R_0 = \frac{1}{((1-v_1)+\sigma v_1)\gamma\alpha_1 + ((1-v_2)+\sigma v_2)(1-\gamma)\alpha_2}. \tag{5.40}$$

The vaccination threshold is always to the right of the epidemic threshold $R_0 = 1$, as can be confirmed by inspection of (5.40).

First, we consider a strategy based on a limited quantity of vaccines corresponding to a given coverage, v. We can then vary the percentage of each risk group covered by the program by fixing $v = v_1\gamma_1 + v_2\gamma_2 = v_1\gamma + v_2(1-\gamma)$ and varying v_2. Naturally, increasing the representation of the high-risk group in the vaccinated sub-population will increase the impact of the program by shifting the vaccination threshold to the right (Britton, 1998). In fact, by substituting $v_1 = \dfrac{v - v_2(1-\gamma)}{\gamma}$ in equation (5.40) and assuming that it defines R_0 as a function of the vaccination coverage for the high-risk group v_2, we obtain

$$\frac{\partial R_0}{\partial v_2} = \frac{(1-\alpha_1)(1-\sigma)}{[((1-v_1)+\sigma v_1)\gamma\alpha_1 + ((1-v_2)+\sigma v_2)(1-\gamma)\alpha_2]^2} > 0 \tag{5.41}$$

Now, we reverse the argument and inspect what coverage we need to attain with a targeted strategy in order to achieve the same effectiveness as the corresponding uniform strategy ($v_1 = v_2$). This will provide an estimation for how many doses we save by targeting the vaccination program to those individuals at higher risk, as a so called *top-to-bottom* strategy (Britton, 1998). Figure 5.8 illustrates the saving associated with targeting. For Figure 5.8(a)

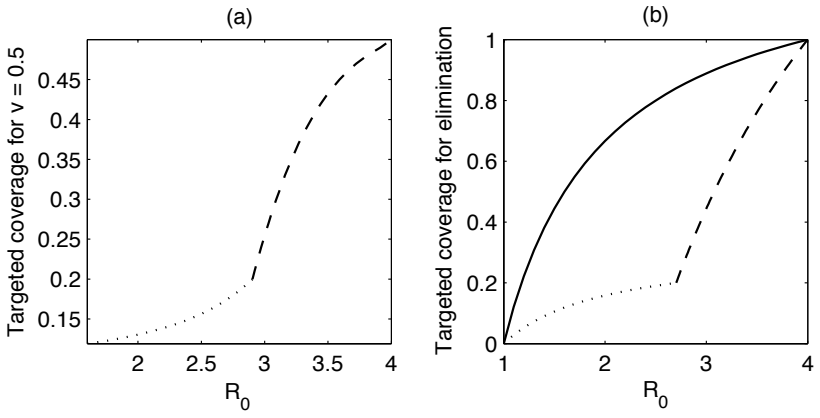


Figure 5.8: Uniform *vs.* targeted vaccination programs. **(a)** Vaccination coverage required for the targeted strategy to obtain the same disease reduction as a uniform strategy with 50% coverage, for each R_0. **(b)** Vaccination coverage required to eliminate the infection for each R_0. Full lines corresponds to the uniform strategy (v_u^*) and broken lines correspond to the targeted strategy (v_t^*), dotted if vaccination is restricted to the high-risk group and dashed if this is complemented by vaccination in the low-risk group. Infection risk distribution is given by $\gamma = 0.8$ and $\alpha_1 = 0.2$.

we use as a reference the reduction in disease prevalence achieved with a uniform vaccination strategy with coverage $v_1 = v_2 = 0.5$. The figure shows the coverage needed for a targeted strategy to achieve the same reduction in disease prevalence as for the uniform vaccination strategy. This result was obtained numerically, by computing, for each β, the vaccination coverage for the targeted strategy v for which the prevalence curves under each vaccination strategy intersect at β. Below the reinfection threshold ($R_0 = 4$) it is always possible to achieve the same reduction using a targeted strategy with lower coverage, while above the reinfection threshold there is no difference (result not shown). Note that this is achieved by vaccinating only a proportion of the high-risk group, if R_0 is low enough and for $v < \gamma_2$ (dotted line in the figure) or by vaccinating completely the high-risk group and a proportion of the low-risk group (dashed lines in the figure).

Figure 5.8(b) represents elimination coverages under different strategies. The critical vaccination coverage to eliminate the infection, for a given R_0, for the uniform strategy (full line) is obtained by substituting $v_1 = v_2 = v$ in equation (5.40) and solving for v:

$$v_u^* = \frac{1 - 1/R_0}{1 - \sigma}. \tag{5.42}$$

For the *top-to-bottom* strategy, we substitute $v_1 = 0$ and $v_2 = v/(1-\gamma)$, for $v \leq 1 - \gamma$, or $v_2 = 1$ and $v_1 = (v - (1-\gamma))/\gamma$, if $v > 1 - \gamma$, in equation (5.40). Solving for v we obtain

$$v_t^* = \frac{(1 - 1/R_0)(1-\gamma)}{(1-\sigma)(1-\gamma\alpha_1)} \text{ or } \; v_t^* = \frac{1 - 1/R_0}{(1-\sigma)\alpha_1} - \frac{(1-\alpha_1)}{\alpha_1}, \tag{5.43}$$

with vaccination of only the high-risk group (dotted line) or both groups (dashed line), respectively. Below the reinfection threshold, the elimination coverage is always lower for the targeted strategy. Above the reinfection threshold, it is impossible to interrupt transmission

and eliminate the infection. In fact, one can easily show that if $R_0 > 1/\sigma$ both v_u^* and v_t^* become greater than one. As previously discussed (Gomes *et al.*, 2004b, 2005) above this threshold only a superior vaccine, capable of inducing an immune response more effective than natural infection, would be efficacious.

5.5.2 Controlling risk profile

We have just described how targeting strategies have the potential to improve vaccination impact. However, this fails for populations above the reinfection threshold. In this section, we analyze the effect of interventions intended to convert high-risk individuals into low-risk. We will see that these interventions that have the potential to reduce disease endemic state even above the reinfection threshold.

We assume that it is possible to reduce the risk of infection of part of the individuals in the high-risk group, either by biomedical or behavioral interventions. We assume that these targeted interventions act on the high-risk individuals by transferring them to the low-risk group, at a rate ϕ. The new model is described by the following set of differential equations

$$\begin{cases} S_1' &= \mu\gamma + \phi S_2 - \lambda\alpha_1 S_1 - \mu S_1 \\ I_1' &= \lambda\alpha_1 S_1 + a\phi I_2 + \sigma\lambda\alpha_1 R_1 - (\tau+\mu) I_1 \\ R_1' &= \tau I_1 + a\phi R_2 - \sigma\lambda\alpha_1 R_1 - \mu R_1 \\ S_2' &= \mu(1-\gamma) - \lambda\alpha_2 S_2 - (\mu+\phi) S_2 \\ I_2' &= \lambda\alpha_2 S_2 + \sigma\lambda\alpha_2 R_2 - (\tau+\mu+a\phi) I_2 \\ R_2' &= \tau I_2 - \sigma\lambda\alpha_2 R_2 - (\mu+a\phi) R_2 \end{cases} \tag{5.44}$$

Remark 5.5.1. *The model with intervention represented by the system* (5.44) *coincides with the original one, system* (5.1)*, when no intervention is implemented ($\phi = 0$) and in the homogenous case ($\alpha_1 = 1$).*

We compare two possible interventions. First, we assume an early intervention, where only the susceptible individuals are able to reduce the infection risk ($a = 0$). Improvement of social conditions for children or immune protection conferred by breast-feeding, can serve as an example. Second, we expand the intervention to all individuals in the high-risk group ($a = 1$). This can be the case of interventions that focus on behavioral changes or treatment follow-up of high-risk individuals.

Region of possible disease elimination.

The possibility to eliminate the disease for a certain population corresponds to the existence and stability of a disease-free equilibrium for a range of transmission intensities, given by R_0. For both early or extended interventions ($a = 0$ or $a = 1$), system (5.44) has a disease-free equilibrium of the form $E_0 = ((\gamma\mu+\phi)/(\mu+\phi), 0, 0, \mu(1-\gamma)/(\mu+\phi), 0, 0)$. For this equilibrium, system (5.44) undergoes a bifurcation in the transmission parameter at $\beta = (\tau+\mu)(\mu+\phi)/(\mu+\alpha_1\phi)$, which corresponds to a shift in the equilibrium stability summarized in the theorem below.

Remark 5.5.2. *The bifurcation point for system* (5.44) *for the disease-free equilibrium corresponds to*

$$R_0 = \frac{\mu+\phi}{\mu+\alpha_1\phi} = ET_\phi, \tag{5.45}$$

where R_0 is the basic reproduction number (5.3) *for the original model* (5.1)*. Moreover, this bifurcation is always to the right of R_0 since $(\mu+\phi)/(\mu+\alpha_1\phi) > 1$, for $\phi > 0$.*

Theorem 5.5.1. *For $\sigma \in (0,1)$, the disease-free equilibrium of system* (5.44)*, E_0, is globally asymptotically stable if $R_0 < ET_\phi$ and it is unstable for $R_0 > ET_\phi$.*

Proof. Lets compute the Jacobian for system (5.44) at E_0

$$\begin{bmatrix} -\mu & -\alpha_1\beta\frac{\gamma\mu+\phi}{\mu+\phi} & 0 & \phi & -\alpha_1\beta\frac{\gamma\mu+\phi}{\mu+\phi} & 0 \\ 0 & \alpha_1\beta\frac{\gamma\mu+\phi}{\mu+\phi} - (\tau+\mu) & 0 & 0 & \alpha_1\beta\frac{\gamma\mu+\phi}{\mu+\phi} + a\phi & 0 \\ 0 & \tau & -\mu & 0 & 0 & a\phi \\ 0 & -\alpha_2\beta\frac{(1-\gamma)\mu}{\mu+\phi} & 0 & -(\mu+\phi) & -\alpha_2\beta\frac{(1-\gamma)\mu}{\mu+\phi} & 0 \\ 0 & \alpha_2\beta\frac{(1-\gamma)\mu}{\mu+\phi} & 0 & 0 & \alpha_2\beta\frac{(1-\gamma)\mu}{\mu+\phi} - (\tau+\mu) - a\phi & 0 \\ 0 & 0 & 0 & 0 & \tau & -(\mu+a\phi) \end{bmatrix}.$$

The eignvalues are $-\mu$, $-(\mu+\phi)$, $-(\mu+a\phi)$, $-(a\phi+\mu+\tau)$ and $\beta(\alpha_1\phi+\mu)/(\mu+\phi)-(\tau+\mu)$. Hence, for $R_0 < ET_\phi$ i.e. for $\beta < (\mu+\phi)(\tau+\mu)/(\alpha_1\phi+\mu)$ all eigenvalues are real and positive and the result follows. □

Early intervention (a=0).

By changing the risk profile of individuals at risk for reinfection, early interventions on susceptible high-risk individuals can give rise to bistable situations, as illustrated in Figure 5.9 and 5.10. Bistable behavior indicates that these interventions have the potential to significantly reduce the disease, even above the reinfection threshold. However, their effect will be negligenciable unless the intervention effort is above a certain critical level.

For intermediate intervention intensity, ϕ, the endemic equilibrium curve exhibits an S-shape as illustrated by the thin curve ($\phi = 0.5$) in Figure 5.9 (a). For a range of R_0, two stable endemic equilibria exist, one low and one high, separated by one unstable endemic equilibrium. Panel (c) shows the corresponding bistability region in the $(R_0, \phi)-$ space marked by A. To improve the impact of this intervention additional measures must be taken in order to bring the disease level below the unstable equilibria that separates the two stable ones or to reduce R_0.

We can say that bistability results from the interplay between population compartments with different susceptibility status: a factor α_1 modifies susceptibility factor to reinfection of individuals in the low-risk group when compared to the population average σ. In the absence of intervention the reinfection threshold is determined by the population average susceptibility factor $R_0 = 1/\sigma$, as previously observed in Section 5.4.1. A widespread strategy of treatment (very high ϕ) would increase significantly the low-risk population moving the reinfection threshold to the right at $R_0 = 1/\alpha_1\sigma$. Formally this threshold is defined for the extreme case $\phi \to +\infty$, where all susceptible individuals belong to the low-risk group. Following the method in Section 2.5 to system (5.44), which in the limiting case ($\phi \to +\infty$) is equivalent to

$$\begin{cases} S_1' &= \mu - \lambda\alpha_1 S_1 - \mu S_1 \\ I_1' &= \lambda\alpha_1 S_1 \sigma\lambda\alpha_1 R_1 - (\tau+\mu) I_1 \\ R_1' &= \tau I_1 - \sigma\lambda\alpha_1 R_1 - \mu R_1 \end{cases} \tag{5.46}$$

we obtain the reinfection sub-model given by

$$\begin{cases} R_1' &= \mu + \tau I_1 - \sigma\lambda\alpha_1 R_1 - \mu R_1 \\ I_1' &= \lambda\alpha_1 S_1 + \sigma\lambda\alpha_1 R_1 - (\tau+\mu) I_1. \end{cases} \tag{5.47}$$

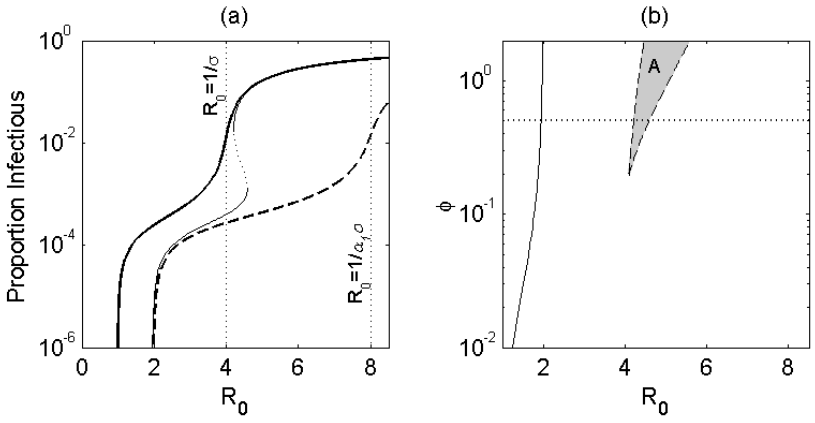


Figure 5.9: Early intervention for $\gamma = 0.5$, $\alpha_1 = 0.5$ **(a)** Equilibrium curves with intervention intensity $\phi = 0.5$ (unstable equilibrium is marked by dotted lines). Heavy full and dashed lines correspond to equilibria with no intervention and the limit case where $\phi \to +\infty$, respectively; **(b)** Two-parameter diagram where bistability region is marked by A. Full line corresponds to $R_0 = ET_\phi$. Horizontal lines correspond to the particular cases represented in (a).

The reinfection sub-model has a unique disease-free equilibrium $E_0 = (1, 0)$. Analyzing the jacobian at E_0 we conclude that a bifurcation on the transmission parameter β occurs at $\beta = (\tau + \mu)/\alpha_1\sigma$, when the disease free equilibrium changes its stability. In terms of the basic reproduction number the bifurcation is attained at $R_0 = 1/\alpha_1\sigma$.

For intermediate intensities the relative size of the recovered classes of each risk group, after the intervention, will then determine which of the reinfection thresholds has more impact in determining the disease level.

These results are applicable to many contexts or diseases. In general, when there is more than one susceptibility group, multiple reinfection thresholds can play a role in defining the disease prevalence. Different susceptibility factors that affect the reinfection rates are crucial to determine the position of the behavior of interest on the transmissibility axis, given by the corresponding reinfection thresholds. Moreover, the success of interventions that alter the relative size of the partially susceptible classes depend on this position. Another example is analyzed in Gomes *et al.* (2007) in the context of post-exposure interventions in tuberculosis.

Another form of bistable behavior may occur from the interplay between the reinfection threshold, $R_0 = 1/\sigma$, and the epidemic threshold, $R_0 = ET_\phi$ for the intervention model. When ET_ϕ is close to the reinfection threshold, bistability of the disease-free and an endemic equilibrium can occur. In fact, when α_1 is low (high heterogeneity) the curve $R_0 = ET_\phi$ intersects region A, which gives rise to a new bistability region, marked by B in Figure 5.10 (b). The upper stable equilibrium exists now to the left of the epidemic threshold $R_0 = ET_\phi$ as illustrated by the case $\phi = 0.5$ in Figure 5.10 (a). This scenario can have important consequences to intervention impact since disease can now be eliminated with additional measures that allow to bring the disease level below the unstable equilibrium that separates the stable ones.

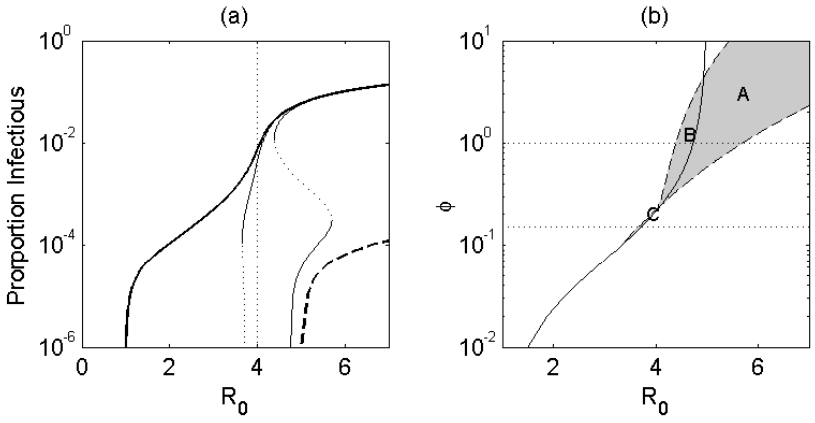


Figure 5.10: Early intervention for $\gamma = 0.8$, $\alpha_1 = 0.2$. **(a)** Equilibrium curves with intervention intensity $\phi = 0.15$ and 0.5, from left to right (full line with unstable equilibrium is marked by dotted lines). Heavy lines correspond to equilibria with no intervention, and dashed heavy lines represent the limit case where $\phi \to +\infty$; **(b)** Two-parameter diagram where bistability region is marked in grey. Bistability of the two endemic equilibria, high and low is marked by A. Bistability of the disease-free and one endemic equilibrium is marked by B and C. Horizontal lines correspond to the particular cases represented in the panels above.

Also for the bifurcation at $R_0 = ET_\phi$, two distinct situations may occur that have different consequences for the intervention impact. The nontrivial equilibrium may exist only to the right of the bifurcation point or to the left, for R_0 in some interval $(ET_\phi - \epsilon, ET_\phi)$ with $\epsilon > 0$. Accordingly, these are called forward (or supercritical) and backward (or subcritical) bifurcations (Dushoff, 1996). For the forward bifurcation the range for disease elimination is maximum, while for the backward bifurcation disease may remain endemic below $R_0 = ET_\phi$. Moreover, note that when the bifurcation is backwards, the unstable nontrivial equilibrium bents forward in a saddle node bifurcation, giving rise to bistability of the disease-free and endemic equilibria, below $R_0 = ET_\phi$ (region C in Figure 5.10 (b)). This situation is illustrated by the thin left curve in Figure 5.10 (a), corresponding to $\phi = 0.15$. It implies that, for a population with transmission intensity in this range, it is only possible to eliminate the disease if additional measures are taken in order to bring disease level below the unstable equilibria. Otherwise disease level is reduced to the level of the stable endemic equilibrium. The result concerning the direction of the bifurcation is summarized in the following theorem.

Theorem 5.5.2. *For system* (5.44) *with* $a = 0$*, the bifurcation of the disease-free equilibrium,* E_0*, at* ET_ϕ *is backwards iff*

$$\frac{\sigma\tau(\tau+\phi)[(1-\gamma)(\alpha_1^2\phi+\mu)+\mu\gamma(1-\alpha_1)^2]}{(\tau+\mu)[\phi^2\alpha_1^2(1-\gamma)+\mu\phi\alpha_1(1-\gamma)(\alpha_1+1)+\mu^2(\gamma(\alpha_1-1)^2+(1-\gamma))]} > 1. \tag{5.48}$$

Proof. By solving each equation from system (5.44), at equilibrium, in order to $I = I_1 + I_2$ and substituting in equation $\dot{I} = 0$ (sum of the second and third equations) we get a polynomial of fourth order $P(\beta, I) = a_4(I)\beta^3 + a_3(I)\beta^3 + a_2(I)\beta^2 + a_1(I)\beta + a_0(I)$. The equilibrium solutions satisfy $P(\beta, I^*) = 0$, where I^* is the equilibrium proportion of infectious. Defining β as a function of I^*, differentiating this equation and setting $(\beta, I^*) = (\beta^*, 0)$, where β^* is

such that $R_0 = ET_\phi$, we obtain

$$\frac{\mathrm{d}\beta}{\mathrm{d}I^*}|_{(\beta^*,0)} = -\frac{(\mu+\phi)(\tau+\mu)}{\mu(1-\gamma)(\mu+\alpha_1\phi)^3}A.$$

With the opposite sign of $A = a\sigma + b$ where $a = (\mu+\phi)\tau((1-\gamma)(\mu+\phi\alpha_1^2) + \mu\gamma(1-\alpha)^2) > 0$; and $b = -(\tau+\mu)[\phi^2\alpha_1^2(1-\gamma) + \mu\phi\alpha_1(1-\gamma)(1+\alpha_1) + \mu^2(\gamma(1\alpha_1)^2 + (1-\gamma))]$. Hence, a backward bifurcation occurs iff $A > 0$ which is equivalent to condition (5.48).

Remark 5.5.3. *We omit the expression of the polynomial P coefficients due to its complexity. Algebraic manipulation was obtained using the symbolic toolbox from* MATLAB 6.5®.

□

Extended intervention (a=1).

When the risk-reduction intervention is applied irrespective of infection status (system (5.44) with $a = 1$) bistability no longer occurs, independently of the intervention intensity (Figure 5.11). In fact, since all classes S, I and R, change at the same rate, only the structure of

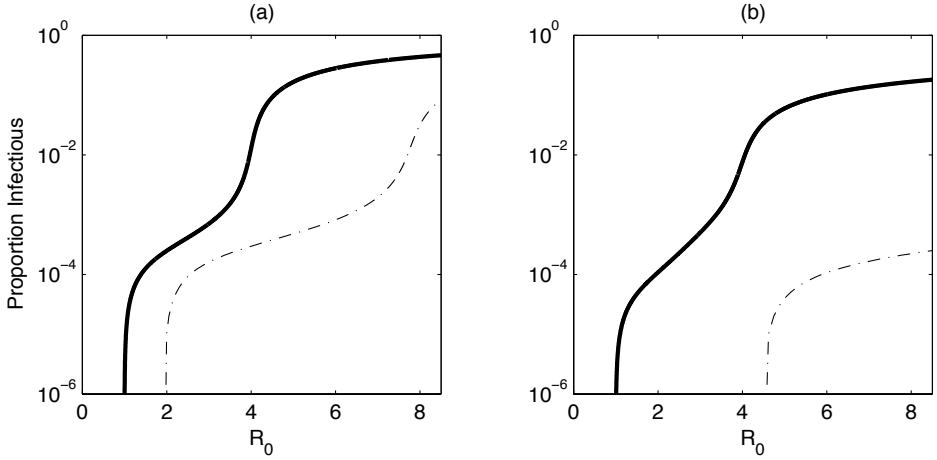


Figure 5.11: Extended intervention. Equilibrium curves for intermediate intervention intensity ($\phi = 0.5$) and for different scenarios for different scenarios for the risk distribution: (a) low ($\gamma = 0.5$, $\alpha_1 = 0.5$); and (b) high distribution variance ($\gamma = 0.8$, $\alpha_1 = 0.2$). Heavy line corresponds to the equilibrium curve with no intervention ($\phi = 0$).

the population and the average population risk are changed. The equations for the total population in each group P_i are

$$\begin{cases} P_1' &= \gamma\mu + \phi P_2 - \mu P_1 \\ P_2' &= (1-\gamma)\mu - (\mu+\phi)P_2. \end{cases} \tag{5.49}$$

Hence the equilibrium solutions are $P_1 = \dfrac{\mu\gamma+\phi}{\mu+\phi}$ and $P_2 = \dfrac{(1-\gamma)\mu}{\mu+\phi}$ and the average risk is $\dfrac{\alpha_1 P_1 + \alpha_2 P_2}{P_1+P_2} = \dfrac{\mu+\alpha_1}{\mu+\phi} = \dfrac{1}{ET_\phi}$. This means that the new disease equilibrium corresponds

to the case where the low-risk group represents a higher proportion of the population, $(\phi + \gamma\mu)/(\phi + \mu)$ instead of γ, and the average risk of infection among the population is reduced from 1 to $1/ET_\phi$.

Below the reinfection threshold, both early and extended interventions have the similar results. Above the reinfection threshold the results of the extended intervention ($a = 1$) are strikingly better. However, note that the intervention effort is not directly comparable. For the early intervention the proportion of individuals covered is ϕS_2, whereas for the extended intervention, is $\phi(S_2 + I_2 + R_2) = \phi(1 - \gamma)$, which is independent of the transmission intensity and always higher. For higher endemic regions, the difference is much more pronounced.

5.6 Discussion

We have previously identified a reinfection threshold in the SIRI model and characterised how this induces a sharp division of the transmissibility axis into two regimes: reinfection is rare below threshold (SIR behavior) and very frequent above (SIS behavior). Here we describe how heterogeneity in innate susceptibility to infection smoothens this transition by making both regimes less extreme. Heterogeneity is always present in nature and it is important to understand how it can affect system behavior both qualitatively and quantitatively.

We perform a systematic analysis of the SIRI model with distributed susceptibility. The most striking result is the prediction that the average rate of reinfection may be higher than the average rate of primary infection, which may seem paradoxical given that primary infection induces life-long partial protection. The rationale behind this result is that infection generates a selection mechanism that skews the susceptibility profiles of the S and R compartments to lower and higher susceptibility, respectively. In other words, selection acts to keep less susceptible individuals in S and more susceptible individuals in R. If this effect is strong enough we have a scenario where, on average, the rate of reinfection (infection out of R) is higher than the rate of primary infection (infection out of S) even though each individual has a risk reduction following primary infection. This mechanism may explain high rates of tuberculosis reinfection recently reported (Verver *et al.*, 2005), as we will describe in Chapter 6.

A rule of thumb has been proposed in infections disease dynamics, whereby 20% of the population is responsible for 80% of all infections due to heterogeneity in susceptibility or infectivity (Woolhouse *et al.*, 1997). However, direct confirmation of this hypothesis requires very large epidemiological studies. For diseases that induce partial immunity, mathematical models such as those proposed here offer the practical alternative of using the ratio between reinfection and primary infection rates as an indirect measure of population heterogeneity.

In the SIRI models with heterogeneous susceptibility, we predict that disease prevalence is lower than the corresponding homogeneous model, as described before for epidemic SIR models (Gart, 1968; Ball, 1985; Anderson & Britton, 1998; Miller, 2007). In other words, to obtain a given level of disease prevalence, the heterogeneous model requires a higher value for the transmission intensity, R_0. This implies that elimination strategies require more effort under wider heterogeneity (Anderson & May, 1991).

The success of vaccination depends then on the ability to target those individuals at higher groups. Generally, there is an additional benefit associated with targeting vaccination strategies, as previously described for the SIR epidemic model (Britton, 1998; Koopman *et al.*,

2005). In the case of the SIRI model, however, the added value of targeting high-risk groups is limited to those regions where transmission is below the reinfection threshold. To overcome this limitation, interventions must be able to change the infection risk of the targeted population. Early interventions, that act on susceptible individuals only, can generate a bistable situation for which an adequate intervention intensity can eliminate the disease or bring it to a lower endemic level for certain regions. If these interventions are more extensively applied irrespective of infection status, then the effect is stronger and the reinfection threshold can be moved to higher transmission intensities.

Chapter 6

The selection hypothesis in tuberculosis

6.1 Introduction

Despite important improvements in tuberculosis treatment, adequately treated patients are still at high risk of developing recurrent pulmonary disease. Recurrent TB is defined as a second episode of TB occurring after a first episode had been considered cured. A review on the recurrence rate of TB for different regions revealed an average of 2,290 cases per 100,000 person-years at 12 months after treatment completion (Panjabi *et al.*, 2007). However, this estimate is biased towards the more common low incidence regions. In high incidence regions, the average TB recurrence rate can reach 7,850 per 100,000. Recurrent TB poses a significant challenge to public health and control programs, as it is associated with drug resistance and treatments with low cure rates. Re-treatment is costly, posing further difficulties in regions with low income that normally have a greater TB burden.

The role of exogenous reinfection with *Mtb*, versus endogenous reactivation (relapse) of latent *Mtb* in the recurrence of pulmonary disease is not completely understood. Deciphering the weigh of each of these mechanisms is of great importance in the choice of the most effective control program. Advances in DNA fingerprinting techniques allowed the genotyping of the *Mtb* causing different disease episodes (McNabb *et al.*, 2002). These methods can reveal whether a new episode of disease is caused by infection with the same strain that caused a previous episode or a different one, permitting a classification into relapse or reinfection, respectively. Despite this correlation not being completely accurate due to the possibility of mixed infections, reinfection with the same strain or laboratory cross-contamination, it can be used as a proxy for the relative frequency of reinfection, and relapse in recurrent TB. In regions with moderate to high endemicity, molecular epidemiological studies have reported that disease caused by reinfection can be responsible for the majority of recurrent cases (Verver *et al.*, 2005). This phenomenon can have strong implications for public health control strategies (Chiang *et al.*, 2005).

Recently, it was observed that reinfection rates correlate with the logarithm of local TB incidence (Wang *et al.*, 2007), and an algorithm based on linear regression was proposed to predict the proportion of reinfection from local incidence. The significantly high recurrence rate and the increased evidence of reinfection contribution to recurrent TB have raised the

issue of better quantifying the infection risk among successfully treated patients. In particular, in an area of South Africa with high disease prevalence it has been reported that the rate of TB reinfection after successful treatment is significantly higher than the rate of new TB (Verver *et al.*, 2005). This observation led the authors to deduce that previous TB episodes induce a form of immune-mediated enhancement of susceptibility to reinfection (Uys *at al.*, 2009).

Based on incidence and reinfection data from 6 regions distributed worldwide, we contrast this explanation with the selection hypothesis, whereby high-risk individuals are overrepresented in the previously infected subpopulation (Yew & Leung, 2005) inflating the rate of reinfection at the population level even though immunity confers partial protection at the individual level. Specifically, we postulate that some individuals are a priory more likely to develop the disease because they are more exposed or have some form of innate susceptibility. In this case, the risk of reinfection in the group of people who previously had TB disease could be higher due to heterogeneity in exposure or innate susceptibility (Austin *et al.*, 2004; Hoal *et al.*, 2004; Sonnenberg *et al.*, 2001; Story *et al.*, 2007).

In this Chapter we a different methodology. We construct two alternative models (with and without heterogeneity) each parameterized by fitting to publish data. The fitting procedure was performed in collaboration with Ricardo Águas using a standard software. Unfortunately, the data consists on only 6 regions which does not allow for a satisfactory statistical discrimination of the models. We advocate for the selection hypothesis based in epidemiological arguments. In the appendix (Section 6.5) we also discuss some of the differences between the models from the structure point of view. The understanding of the behavior of the heterogeneous model results also from the analysis of the simple SIRI model presented in the previous chapter (Section 5.4).

6.2 Methods

6.2.1 The model

Our purpose here is to get data comparable results therefore we have to refine the simple tuberculosis model used in Chapter 3. We choose as reference the model in Gomes *et al.* (2007), where a new compartment is considered for recently exposed individuals harboring a primary infection and where distinction between latent and previously treated classes individuals is made, allowing to differentiate reinfection and reactivation events.

The assumption on partial immunity is relaxed to allow the reinfection rate to be any factor of first infection. We further expand the model to accommodate host heterogeneity in susceptibility to infection, using the same formulation as in the previous chapter. The population is divided into two risk groups. Within each group (indexed by $i = 1, 2$), individuals are classified, according to their infection history, into susceptible (S_i), recently exposed harboring a primary infection (P_i), latent (L_i), active pulmonary tuberculosis (I_i) and recovered (R_i). Figure 6.1 provides a schematic representation of the model. We denote by γ , the proportion of the population assigned to group 1, following that a proportion $1-\gamma$ is assigned to group 2, that is $S_1 + P_1 + L_1 + I_1 + R_1 = \gamma$ and $S_2 + P_2 + L_2 + I_2 + R_2 = 1 - \gamma$. The force of infection acting on each risk group is $\lambda_i = \alpha_i \beta I$, where $I = I_1 + I_2$, β is a transmission coefficient specific to each population, and α_i represent the risk factors that differentiate the

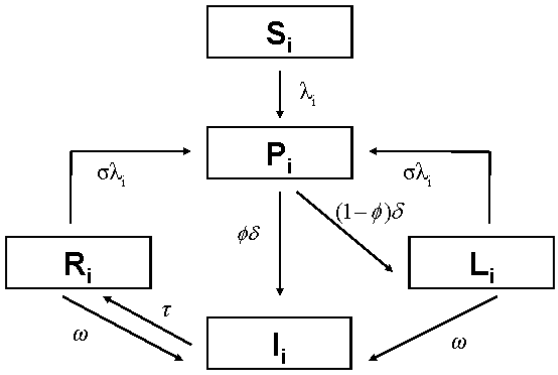


Figure 6.1: Tuberculosis model diagram. Individuals are classified according to infection history into susceptible (S_i), primary infection (P_i), latent (L_i), active pulmonary tuberculosis (I_i) and recovered (R_i), where the index ($i = 1, 2$) designates the risk group. Parameters are the rate of progression from primary infection (δ), the proportion progressing to pulmonary disease (ϕ), the rate of endogenous reactivation of latent infections (ω), the rate at which infectious individuals are detected and treated (τ) and the relative risk of reinfection over first infection (σ). The force of infection applying to each risk group is λ_i.

two risk groups. To standardize the analysis and interpretation, we normalize the average risk factor such that $\gamma\alpha_1 + (1-\gamma)\alpha_2 = 1$. The rate of reinfection is affected by the reinfection factor, σ, that corresponds to the relative risk of reinfection in relation to first infection. The model can be written as a system of 2×5 differential equations

$$\begin{cases} S_i' &= \mu\gamma_i - (\lambda_i + \mu)S_i \\ P_i' &= \lambda_i S_i + \sigma\lambda_i(L_i + R_i) - (\delta + \mu)P_i \\ I_i' &= \phi\delta P_i + \omega L_i + \omega R_i - (\tau + \mu)I_i \\ L_i' &= (1-\phi)\delta P_i - (\sigma\lambda_i + \omega + \mu)L_i \\ R_i' &= \tau I_i - (\sigma\lambda_i + \omega + \mu)R_i. \end{cases} \tag{6.1}$$

Individuals are born at a fixed rate, μ, and enter S_1 or S_2 in the fixed proportions, $\gamma_1 = \gamma$ and $\gamma_2 = 1 - \gamma$, respectively. For both risk groups, susceptible individuals, when infected, move to the primary infection compartment, which they leave at a rate, $\delta = 12\ yr^{-1}$. A fraction of infections, $\phi = 0.05$, progresses directly to disease, while the majority is able to contain the infection moving to the latent class. Infectious individuals recover by treatment to the recovered class at a rate, $\tau = 2\ yr^{-1}$, which reflects the average time to detection and smear conversion after treatment initiation. Both latent and recovered infections can further progress to disease upon endogenous reactivation or exogenous reinfection. We assume the simplest possible scenario where both of these processes are equivalent for latent and recovered classes. Populations differ in the transmission potential given by β. Parameters μ, ϕ, δ and τ are fixed and their values are the same as in Gomes *et al.* (2007). The remaining parameters are estimated.

6.2.2 Basic reproduction number

The basic reproduction number, R_0, is calculated using the next generation operator described in (van den Driessche & Watmough, 2002).

Accordingly, the infected classes are P_i, L_i, I_i and R_i, provided that $\omega \neq 0$. So we write

system (6.1) as

$$X' = f(X) \Leftrightarrow X' = \mathcal{F}(X) - \mathcal{V}(X), \tag{6.2}$$

where X is the ordered vector of the state variables, $\mathcal{F} = (\alpha_1\beta IS_1, 0, 0, 0, \alpha_2\beta IS_2, 0, 0, 0, 0, 0)^T$ is the rate of appearance of new infections in each class and the disease-free equilibrium is $X_0 = (0, 0, 0, 0, 0, 0, 0, 0, \gamma, 1-\gamma)$. The derivatives of $\mathcal{F}$ and $\mathcal{V}$ with respect to the infected classes at X_0 are

$$F = \begin{bmatrix} 0 & 0 & \alpha_1\beta\gamma & 0 & 0 & 0 & \alpha_1\beta\gamma & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_2\beta(1-\gamma) & 0 & 0 & 0 & \alpha_2\beta(1-\gamma) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$V = \begin{bmatrix} \delta+\mu & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -(1-\phi)\delta & \omega+\mu & 0 & 0 & 0 & 0 & 0 & 0 \\ -\phi\delta & -\omega & \tau+\mu & -\omega & 0 & 0 & 0 & 0 \\ 0 & 0 & -\tau & \omega+\mu & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta+\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -(1-\phi)\delta & \omega+\mu & 0 & 0 \\ 0 & 0 & 0 & 0 & -\phi\delta & -\omega & \tau+\mu & -\omega \\ 0 & 0 & 0 & 0 & 0 & 0 & -\tau & \omega+\mu \end{bmatrix}.$$

Now, the basic reproduction number is defined as the spectral radius of the next generation matrix, FV^{-1}:

$$R_0 = \frac{\beta\delta(\phi\mu+\omega)}{\mu(\mu+\delta)(\tau+\omega+\mu)}. \tag{6.3}$$

Using the direct calculations as in Hethcote (2000) we can rewrite R_0 with a more manful expression:

$$R_0 = \beta\Big(\frac{\phi\delta}{\delta+\mu} + \frac{(1-\phi)\delta}{\delta+\mu}\frac{\omega}{\omega+\mu}\Big)\ \frac{1}{\tau+\mu}\ \frac{(\tau+\mu)(\omega+\mu)}{\mu(\tau+\omega+\mu)}. \tag{6.4}$$

We can see that R_0 is proportional to the transmission coefficient, β, and encompasses the two alternative paths to disease progression corresponding to the two terms inside brackets. The average infectious period, $1/(\tau+\mu)$, is multiplied by the average time to reactivation of the recovered individuals. Note that the expression for R_0 is model dependent, so it should be interpreted as a way to classify different populations according to their potential for transmission, under the assumptions made. R_0 is a threshold parameter for endemicity, as for $R_0 > 1$ the disease free equilibrium becomes unstable and an endemic stable equilibrium emerges.

6.2.3 The data

This study includes data points obtained from Wang *et al.* (2007) corresponding to the reinfection proportion in recurrent TB and local TB incidence for 6 regions distributed worldwide.

The data points in (Wang *et al.*, 2007) were gathered by systematic literature review, following a set of inclusion criteria defined by the authors. Recurrent TB was defined as cases with culture positive after bacteriologically confirmed cure or complete treatment for the first episode. Reinfection and reactivation were distinguished by comparing *Mtb* DNA fingerprinting of the initial and recurrent episodes. The proportion of reinfection was defined as patients with reinfection over all with recurrent TB. We excluded the point corresponding to the Netherlands (de Boer & Soolingen, 2000) since the proportion of reinfection has been revised in a more recent study conducted by the same authors (de Boer *et al.*, 2003). After careful analysis of the original papers we have corrected the value for the reinfection proportion for Houston to 24% (El Sahly, 2004). For Cape Town, we used the average total notification incidence of TB, 761 per 100,000 (supplementary material in (Verver *et al.*, 2005)) instead of the new TB incidence previously used to represented the local TB incidence (Wang *et al.*, 2007). Finally, we have included the point corresponding to the study conducted in Taiwan (Wang *et al.*, 2007), which in the original paper was used to confirm the prediction from a linear regression model.

6.2.4 Measures of TB incidence

Following the criteria used in the data collection, we classify a recurrent TB case as any individual who enters the infectious compartment after having gone through the recovered class. There are two alternative pathways: (i) reactivation while in the recovered class; or (ii) reinfection with progression to active pulmonary disease (direct or following a latent period). From system (6.1), these are formally defined by

$$reACT = \omega \sum_i R_i, \tag{6.5}$$

$$\begin{aligned} reINF = &\sum_i \sigma\beta I\alpha_i R_i \Big(\frac{\phi\delta}{\delta+\mu} + \frac{(1-\phi)\delta}{\delta+\mu} \frac{\omega}{\omega+\mu+\sigma\beta I\alpha_i} \Big) \times \\ &\Big(\frac{(\omega+\mu+\sigma\beta I\alpha_i)(\delta+\mu)}{(\omega+\mu)(\delta+\mu)+\sigma\beta I\alpha_i(\mu+\phi\delta)} \Big), \end{aligned} \tag{6.6}$$

for i = 1,2, where I and R_i are equilibrium values. The expression for $reINF$ is given by the rate at which recovered individuals are reinfected, $\sigma\beta I\alpha_i$, times the probability to survive the exposed period either progressing directly to disease, $\dfrac{\phi\delta}{(\delta+\mu)}$, or going first through an extended latent period, $\dfrac{(1-\phi)\delta\omega}{(\delta+\mu)(\omega+\mu+\sigma\beta I\alpha_i)}$, and by the term $\dfrac{\omega+\mu+\sigma\beta I\alpha_i)(\delta+\mu)}{(\omega+\mu)(\delta+\mu)+\sigma\beta I\alpha_i(\mu+\phi\delta)}$ which accounts for the chance of being again reinfected while latent, going back to the exposed class.

The proportion of reinfection, at equilibrium, is given by

$$p = \frac{reINF}{reINF + reACT}. \tag{6.7}$$

TB incidence, defined by the number of cases per 100,000, is calculated from equation $I' = \phi\delta(\sum_i P_i) + \omega\sum_i(L_i + R_i) - (\tau+\mu)I$ at equilibrium, as the proportion of cases entering the infectious class times 100,000 cases:

$$y = (\tau+\mu)\ I \times 100,000. \tag{6.8}$$

6.2.5 Fitting procedure

Parameters are estimated under two sets of assumptions: homogeneous susceptibility ($\alpha_1 = \alpha_2 = 1$); and heterogeneous susceptibility ($\alpha_1 < \alpha_2$). In both situations we assume that the rate of endogenous reactivation (ω), the reinfection factor (σ), and the parameters regulating heterogeneity (α_i and γ), do not vary between populations (global parameters). The force of infection, λ (or λ_i), is assumed to be a linear parameter and it is allowed to vary between regions (local parameters). Parameters are then estimated as the best fit to the proportion of recurrences due to reinfection (p) and local incidence (y) for the 6 data points, simultaneously. We assume that the data observations were performed at a time when TB transmission has reached a stationary state. This assumption allowed us to ensure the robustness of the parameter estimates. Hence, we run the model in time until equilibrium is reached, and compare the resulting values of p and y with the data points for each region. This is embedded in a least squares minimization fitting method using the Berkeley Madonna software v8.3.6c. The method minimises the sums of the squares of residuals, which is the difference between the model prediction and the data output at each data point. We assume a Gaussian distribution for scatter of residuals and the same standard deviation for all points. Initial guesses for the parameters are chosen a priori from a biologically plausible range of values, in such a manner as to serve as boundaries for the most likely value for those parameters. Guesses are fined tuned iteratively, according to the resulting estimates. The ideal fit is the one that results in the least sum of square of residuals, and gives estimates for the parameters which are within our plausibility range.

The strategy, adopted here, of simultaneously fitting datasets for different population leaving the force of infection (λ) as the only regions specific parameter has proven successful in studies for other diseases (White *et al.*, 2007; Águas *et al.*, 2008).

From the estimated region-specific forces of infection we derive the region-specific transmission coefficients by dividing the force of infection by the total prevalence following the relation $\lambda_i = \beta\alpha_i I$. So, formally we have:

$$\beta_j = \frac{\lambda_j}{\alpha_i I_j}, \text{ for } i = 1, 2 \text{ and } j = 1, \ldots, 6, \tag{6.9}$$

where $I_j = \frac{y_j}{(\tau+\mu)100{,}000}$ is the region specific proportion of infectious individuals, obtained from equation (6.8). From relation (6.3) we can also compute the region specific reproduction numbers:

$$R_{0,j} = \frac{\beta_j \delta(\phi\mu + \omega)}{\mu(\mu + \delta)(\tau + \omega + \mu)}, \ j = 1, \ldots, 6.. \tag{6.10}$$

6.2.6 Ratio of reinfection over new TB

We define ρ as the ratio of the rate of reinfection among successfully treated patients, $\frac{reINF}{R_1 + R_2}$, over the rate of new TB among never-infected individuals, $\frac{newINF}{S_1 + S_2}$, where $reINF$ is defined

above and similarly we define $newINF$ as

$$newINF = \sum_i \beta I \alpha_i S_i \Big(\frac{\phi\delta}{\delta+\mu} + \frac{(1-\phi)\delta}{\delta+\mu} \frac{\omega}{\omega+\mu+\sigma\beta I\alpha_i} \Big) \Big(\frac{(\omega+\mu+\sigma\beta I\alpha_i)(\delta+\mu)}{(\omega+\mu)(\delta+\mu)+\sigma\beta I\alpha_i(\mu+\phi\delta)} \Big), \tag{6.11}$$

Note that even reinfected individuals who have never experienced disease before are classified as new cases. More formally, we calculate ρ as

$$\rho = \frac{reINF(S_1+S_2)}{newINF(R_1+R_2)}. \tag{6.12}$$

Interestingly, for the homogeneous model, we have $\rho = \sigma$, since $\alpha_1 = \alpha_2 = 1$. For the heterogeneous model, the rates of reinfection and new infection are weighted according to the equilibrium susceptibility profiles of the R and S compartments, respectively. Data and calculations for Cape Town are shown in Table 6.3.

Based on our previous study of the SIRI heterogeneous model behavior (Chapter 5) we propose ρ as an alternative measure that can further distinguish the competing hypotheses. For the simple SIRI model ρ corresponds to Y_2/Y_1 (equations 5.37 and 5.38), illustrated in Figure 5.6 (b) for different values of α_1. In populations for which susceptibility to infection is sufficiently heterogeneous, the selection hypothesis distinguishes itself from homogeneous susceptibility, for which the ratio is constant $\rho = \sigma$, by predicting a much higher reinfection rate for low to intermediate transmission intensities due to selection of high-risk individuals to the recovered sub-population.

6.3 Results

We construct a mathematical model for TB transmission (adapted from Gomes *et al.* (2007)) where increased attention is given to recurrent TB. The model is based on the assumption that individuals with a latent infection, or that have recovered from an active disease episode after effective treatment, can be reinfected at a rate that is proportional to the rate of first infection, with multiplicity factor σ. We also consider the possibility that susceptibility to infection can be heterogeneously distributed among the population. The total population is divided into two susceptibility groups with distinct risk factors, $\alpha_1 \leq \alpha_2$. The low-risk group constitutes a proportion, γ, of the population which is constant over time. Within each group group, individuals are classified according to their infection history (see Figure 6.1).

We analyze the differences in the contribution of reinfection and reactivation to recurrent TB across distinct regions/countries. The model is parameterized by the transmission coefficient (β), which differentiates regions/countries according to socioeconomic and environmental factors and impacts on the force of infection (λ, see Methods). The two alternative hypotheses for the inflation of reinfection rates are then contrasted by two distinct model implementations. The hypothesis of immune-mediated enhancement is realized in a homogeneous host population ($\alpha_1 = \alpha_2 = 1$, $\sigma > 1$), while the selection hypothesis invokes a heterogeneous host population to reconcile high reinfection rates with the view of partial immune protection ($\alpha_1 < \alpha_2$, $\sigma < 1$). The two model versions are parameterized on published data for the proportion of reinfection in recurrent TB versus TB incidence across a range of

communities from low to high incidence (Wang *et al.*, 2007).

As in previous models (White *et al.*, 2007; Águas *et al.*, 2008) we have considered the force of infection (λ) as the local parameter that capture the differences in transmission among regions, the remaining being set as global parameters. Tables 6.1 and 6.2 show the estimated

Table 6.1: Estimated Global Parameters.

Symbol	Definition	Homogeneous Model	Heterogeneous Model
σ	reinfection factor	6.56	0.617
ω	rate of endogenous reactivations	0.00425	0.00779
γ	proportion low-risk group	1 (imposed)	0.937
α_1	low risk factor	1(imposed)	0.0690
α_2	high risk factor	1(imposed)	14.8

global and local parameters, respectively, according to the procedure described in Methods. Concerning the region-specific transmission parameters we have, initially, estimated the forces

Table 6.2: Estimated region-specific force of infection (λ or λ_i) and derived transmission coeffitient (β) and basic reproduction number R_0.

US & Canada	Lombardy, Italy	Houston, US	Taiwan	Madras, India	Cape Town, South Africa	
			Homogeneous	model		
λ	0.00014733	0.00050992	0.00062570	0.0018278	0.001006	1 0.023296
β	5.8231	5.8323	5.8352	5.8631	5.8451	6.1224
R_0	1.0013	1.0029	1.0034	1.0082	1.0051	1.0527
			Heterogeneous	model		
λ_2	0.0012664	0.0053642	0.0059830	0.045075	0.016876	2.0896
β	3.6691	4.4998	4.6211	10.604	6.5942	40.275
R_0	1.0802	1.3247	1.3604	3.1218	1.9413	11.856

of infection and then calculated the corresponding values for the transmission coefficients and basic reproduction numbers (R_0), through equations (6.9) and (6.10).

Figure 6.2 shows the equilibrium solutions of the model that best fit the proportion of recurrent TB attributed to reinfection (p) and local TB incidence (y) for the 6 study populations included in this study (see inclusion criteria in Methods). When innate susceptibility to infection is homogenous (Figure 6.2(a,c)) the estimated value for the reinfection factor ($\sigma = 6.56$) suggests some form of immune-mediated enhancement whereby immunity renders individuals more susceptible to subsequent infections (Uys *at al.*, 2009). Contrastingly, when two susceptibility groups are permitted (Figure 6.2(b,d)) the estimated value for the reinfection factor ($\sigma = 0.62$) is compatible with the more consensual view that immunity confers partial protection against subsequent reinfections (Smith *at al.*, 1994; Dye *et al.*, 1998). The estimates for the heterogeneity parameters indicate that susceptibility is considerably higher

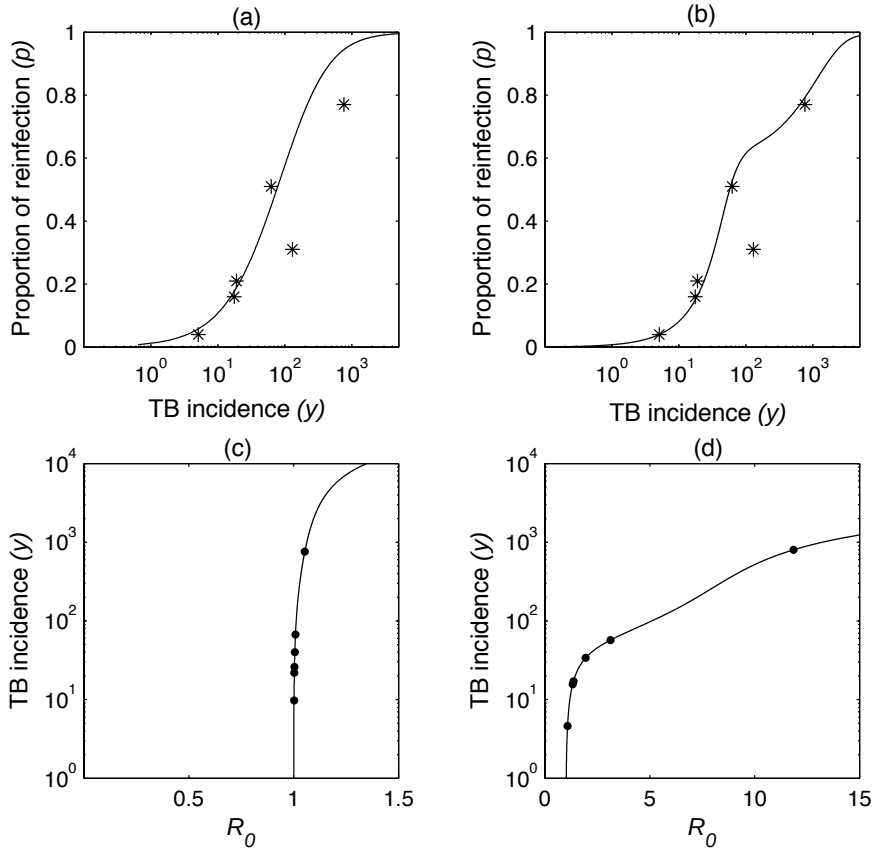


Figure 6.2: Comparison of model and data. **(a,b)** Proportion of reinfection in recurrent tuberculosis as a function of local incidence (log). The data points correspond to 6 regions (% reinfection, incidence per 100,000 person-years): US & Canada (4%, 5.1); Lombardy, Italy (16%, 17.5); Houston, US (24%, 19); Taiwan (51%, 62.4); Madras, India (31%, 130) and Cape Town, South Africa (77%, 761). The curves correspond to (a) homogeneous model and (b) heterogeneous model equilibria, using local parameters according to Table 6.1. **(c,d)** TB incidence per 100,000 person-years, at equilibrium, as a function of R_0 for the homogeneous model (c) and the heterogeneous model (d). The dots mark the transmission coefficients for the 6 regions, as in Table 1, and the associated incidences predicted by the model.

($\alpha_2 = 14.8$ times higher than population average) among a small sub-group consisting of 6% $(1-\gamma)$ of the population. This is in agreement with TB transmission, especially in regions of low to moderate transmission, where TB is confined to particular risk groups (such as homeless, immigrants or prisoners) with sporadic small outbreaks in the general population (Story *et al.*, 2007; Nardell *et al.*, 1986). Both scenarios indicate reactivation rates ($\omega = 0.00425$ and $\omega = 0.00779$, respectively) on the upper range of previously published estimates (Dye *et al.*, 1998; Sutherland *et al.*, 1982). The equilibrium curves are parameterized by the local transmission coefficient, showing a marked nonlinear relation between the proportion of reinfection and local incidence, not capture by previous studies (Wang *et al.*, 2007).

Figure 6.2(c,d) shows the same equilibrium curves plotted as incidence in terms of R_0. The 6 study populations are positioned according to the local TB incidence. In the homogeneous case (Figure 6.2(c)) despite the range of TB incidence observed, the values found for the basic

reproduction numbers accumulate just above the epidemic threshold ($R_0 = 1$), which is not compatible with the notable persistence of TB transmission under the most aggressive control measures. In the heterogeneous scenario (Figure 6.2(d)) the estimates for R_0 correspond to a wider range of transmission coefficients in accordance with previously estimated values (Trunz *et al.*, 2006), showing a noticeable separation between Cape Town and the remaining regions (US/Canada, Lombardy, Houston, Taiwan, Madras).

Concerns with the intensity of transmission in Cape Town have motivated follow-up studies to characterize the contribution of reinfection (Verver *et al.*, 2005; Uys *at al.*, 2009). Patients with reported TB in the area were followed for an average duration of 5 years, and the incidence rate of TB attributable to reinfection after successful treatment was estimated as seven times higher than the rate of new TB (four times higher when incidence rates are adjusted for age).

We have calculated the ratio, ρ (6.12), of the two rates (reinfection TB over new TB) with outputs from the model, using the local parameters estimated for Cape Town, and obtained $\rho = 6.6$ in the homogeneous implementation and $\rho = 5.4$ when two risk groups are considered (see Table 6.3), both within the range 4–7 predicted by (Verver *et al.*, 2005; Uys *at al.*, 2009). We have calculated the ratio, ρ, over the entire range of local TB incidences included in this

Table 6.3: Rates of new TB and reinfection TB using parameter values estimated for the Cape Town region. The values obtained for the ratio, ρ, are in the range 4–7 estimated in (Verver *et al.*, 2005; Uys *at al.*, 2009).

	Homgeneous Model (prediction)	Heterogeneous Model (prediction)
Rate new TB* $newINF/(S_1 + S_2)$	1095	579
Rate reinfection TB* $reINF/(R_1 + R_2)$	7180	3114
ρ	6.6	5.4

*Cases per 100,000 person-years.

study. The resulting curves obtained with both model implementation reveal different trends (Figure 6.3) suggesting a potential criterion for model discrimination and hypothesis testing. While the homogeneous model predicts a constancy of reinfection TB over new TB across the entire range of transmission intensities, the heterogeneous models predicts that regions of low to moderate transmission intensity support relatively higher reinfection rates. As for the simple SIRI model we study in the previous section, this is explained by the gradient in the selective pressure for high-risk individuals to the recovered subpopulation (see Methods). The rationale is that infection tends to affect individuals at higher risk, skewing the distribution of recovered individuals towards higher susceptibility and inflating the rates of reinfection. Transmission intensity tends to homogenize the two distributions making this effect less pronounced. Cape Town is, to our knowledge, the only study reporting a sufficiently long follow-up of successfully treated patients to permit a reliable estimation of this ratio. The nearest is the study of US & Canada (Jasmer *et al.*, 2004) based on a follow-up

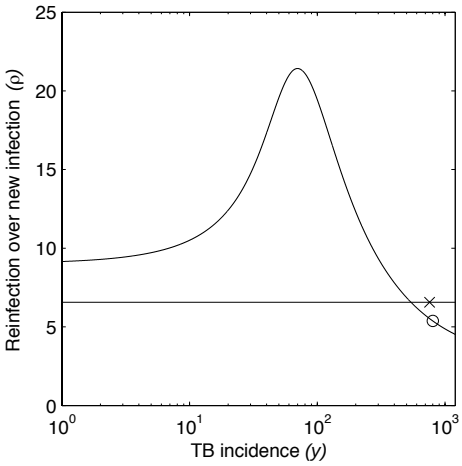


Figure 6.3: Ratio between reinfection TB and new TB. Model predictions for the ratio, ρ, of reinfection TB after successful treatment and new TB as a function of TB incidence per 100,000 person-years. Symbols ($\times$) and ($\circ$) correspond to the homogeneous ($\sigma = 6.56$) and heterogeneous ($\sigma = 5.38$) models, respectively, for the region of Cape Town, using local parameters as in Tables 6.1 and 6.2.

of 2 years, which suggests that the rate of reinfection TB is nearly thirty times higher than the rate of new TB. Although more date is necessary to validate the trends predicted here, preliminary results are consistent with the selection hypothesis.

6.4 Discussion

We propose a minimal model for TB transmission to describe the relative contributions of reinfection and reactivation to recurrent TB across a range of transmission intensities. A nonlinear relation between the proportion of reinfection and the local incidence is derived by fitting this mechanistic model to a dataset compiled by (Wang *et al.*, 2007) under two different assumptions on the distribution of susceptibility to infection among the population: homogeneous ($\alpha_1 = \alpha_2 = 1$) and heterogeneous ($\alpha_1 < \alpha_2$). In both cases, we assume that the reinfection rate is a multiple of the rate of first infection, with a reinfection factor, σ. Two alternative hypotheses were confronted to explain recent results from molecular epidemiological studies indicating that rates of reinfection tuberculosis are higher than rates of new tuberculosis. For the homogeneous model, the estimated parameters suggest that increased reinfection results from some form of immune-mediated enhancement in susceptibility after successful treatment ($\sigma > 1$). For the heterogeneous model, results suggest that infection confers partial immunity ($\sigma < 1$) to subsequent infections and that high susceptibility to infection is restricted to a small group of the population. The increased rate of reinfection is then explained by a selective pressure imposed by infection on the more susceptible group, sewing the distribution of the recovered subpopulation towards higher susceptibility. Although both model implementations reproduce a previously reported correlation between tuberculosis reinfection proportion and local incidence (Wang *et al.*, 2007), they lead to contrasting conclusions regarding adaptive immunity. We propose the selection hypothesis to reconcile the more consensual view that infection with *Mtb* confers partial protection to subsequent reinfection (Smith *at al.*, 1994;

Dye *et al.*, 1998) with recent reports of rates of reinfection that are higher that rates of new infection (Verver *et al.*, 2005; Jasmer *et al.*, 2004). We have obtained estimates for the basic reproduction number in 6 distinct regions that are compatible with previous estimates (Trunz *et al.*, 2006). As a criterion for indirectly testing the selection hypothesis for tuberculosis we propose that rates of reinfection and new infection are obtained under different transmission intensities and confronted with the trend predicted by the model. More specifically, the model predicts that rates of reinfection are relatively higher under low to intermediate transmission intensities due to selection of high-risk individuals to the recovered subpopulation.

6.5 Appendix

In this section we extend the analysis of the competing models and complete some of the arguments in favor of the selection hypothesis.

Nonlinear relation between the proportion of reinfection and local incidence

Apart from the scarce data available to be able to compare these models statistically, there is another difficulty in validating or refusing one of them. This difficulty resides in the fact that the comparison is based on the proportion of reinfection as compared to reactivation in recurrent cases. In fact, since both models assume that reactivation rate (ω) is independent of disease level in opposition to reinfection, that is proportional to the proportion of infectious individuals ($\sigma\beta I$), it will always result a nonlinear relation between the proportion of reinfection and local incidence. Above a certain transmission level and therefore a certain

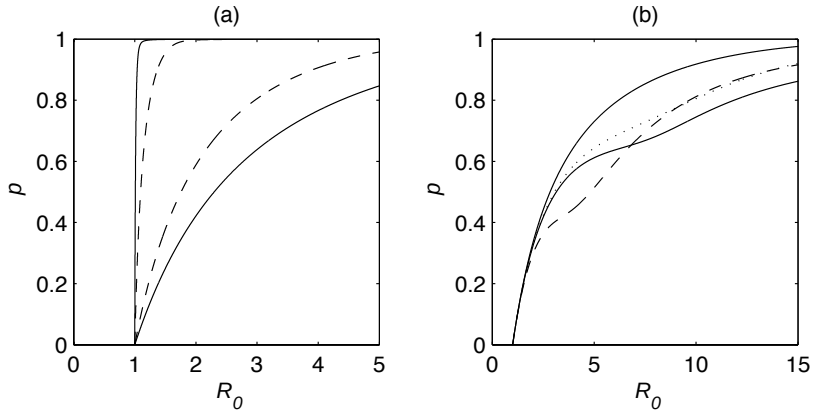


Figure 6.4: Change in the proprotion of reinfection in recurrent TB with R_0. **(a)** Proportion of reinfection for the homogeneous model with different values of the reinfection factor σ: 6.56 (estimated values), 4, 1 and 0.62, from top to bottom. **(b)** Proportion of reinfection for the heterogeneous model with different distribution of the infection risk (γ, α_1): bottom full curve (0.937,0.069) (estimated values), dashed curve (0.937,0.5), dotted curve (0.5,0.069) and top full curve (1,1) (homogeneous distribution). Remaining model parameters according to Table 6.1.

level of endemicity, reinfection is very common and it surpasses reactivation. The shape of the (y, p) curve is mainly driven by the model structure and robust to parameter changes, unless a more extreme infection-risk distribution is chosen. However, if instead of considering

disease incidence on y-axis we consider transmission intensity (R_0) we observe that the range we obtain for R_0 can be very different.

Figure 6.4 illustrates how the proportion of reinfection (p) changes with the reproduction number (R_0) for different values of the reinfection factor (σ) for the homogeneous model (panel (a)) and for different assumptions on the infection risk distribution among population groups (panel (b)). For the homogeneous model, the reinfection factor determines the range of R_0 observed, maintaining the same curve shape. Now, fixing σ, the introduction of heterogeneity in the infection risk contributes to smoothen the reinfection impact as we have described for the simple SIRI model, in Section 5.4. It increases reinfection for low transmission settings due to the selection of more susceptible individuals and it reduces reinfection likelihood when compared to the homogeneous model (top curve in panel (b)), especially when common.

These observations have motivated us to use the estimated local R_0 as another criteria to compare the models, as discussed in Section 6.3. The best fitted parameters suggest a very narrow range of R_0 for the homogeneous model and considerable wider range for the heterogeneous one, which is in better agreement with published data (Trunz *et al.*, 2006).

Enhancement in susceptibility after successful treatment and bistability

Previous models have raised the possibility of the existence of multiple equilibria in tuberculosis, driven by the reinfection process (Feng *et al.*, 2000). However these were not consensual, since multiple equilibria existence depend on the assumption that already-infected individuals must be more likely to get TB and become infectious than uninfected individuals (Lipsitch & Murray, 2003). Here, under the assumption of homogeneous susceptibility to infection, enhancement in susceptibility after successful treatment can also lead to a multiple equilibria situation for $R_0 < 1$, provided that the reinfection factor σ, be sufficiently large. Moreover, independently of the other model parameters it can only happen assuming increased risk ($\sigma > 1$). Theorem 6.5.1 at the end of this section summarizes this result for system (6.1) with $\alpha_1 = \alpha_2 = 1$.

An important aspect of this result is its impact in the interpretation of the fitting results. For technical reasons, for the fitting procedure it is common to use the force of infection λ, as a linear parameter instead of the transmission coefficient β, as in the original formulation $\lambda = \beta I$ (Uys *at al.*, 2009; White *et al.*, 2007; Águas *et al.*, 2008). The correspondent dynamical system (system (6.1) with $\beta I \equiv \lambda$) becomes linear and the backward bifurcation is no longer possible. Hence, the estimation procedure can lead to parameter values that give unreasonable results for the nonlinear dynamical system, attributing observed incidences to unstable equilibria.

In this Chapter we used the same simplification aforementioned. Thoerem 6.5.1 allowed us to accept the results obtained by the fitting procedure, by posteriorly inspecting the condition for the existence of backward bifurcation for the estimated parameters. This subject certainly deserves further investigation.

Note that the obtained reinfection factor for the homogeneous model is close to the critical: $\sigma = 6.56 < 7.03 = \sigma_c$.

Theorem 6.5.1. *For system* (6.1) *with* $\alpha_1 = \alpha_2 = 1$ *a backward bifurcation at* $R_0 = 1$ *occurs if* $\omega < \delta\phi$ *for*

$$\sigma > \sigma_c = \frac{(\mu\phi + \omega)(\delta + \mu)(\omega + \mu + \tau)}{\mu(\mu(1-\phi) + \tau)(\delta\phi - \omega)} (> 1). \tag{6.13}$$

Remark 6.5.1. *Note that for the backward bifurcation to occur it is also required that the reactivation rate be lower than the rate of progression from primary infection directly to active TB ($\omega < \delta\phi$), which is reasonable from the epidemiological point of view.*

Proof. System (6.1) with $\alpha_1 = \alpha_2 = 1$ can be reduced to a system of 5 differential equations

$$\begin{cases} S' &= \mu - (\lambda + \mu)S \\ P' &= \lambda S + \sigma\lambda(L + R) - (\delta + \mu)P \\ I' &= \phi\delta P + \omega L + \omega R - (\tau + \mu)I \\ L' &= (1 - \phi)\delta L - (\sigma\lambda + \omega + \mu)L \\ R' &= \tau I - (\sigma\lambda + \omega + \mu)R \end{cases} \tag{6.14}$$

where $\lambda = \beta I$ is the force of infection. Since we assumed that the reactivation from recovered and latent classes is the same ω it can be furthered simplified into only 4 equations by collapsing these classes into one $RL = (R + L)$

$$\begin{cases} S' &= \mu - (\lambda + \mu)S \\ P' &= \lambda S + \sigma\lambda RL - (\delta + \mu)P \\ I' &= \phi\delta P + \omega RL - (\tau + \mu)I \\ RL' &= \tau I + (1 - \phi)\delta P - (\sigma\lambda + \omega + \mu)RL \end{cases} \tag{6.15}$$

Now by solving each equation, at equilibrium, in order to I and substituting in the third equation we get a polynomial of second order $P(\beta, I) = a_2(I)\beta^2 + a_1(I)\beta + a_0(I)$, for which the equilibrium solutions satisfy $P(\beta, I^*) = 0$, where I^* is the equilibrium proportion of infectious and

$$\begin{aligned} a_2 &= \sigma I[(\delta\phi + \mu + \tau)I - \phi\delta] \\ a_1 &= I[(\mu + \delta)(\omega + \tau + \mu) + \sigma\mu(\tau + \delta\phi + \mu) - \delta(\mu\phi + \omega)] \\ a_0 &= \mu(\mu + \delta)(\mu + \tau + \omega). \end{aligned}$$

Treating β as a function of I^*, differentiating this equation and setting $(\beta, I^*) = (\beta^*, 0)$, where β^* is such that $R_0 = 1$, we obtain

$$\frac{\mathrm{d}\beta}{\mathrm{d}I^*}|_{(\beta^*,0)} = \frac{\mu(\delta + \mu)(\mu + \tau + \mu)}{(\mu\phi + \omega)^3\delta^2}[\mu(\mu(1 - \phi) + \tau)(\omega - \delta\phi)\sigma + (\mu\phi + \omega)(\delta + \mu)(\omega + \mu + \tau)].$$

Hence, a backward bifurcation occurs if $\omega < \delta\phi$ for $\sigma > \sigma_c = \dfrac{(\mu\phi + \omega)(\delta + \mu)(\omega + \mu + \tau)}{\mu(\mu(1 - \phi) + \tau)(\delta\phi - \omega)}$.

In particular, we can show that if the bifurcation at $R_0 = 1$ is backward then $\sigma > 1$. Considering the case where $\omega < \delta\phi$, the critical value for σ increases with $\omega \in [0, \delta\phi[$ since

$$\frac{\partial\sigma_c}{\partial\omega} = \frac{-\omega^2 + 2\phi\delta\omega + \mu^2\phi + \mu\delta\phi^2 + \tau\mu\phi + \delta\phi\mu + \tau\delta\phi}{\mu(\delta\phi - \omega)^2(\mu(1 - \phi) + \tau)} > 0$$

for which the numerator is a polynomial of second degree $p(\omega)$ with negative second derivative that verifies $p(0) > 0$ and $p(\delta\phi) > 0$. Hence, the minimum values for σ_c is attained for $\omega = 0$,

$$min\{\sigma_c\}_{\omega\in[0,\delta\phi[} = 1 + \frac{\mu(\mu + \tau + \delta\phi)}{\mu\delta(1 - \phi) + \delta\tau}$$

which is still above one. □

Chapter 7

Conclusions and prospects

This thesis has been developed in two lines of work: heterogeneity and partial immunity. Tuberculosis has served as biological problem that driven the progress of the work. We present here how theses results provide several possibilities for further research and applications.

In Chapter 2, a simple framework is developed to extend the concept and computation of the Reinfection Threshold (RT) for infectious diseases with partial immunity, described by deterministic models. We interpret the RT as the transmission level above which is possible to sustain transmission in a partially protected population. For its computation we define the reinfection sub-model from the original one by separating reinfection from other immune processes. The RT marks important changes in the system dynamical behavior and in different occasions we saw that reinfection imposes limitations to interventions impact marked by the RT, such as for vaccination.

When more than one susceptibility level exists in the population, reinfection is no longer characterized solely by one threshold in transmission. Instead, multiple RT exist, associated to the contribution of each susceptibility class to the overall dynamics. Their impact to the model behavior can go unnoticed unless perturbed. Interventions that can alter the susceptibility profile of the population have the potential to reveal these RTs, by creating regions of bistability (section 5.5.2). The knowledge of their position can help to design better intervention programs, able to overcome classical interventions limitations as uniform vaccination programs.

This framework offers a systematic way to identify these regions of interest on the transmissibility axis. Furthermore, it is widely applicable to other infectious disease for which immunity is not fully protective.

Chapter 3 we address the problem of drug resistance in tuberculosis. This can be seen as an example of heterogeneity in the pathogen population. Reinfection in conjunction with heterogeneity proves to be an important factor in the determination of the tuberculosis epidemiological landscape. It imposes a new threshold for tuberculosis transmission, above which resistant strains dissemination is facilitated, superinfection threshold. Consequently, drug resistance control would benefit from a change in the interventions focus, from reducing drug acquisition to blocking transmission of specially fitted resistant strains, depending on the epidemiological setting.

Reinfection also opens the possibility of infections with more then one strain, with different

drug-susceptibility and from distinct lineages. The question about how these mixed infections can contribute to transmission motivated a change in model scale, from host-population models to within-host models. The goal is to try to reveal some of the mechanisms underling strain competition in mixed infections within the host and use those results together with the epidemiological model as a way to infer how the competition at individual level affects the spread of drug resistant strains at the population level. Chapter 4 constitutes a first effort to make the link between within and between host strain competition.

We were able to reproduce the patterns of drug-susceptibility described in the literature at the individuals level and link these different disease manifestations with the relative fitness of competing strains. Based on the duration of infection and on the strain frequency we reparameterized the epidemiological model. This way we could draw a theoretical epidemiological landscape depending of the relative within-host fitness of strains.

In this process, we were confronted with the lack of studies on the reinfection process, at the within-host level. So far, most of the experimental or modeling work concentrates itself in the immune response to primary infection or drug resistance acquisition. Thus, we confined our model to a stage where both strains are already present and active, ignoring the initial process (reinfection and latency). A more challenging task, would be to model the reinfection process and subsequent progression from latency to active disease. It is known that, during latency, most bacteria load remains isolated in complex cellular structures, the granulomas, formed by the immune response. However, little is known about the mechanisms that affect the immune response upon a new infection (reinfection). How these structures are affected by the new strain and if this enhances or impairs the immune response and the progression to active disease. Experimental knowledge is accumulating and new techniques allow a better understanding of the immune processes involved. Mathematical within-host models are suitable to test alternative hypothesis on the immune response to the challenge with a second strain, confronting the model results with the available data. Models that describe the immune response using dynamical systems, in the line of Gammack *et al.* (2005) provide an adequate framework that can be extended to this problem. With these more detailed models we hope to address not only the question of mixed infections but also to clarify the meaning of partial immunity on TB transmission.

Most of simple, theoretical models considers 'typical' individuals and 'average' behaviors assuming that the underlying heterogeneity does not affect significantly the model outcome. However, in some cases heterogeneity itself can be one of the problem determinants. In Chapter 5, we introduce heterogeneity in the host population by considering distinct groups with different susceptibility to infection. Infection tends to affect individuals at higher risk and as a consequence the high-risk individuals accumulate in the recovered class, implying a higher rate of reinfection among this group. One of the heterogeneity signatures we find is, in fact, an increased reinfection rate, in low to moderate transmission settings, for which reinfection is expected to be low under the hypothesis of partial immunity. The idea is applied for the particular case of tuberculosis in Chapter 6.

The simple framework proposed can capture the essence of heterogeneity while keeping a simple structure suitable of analytical treatment. Heterogeneity in other disease related processes can be investigated and its consequences to data interpretation drawn. Further complexity could be added to the model when the general behavior is understood, especially when dealing with concrete applications where data on heterogeneity is available.

Recurrent episodes of tuberculosis can be due to endogenous reactivation or exogenous reinfection, and the discrimination of these two processes is crucial to control planning. A number of studies, based on molecular typing of *Mtb*, have alerted for the relative contribution of reinfection (Wang *et al.*, 2007). It has been stipulated that, for some endemic regions, rates of reinfection tuberculosis after successful treatment are higher than rates of new tuberculosis (Verver *et al.*, 2005). And, more recently, it has been suggested that these observations are compatible with the hypothesis that individuals who had tuberculosis before become more susceptible to reinfection (Uys *at al.*, 2009). We apply the results from Chapter 5 and propose the selection hypothesis for tuberculosis. To reconcile high reinfection rates with the more consensual view that infection confers some degree of protection that reduces the individual susceptibility to reinfection, we postulate that some individuals are a priory more likely to develop the disease because they are more exposed or have some form of innate susceptibility. As infection tends to affect individuals at higher risk, the distribution of recovered individuals is skewed towards higher susceptibility inflating the rates of reinfection. Using a mathematical model, the two alternative hypotheses, which lead to contrasting conclusions regarding adaptive immunity in tuberculosis, are confronted with data from six regions representing distinct transmission intensities distributed worldwide. We show that only the selection hypothesis is compatible with previous estimates for the basic reproduction number and propose a criterion for further validation.

One of the most challenging proposals that come out of this work is the selection hypothesis. We hope to continue the work by further characterize the source of heterogeneity. HIV seems to be one of the possible candidates to differentiate the population into different infection susceptibility groups. It is known that coinfection with HIV is associated with increased risk for recurrent TB (Panjabi *et al.*, 2007). When clarified, which factor or factors discriminate the risk groups we can consider heterogeneity parameters for the different regions instead of assume them globally and independent of the epidemiological setting. What will probably contribute to a better agreement with data. Parallel to that, more data on the rates of reinfection and new infection must be gathered for different transmission intensities and confronted with the trend predicted by the model under the selection hypothesis.

Bibliography

Águas R, Gonçalves G, Gomes MGM. 2006. Pertussis: Increasing disease as a consequence of reducing transmission. *Lancet Inf Dis.* **6**:112-117.

Águas R, White LJ, Snow RW, Gomes MGM. 2008. Prospects for malaria eradication in sub-Saharan Africa. *PLoS ONE* **3**(3):e1767.

Alavez-Ramirez J, Castellanos JR, Esteva L, Flores JA, Fuentes-Allen JL, Garcia-Ramos G, Gomez G, Lopez-Estrada J. 2006. Within-host population dynamics of antibiotic-resistant M. tuberculosis, *Math Med Biol.* **24**(1):35-56.

Anderson H, Britton T. 1998. Heterogeneity in epidemic models and its effect on the spread of infection. *J Appl Prob.* **35**:651-61.

Anderson RM, May RM. 1985. Age-related changes in the rate of disease transmission: implications for the design of vaccination programmes. *J Hyg (Lond).* **94**(3): 365-436.

Anderson RM, May RM. 1991. Infectious diseases of humans. Oxford Science Publication, Oxford.

Antia R, Koella J, Perrot V. 1996. Models of the within-host dynamics of persistent mycobacterial infections, *Pro R Soc Lond B.* **263**:257-63.

Austin DJ, Anderson RM. 1999. Studies of antibiotic resistance within the patient, hospitals and the community using simple mathematical models. *Phil Trans R Soc Lond B.* **354**:271-38.

Austin JF, Dick JM, Zwarenstein M. 2004. Gender disparity amongst TB suspects and new TB patients according to data recorded at South African Institute of Medical Research laboratory for the Western Cape Region of South Africa. *Int J Tuberc Lung Dis.* **8**:435-39.

Austin DJ, Kakehashi M, Anderson RM. 1997. The transmission dynamics of antibiotic-resistant bacteria: the relationship between resistance in commensal organisms and antibiotic consumption. *Proc R So. B.* **264**:1629-38.

Austin DJ, White NJ, Anderson RM. 1998. The dynamics of drug action on the within-host population growth of infectious agents: Melding pharmacokinetics with pathogen population dynamics. *J Theor Biol.* **194**:313-39.

Ball F. 1985. Deterministic and stochastic epidemics with several kinds of susceptibles. *Adv Appl Prob.* **17**(1):1-22.

Boldin B, Diekmann O. 2008. Superinfections can induce evolutionarily stable coexistence of pathogens. *J Math Biol.* **56**(5):635-72.

Blower SM, Chou T. 2004. Modeling the emergence of the 'hot zones': tuberculosis and the amplification dynamics of drug resistance. *Nature Med.* **10**:1111-16.

Blower SM, Gerberding J. 1998. Understanding, predicting and controlling the emergence of drug tuberculosis: a theoretical framework. *J Mol Med.* **76**:624-36.

Blower SM, McLean AR, Porco TC, Small PM, Hopewell PC, Sanchez MA, Moss AR. 1995. The intrinsic transmission dynamics of tuberculosis epidemics.*Nat Med.* **1**(8):815-21.

Blower SM, Small PM, Hopewell PC. 1996. Control Stretagies for tuberculosis epidemics: New models for old problems. *Science.* **273**(5274):497-500.

Bonhoeffer S. 2002. Managing antibiotic resistance: What models tell us? In Diekmann, U., Metz, J. A. J., Sabelis,M. W., Sigmund, K., editors, Adaptive Dynamics and Infectious Diseases: "In Pursuit of Virulence Management", 326-338. Cambridge University Press, Cambridge, UK.

Boni MF, Feldman MW. 2005. Evolution of antibiotic resistance by human and bacterial niche construction. *Evolution.* **59**(3):477-91.

Breban, R, Blower, S. 2005. The Reinfection Threshold does not exist. *J Theor Biol.* **235**:151-2.

Britton, T. 1998. On critical vaccination coverage in multiple epidemics. *J Appl Prob.* **35**:1003-6.

Castillo-Chavez C, Feng Z. 1997. To treat or not to treat: the case of tuberculosis. *J Math Biol.* **35**:629-56.

Caswell H. 2001. "Matrix population models: Construction, Analysis, and Interpretation", Sinauer Ass., Sunderland, Mass., US.

Chiang CY, Riley LW. 2005. Exogenous reinfection in tuberculosis. *Lancet Infec Dis.* **5**:629-36.

Cohen T, Murray M. 2004. Modelling epidemics of multidrug-resistant *M. Tuberculosis* of heterogeneous fitness, *Nature Medicine.* **10**:1117-21.

Cohen T, Sommers B, Murray M. 2003. The effect of drug resistance on the fitness of *Mycobacterium tuberculosis*, *Lancet Infec Dis.* **3**:13-21.

Cosma C, Humbert O, Ramakrishnan L. 2004. Superinfecting mycobacteria home to established tuberculous granulomas. *Nature Immun.* **5**(8):828-35.

Cosma C, Humbert O, Sherman DR, Ramakrishnan, L. 2004. Trafficking of superinfecting *Mycobacterium* organisms into established granulomas occurs in mammals and is independent of the Erp and ESX-1 Mycobacterial virulence Loci. *JID.* **198**:1851-5.

Coutinho FAB, Massad E, Lopez LF, Burattini MN, Struchiner CJ, Azevedo-Neto RS. 1999. Modelling heterogeneities in individual frailties in epidemic models. *Math Comp Mod.* **30**:97-115.

Crofton J, Chaulet P, Mahaler D. 1997. Guidelines for the management of drug resistant tuberculosis. *WHO/TB/96.21* (Rev/1) p8.

de Boer AS, Borgdorff MW, Vynnycky E, Sebek MM, van Soolingen D. 2003. Exogenous re-infection as a cause of recurrent tuberculosis in a low-incidence area. *Int J Tuberc Lung Dis.* **7**:145-152.

de Boer AS, van Soolingen D. 2000. Recurrent tuberculosis due to exogenous reinfection. *N Engl J Med.* **342**:1050-1051.

Dhooge A, Govaerts W, Kuznettsov YA. 2003. MATCONT: A MATLAB package for numerical bifurcation analysis of ODE's. *ACM Trans. Math. Software* **29**:141-64.

Diekmann O, Heesterbeek JAP, Metz JAJ. 1990. On the definition and computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J Math Biol.* **28**:365-382.

Diekmann O, Heesterbeek JAP. 2000. Mathematical epidemiology of infectious diseases. Model building, analysis and interpretation. Wiley Series in Mathematical and Computational Biology. John Wiley & Sons, La., Chichester.

Dietz K. 1993. The estimation of the basic reproduction number for infectious diseases. *Stat Methods Med Res.* **2**:23-41.

Dye C, Espinal MA. 2001. Will tuberculosis become resistant to all antibiotics? *Proc R Soc Lond B.* **268**:45-52.

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

Dye C, Williams BG. 2000. Criteria for the control of drug-resistant tuberculosis. *Proc Natl Acad Sci USA.* **97**:8180-85.

Dye C, Williams B, Espinal MA, Raviglione M. 2002. Erasing the world's slow stain: strategies to beat multidrug-resistant tuberculosis, *Science.* **295**:2042-46.

Dushoff J. 1999. Host heterogeneity and disease endemicity: a moment-based approach. *Theor Pop Biol.* **56**:325-35.

Dushoff J. 1996. Incorporating immunological ideas in epidemiological models. *J Theor Biol.* **180**(3):181-7.

El Sahly et al. 2004. Recurrent tuberculosis in Houston, Texas: a population- based study. *Int J Tuberc Lung Dis.* **8**:333-340.

Espinal MA, Kim SJ, Suarez PG, Kam KM, Khomenko AG, Migliori GB, Baéz J, Kochi A, Dye C, Raviglione MC. 2000. Standard short-course chemotherapy for drug-resistant tuberculosis: treatment outcomes in 6 countries. *JAMA.* **283**(19):2537-45.

Feng Z, Castillo-Chavez C, Capurro A. 2000. A Model for Tuberculosis with Exogenous Reinfection. *Theor. Pop. Biol.* **57**:235-47.

Fine PE, Clarkson JA. 1982. Measles in England and Wales–II: The impact of the measles vaccination programme on the distribution of immunity in the population. *Int J Epidemiol.* **11**(1):15-25.

Gagneux S, long DC, Small P, Van T, Schoolnik GK, Bohannan BJM. 2006. The competitive cost of antibiotic resistance in *Mycobaterium tuberculosis*. *Science*. **312**:1944-46.

Galvani A, May RM. 2005. Dimensions of superspreading. *Nature*. **438**:293-5.

Gammack D, Ganguli S, Marino S, Segovia-Juarez J, Kirschner D. 2005. Understanding the immune response in tuberculosis using different mathematical models and biological scales. *Multiscal Model Simul.* **3**(2):312-45.

Gart JJ. 1968. The mathematical analysis of an epidemic with two kinds of susceptibles. *Biometrics.* **24**:557-66.

Gilchrist M, Coombs D. 2006. Evolution of virolence: Interdependence, constrains and selection using nested models. *Theor Popul Biol.* **69**(2):145-53.

Gökaydin D, Oliveira-Martins JB, Gordo I, Gomes MGM, 2007. Reinfection thresholds regulate pathogen diversity: The case of influenza. *Royal Society Interface*. **4**:137-142.

Gomes MGM, Rodrigues P, Hilker F, Mantilla-Beniers NB, Muehlen M, Paulo A, Medley G. 2007. Implications of partial immunity on the prospects for tuberculosis control by post-exposure interventions. *J. Theor. Biol.* **248**,(4):608-17.

Gomes MGM, White LJ, Medley GF. 2005. The reinfection threshold. *J Theor Biol.* **236**:111-13.

Gomes MGM, Franco AO, Gomes MC, Medley GF. 2004. The reinfection threshold promotes variability in tuberculosis epidemiology and vaccine efficacy. *Proc R Soc Lond B.* **271**:617-23.

Gomes MGM, White LJ, Medley GF. 2004. Infection, reinfection, and vaccination under suboptimal immune protection: epidemiological perspectives. *J Theor Biol.* **235**(2):151-2.

Heesterbeek JAP. 2002. A brief history of R0 and a recipe for its calculation. *Acta Biotheor.* **50**(3):189-204.

Hethcote H. 2000. The mathematics of infectious diseases. *SIAM Review.* **42**(4):599–653.

Hethcote H. 1996. Modeling heterogeneous mixing in infectious disease dynamics. In: Models for infectious diseases, their structure and relation to data. Cambridge University Press, p - 215-238.

Hoal EG, Lewis L-A, Jamieson SE, Tanzer F, Rossouw M, Victor T, Hillerman R, Beyers N, Blackwell JM, van Helden PD. 2004. SLC11A1 (NRAMP1) but not SLC11A2 (NRAMP2) polymorphisms are associated with susceptibility to tuberculosis in a high-incidence community in South Africa. *Int J Tuberc Lung Dis.* **8**:1464–1471.

Hyman MJ, Li J. 2000. An intuitive formulation for the reproductive number for the spread of diseases in heterogeneous populations. *Math Biosc.* **167**:65-86.

Hyman MJ, Li J. 2005. Differential susceptibility epidemic models. *J Math Biol.* **50**:626-44.

Iseman MD. 2002. Tuberculosis therapy: past, present and future. *Eur Respir J Suppl.* **36**:87-94.

Jasmer RM, Bozeman L, Schwartzman K, Cave MD, Saukkonen JJ, Metchock B, Khan A, Burman WJ. 2004. Recurrent tuberculosis in the United states and Canada: Relapse or reinfection? *Am J Respir Crit Care Med.* **170**:1360-1366.

Kermack WO, McKendrick AG. 1927. A contribution to the mathematical theory of epidemics. *Proc R Soc A.* **115**: 700-721.

Koopman JS, Simon CP, Riolo CP. 2005. When to control endemic infections by focusing on high-risk groups. *Epidemiology* **16**(5):621-27.

Lipsitch M, Levin B. 1997. The Population Dynamics of Antimicrobial Chemotherapy. *Antimicrob Agents Chemother.* **41**(2):363-73.

Lipsitch M, Levin B. 1998. Population dynamics of tuberculosis treatment: mathematical models of the roles of non.compliance and bacterial heterogeneity in the evolution of drug resistance. *Int J Tuberc Dis.* **2**(3):187-99.

Lipsitch M, Murray MB. 2003. Multiple equilibria: Tuberculosis transmission require unrealistic assumptions. *J Math Biol.* **63**:169-170.

Miller JC. 2007. Epidemic size and probability in populations with heterogeneous infectivity and susceptibility. *Phys Rev E.* **76**:010101.

Murray CJL, Salomom JA. 1998. Modeling the impact of global tuberculosis control strategies. *Proc Natl Acad Sci USA.* **95**:13881-86.

Murphy BM, Singer BH, Anderson S, Kirschner D. 2002. Comparing epidemic tuberculosis in demographically distinct heterogeneous populations. *Math Biosc.* **180**:161-85.

Murphy BM, Singer BH, Kirschner D. 2003. On treatment of tuberculosis in heterogeneous populations. *J Theor Biol.* **223**:391-404.

Nardell E, McInnis B, Thomeas B, Weidhaas S. 1986. Exogenous reinfection with tuberculosis in a shelter for the homless. *N Engl J Med.* **315**:1570-1575.

McNabb SJ, Braden CR, Navin TR. 2002. DNA fngerprinting of Mycobacterium tuberculosis: lessons learned and implications for the future. *Emerg Infect Dis.***8**:1314-1319.

Nikolaou M, Tam VH. 2006. A new modeling approach to the effect of antimicrobial agents on heterogeneous microbial populations. *J. Math. Biol.* **52**:154-82.

Pablos-Mendez A, Gowda DK, Friedman TR. 2002. Controlling multidrug-resistant tuberculosis and access to expensive drugs: a rational framework, *Bull. of WHO* **80**(6):489-95.

Panjabi R, Comstock GW, Golub JE. 2007. Recurrent tuberculosis and its risk factors: adequately treated patients are still at high risk. *Int. J. Tuberc. Lung Dis.* **11**(8):828-37.

Rodrigues P, Margheri A, Rebelo C, Gomes MGM. 2009. Heterogeneity in susceptibility to infection can explain high reinfection rates. *J Theor Biol.* **259**: 280-290.

Rodrigues P, Gomes MG, Rebelo C. 2007. Drug resistance in tuberculosis - a reinfection model, *Theor Pop Biol.* **71**: 196-212.

Smith DL, Dushoff J, Snow RW, Hay SI. 2005. The entomological inoculation rate and *Plasmodium falciparum* infection in Africa children. *Nature* **438**(24):492-5.

Smith PG, Moss AR. 1994 in Tuberculosis: Pathogenesis, Protection and control, ed Bloom, BR (ASM Press, Washington), pp 47-59.

Sonnenberg P, Murray J, Glynn JR, Shearer S, Kambashi B, Godfrey-Faussett P. 2001. HIV-1 and recurrence, relapse, and reinfection of tuberculosis after cure: a cohort study in South African mineworkers. *Lancet.* **358**:1687-1693.

Stollenwerk N, Martins J, Pinto A, 2007. The phase transition lines in pair approximation for the basic reinfection model SIRI. *Phys Lett A.* **371**:379-388.

Story A, Murad S, Verheyen M, Roberts W, Hayward AC. 2007. Tuberculosis in London: The importance of homelessness, problem drug use and prison. *Thorax* **62**:667-71.

Sutherland I, Svandová E, Radhakrishna S. 1982. The development of clinical tuberculosis following infection with tubercle bacilli. *Tuberclo.* **63**:255-68.

Trunz BB, Fine PEM, Dye C. 2006. Effect of BCG vaccination on childhood tuberculous meningitis and miliary tuberculosis worldwideÊ: a meta-analysis and assessment of cost-effectiveness. *The Lancet.* **367**Ê:1173-1180.

Urlichs T, Kaufmann S. 2002. Mycobacterial persistence and immunity. *Front Biosci.* **7**(5):458-69.

Uys P, van Helden PD, Hargrove JW. 2009. Tuberculosis reinfection rate as a proportion of total infection rate correlates with the logarithm of the incidence rate: a mathematical model. *J R Soc Interface.* **6**:11-15.

van Boven M, de Melker HE, Schellekens JFP, Kretzschmar M. 2001. A model based evaluation of the 1996-7 pertussis epidemic in the Netherlands. *Epidemiol Infect.* **127**:73-85.

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

van Rie A, Richardson M, Johnson R, van der Spuy GD, Murray EJ, Beyers N, van Pittius NC, van Helden PD, Warren RM. 2004. Reinfection and mixed infection cause changing *Mycobacterium tuberculosis* drug-resistance patterns. *Am J Respir Crit Care Med.* **172**(5):636-42.

van Rie A, Warren R, Richardson M, Victor T, Gie R, Enarson DA, Beyers N, van Helden PD. 1999. Exogenous reinfection as a cause of recurrent tuberculosis after curative treatment. *N Engl J Med.***341**:1174-79.

Veliov VM. 2005. On the effect of population heterogeneity on dynamics of epidemic diseases. *J Math Biol.* **51**:123-43.

Verver S, Warren R, Beyers N, Richardson M, van der Spuy G, Borgdorff MW, Enarson DA, Behr MA, van Helden PD. 2005. Rate of reinfection tuberculosis after successful treatment is higher than rate of new tuberculosis. *Am J Respir Crit Care Med.* **171**:1430-35.

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

Wang JY, Lee LN, Lai HC, Hsu HL, Liaw HL, Hsueh PR, Yang PC. 2007. Prediction of the tuberculosis reinfection proportion from the local incidence. *JID.* **196**(2):281-8.

Warren RM, Victor CT, Streicher ME, Richardson M, Beyers N, van Pittius NC, van Helden PD. 2004. Patients with active Tuberculosis often have different strains in the same sputum specimen. *Am J Respir Crit Care Med.* **169**:610-14.

Webb G, D'Agata E, Magal P, Ruan S. 2005. A model of antibiotic-resistant bacterial epidemics in hospitals. *PNAS.* **102**(37):13343-8.

White LJ, Mandl JN, Gomes MGM, Bodley-Tickell AT, Cane PA, Perez P, Siqueira MM, Portes SA, Straliotto SM, Waris M, Medley GF, Nokes DJ. 2007. Understanding the transmission dynamics of RSV: Multiple time series and nested models. *Math Biosc.* **209**:222-239.

Wiggins S. 1990. Introduction to Applied Nonlinear Dynamical Systems and Chaos, Springer, Berlin.

WHO/IUATLD Global Working Group on Antituberculosis Drug Resistance Surveillance. 1998. Guidelines for surveillance of drug resistance in tuberculosis, WHO Geneva/IUATLD Paris. *Int J Tuberc Lung Dis.***2**:72-89.

WHO/IUATLD 2004. Global Project on Anti-tuberculosis Drug Resistance Surveillance. Anti-Tuberculosis Drug Resistance in the World: third global report, WHO/HTM/TB/2004.343. Geneva, Switzerland.

World Health Organization 2005. Global Tuberculosis Control: Surveillance, Planning, Financing. WHO/HTM/TB/2005.349. Geneva, Switzerland.

World Health Organization. 2007. The Global MDR-TB and XDR-TB response plan 2007-2008. Geneva, World Health Organization.

World Health Organization. 2009. Global tuberculosis control - epidemiology, strategy, financing. WHO/HTM/TB/2009.411. Geneva, Switzerland.

Woolhouse MEJ, Dye C, Etard JF, Smith T, Charlwood JD, Garnett GP, Hagan P, Hii JLK, Ndhlovu PD, Quinnell RJ, Watts CH, Chandiwana SK, Anderson RM. 1997. Heterogeneities in the transmission of infectious agents: implications for the design control programs. *Proc Natl Acad Sci USA.* **94**:338-42.

Yew WW, Leung CC. 2005. Are some people not safer after successful treatment of tuberculosis? *Am J Respir Crit Care Med.* **171**(12):1324-5.

List of Figures

List of Tables

Abbreviations

TB	Tuberculosis	pag. 11
SIS	Susceptible-Infectious-Recovered model	pag. 11
SIR	Susceptible-Infectious-Susceptible model	pag. 11
RT	Reinfection threshold	pag. 12
WHO	World Health Organization	pag. 13
HIV	Human immunodeficiency virus	pag. 13
BCG	Bacille Calmette-Guérin vaccine	pag. 13
Mtb	*Mycobacterium tuberculosis*	pag. 13
SIRI	Susceptible-Infectious-Recovered-Infectious model	pag. 14
MDR	Multi-drug resistant	pag. 34
XDR-TB	Extensively drug-resistant tuberculosis	pag. 34
DOTS	Directly Observed Treatment Short-course program	pag. 50
DOTS-Plus	Extension of DOTS program	pag. 50

(in order of appearance in the text)