A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

http://wrap.warwick.ac.uk/166149

Copyright and reuse:
This thesis is made available online and is protected by original copyright.
Please scroll down to view the document itself.
Please refer to the repository record for this item for information to help you to cite it.
Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications

Modelling the interactions between *Mycobacterium tuberculosis* and innate alveolar macrophages

by

Cameron Lack

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in

Mathematics for Real-World Systems

University of Warwick

EPSRC & MRC Centre for Doctoral Training

in Mathematics for Real-World Systems

March, 2021

Contents

List of Figures xi

List of Tables xxv

Acknowledgments xxix

Declaration xxxi

Abstract xxxiii

1 Introduction 1

- 1.1 Global epidemiology 2
- 1.2 Understanding the within-host dynamics 5
- 1.3 Importance of early dynamics 9
- 1.4 Disease progression 9
- 1.5 Description of experiment 10
- 1.6 Methodology 14
  - 1.6.1 Mechanistic modelling 15
  - 1.6.2 Ordinary differential equations 15
  - 1.6.3 Agent based modelling 17
  - 1.6.4 Parameter inference 18
  - 1.6.5 Sobol sensitivity analysis 19
  - 1.6.6 Goodness of fit 25
  - 1.6.7 Identifiability 26
- 1.7 Literature review 27
  - 1.7.1 Biology and pathogenesis of tuberculosis infections 27
  - 1.7.2 Mathematical modelling of tuberculosis infections 35

1.8 Thesis Summary 39

2 Implicit intracellular bacteria 41

2.1 Model description 41

2.1.1 Mechanisms 42

2.1.2 Model definition 44

2.2 Data 47

2.3 Parameter Fitting 49

2.3.1 Macrophage division 49

2.3.2 Macrophage death 50

2.3.3 Bacteria growth 51

2.3.4 Infection 53

2.3.5 Infected macrophage division 56

2.3.6 Parameter values 58

2.4 Model results 59

2.4.1 Model output 59

2.4.2 Sensitivity 64

2.4.3 Identifiability 68

2.5 Conclusions 68

3 Intracellular bacteria and Erlang growth 71

3.1 Erlang growth 72

3.2 Model description 74

3.2.1 Mechanisms 74

3.2.2 Model definition 75

3.3 Data 80

3.4 Parameter Fitting 80

3.4.1 Macrophage birth death process 81

3.4.2 Intracellular bacteria growth 82

3.4.3 Infection 84

3.4.4 Infected macrophage division 84

3.5 Model results . . . 84
3.5.1 Model output . . . 86
3.5.2 Sensitivity . . . 88
3.5.3 Identifiability . . . 92
3.5.4 Compartment choice post infection . . . 92
3.5.5 Predictions . . . 95
3.6 Conclusions . . . 99

4 Stochastic agent based model 103
4.1 From continuous to discrete populations . . . 103
4.2 Data . . . 105
4.3 Extracellular growth . . . 107
4.4 Intracellular growth . . . 108
4.5 Infection rate . . . 109
4.6 Macrophage birth death process . . . 111
4.7 Infected macrophage division . . . 120
4.8 Model parameters . . . 121
4.9 Model . . . 123
4.10 Model results . . . 126
4.10.1 Model output . . . 126
4.10.2 Sensitivity analysis . . . 129
4.11 Computational experiments . . . 132
4.11.1 Effect of intracellular growth rate . . . 132
4.11.2 Effect of probability of growth inhibition . . . 134
4.12 Conclusions . . . 135

5 Stochastic model with aggregated extracellular bacteria 139
5.1 Data . . . 140
5.2 Model description . . . 140
5.2.1 Mechanisms . . . 141

5.3 Parameter Fitting 142
    5.3.1 Aggregated extracellular growth 142
    5.3.2 Infection 143
    5.3.3 Post infection death rate 148
    5.3.4 Infected macrophage divisions process 149
    5.3.5 Maximum number of intracellular bacteria 151
5.4 Model definition 157
5.5 Model results 158
    5.5.1 Time series 158
    5.5.2 Validation 160
    5.5.3 Sensitivity analysis 165
    5.5.4 Experiments 168
5.6 Conclusions 172

6 Discussion **175**

6.1 Literature review 177
6.2 Non spatial model 178
6.3 Infection process 179
6.4 Macrophage division 182
6.5 Data 183
6.6 Model fitting 184

7 Conclusions **187**

A Colour coded equations **191**

B Model code for chapter 4 **193**

C Additionally considered mechanisms **197**

C.1 Load dependent death rate 197
    C.1.1 Investigating the relationship of d_a and d_b 198
    C.1.2 Weibull 200

C.2 Infection uptake distribution 202
  C.2.1 Binomial distribution 202
  C.2.2 Uniform distribution 203
  C.2.3 Truncated uniform distribution 204

Abbreviations **205**

Bibliography **207**

List of Figures

1.1 The distribution of the estimated global incidence of tuberculosis (TB) (**left**) compared to the distribution of the global population (**right**) as reported in the Global Tuberculosis Report 2019 [1]. The 5 countries with the highest absolute incidence account for roughly 55% of the the global incidence, while only accounting for 45% of the human population. The incidence is estimated per 100,000 population, predominantly from case notification data combined with either expert opinion or standard adjustment. With the exception of China, all of these countries are significantly over-represented for burden of TB demonstrating the imbalance in how the disease is managed globally. 3

1.2 The estimated incidence of TB is not uniformly distributed globally. South East Asia and sub Saharan Africa demonstrate the highest levels of active disease (**top**) [1]. As coinfection with Human Immunodeficiency Virus (HIV) is a significant risk factor, the global incidence of HIV per 100,000 population is also shown (**bottom**) [12]. This risk is particularly evident in southern Africa where HIV prevalence in new and relapsed TB cases can exceed 50% [1]. Figures produced using data from references [1, 12]. 4

1.3 (On previous page) TB exhibits a wide variety of responses in the host, as well as heterogeneity in its route from exposure to active infection. The exposure process can be broadly categorised into three stages: natural clearance, adaptive immune control, and active disease (**top**). An interferon-γ release assay (IGRA) test is able to differentiate between innate clearers and adaptive clearers [19] however an exposed individual that does not manage to clear the infection can progress to active disease either directly (primary TB infection) or after containing the infection for a period as latent TB infection (LTBI) (secondary TB infection) (**bottom**). Figure adapted from references [5, 20, 21]. 8

1.4 A snapshot of some macrophages, one of which contains intracellular bacteria (circled in red). The level of bacteria is measured as area of fluorescence since individual bacteria cannot be visually resolved. . . 11

1.5 Macrophages and bacteria are introduced on a Petri dish and the interactions between them are tracked. Of particular interest is the infection of macrophages, intracellular growth of bacteria, and eventual release of bacteria back into the extracellular space. 12

1.6 Progressive iterations of the Saltelli sampler fill in the parameter space with the goal of covering as much of the space as possible with as few parameter sets as possible, allowing the calculation of the global sensitivity indices to converge quickly. Shown is the coverage of the parameter space for a two parameter model. The last two frames (bottom middle and bottom right) show how the Saltelli sampler compares with the Latin Hypercube sampler. 22

1.7 Sampling the hypercube $[-\pi, \pi]^3$ using the Saltelli sampler and evaluating the Ishigami function over the resultant samples. The displayed output is passed to the Sobol sensitivity analysis along with the output parameters to determine the sensitivity of the model to each of the parameters. (**Top**) The Saltelli sampler evenly samples all three dimensions with good coverage even when projecting on to lower dimensions. (**Bottom**) The Ishigami function displays strong non-linearity and non-monotonicity with respect to the input parameters. 23

1.8 Sobol sensitivity indices calculated for the Ishigami function. Parameter x_1 has the most effect on the model output. Parameter x_2 is naturally independent of x_1 and x_3. Parameter x_3 only affects the model output in conjunction with x_1. 24

2.1 Schematic of the model showing the mechanisms included and the possible interactions between cell types. Healthy macrophages get infected by bacteria and become infected macrophages. Macrophages follow a birth (division) death cycle, while bacteria grow exponentially. Infected macrophages divide to produce both infected and healthy macrophages. And finally, when infected macrophages die, they release their intracellular load back into the extracellular space. . 44

2.2 Raw data of extracellular bacteria growth. Note the log scale on the y-axis. Clearly these data are non-stationary and the constant trend in the log scale demonstrates exponential growth. 48

2.3 Division wait times for both healthy and infected macrophages are recorded, and demonstrate distinct distributions. Despite a poor fit, for simplicity of the model an exponential distribution is assumed and fitted using maximum likelihood and shown against the data along with its 95% confidence interval. Better fitting distributions will be investigated in future chapters. 50

2.4 Birth to death wait times for both healthy and infected macrophages are recorded, and demonstrate similar distributions. Exponential distributions fit the data well. Since birth rates are distinct for both healthy and infected macrophages, death times will be kept distinct. . 51

2.5 Data show exponential growth of extracellular bacteria, to which an exponential growth curve can be fitted. Averaging over the fitted exponential growth rates weighted by their errors results in the kernel density estimate. This distribution is used to find the optimal value for the extracellular growth rate. 52

2.6 Distribution of time to infection wait times for healthy macrophages that were observed to become infected. Time to infection is the time of infection subtract time of birth. Macrophages that were born infected are not included. 55

2.7 Likelihood function of the rate of infection β given birth and infection times of macrophages. 55

2.8 Frequency of infected cell division which results in 1 or 2 infected daughter cells. 57

2.9 Likelihood and 95% confidence interval for probability of infected-infected division. 57

2.10 Integration of equations 2.1 to 2.3. Solid lines represent parameters set to their maximum likelihood estimates (MLEs), while the shaded area is created by varying the parameters within their confidence intervals. Parameter sets were created using latin hypercube sampling (LHS) with 2000 sets of parameters chosen. The initial conditions for the model are $M_H = 9$, $M_I = 1$ and $E = 1$, this is in contrast to the experimental initial conditions in which macrophages were infected with a multiplicity of infection of 1. 60

2.11 **Left** Base model as seen in figure 2.10 without the confidence intervals. **Centre** Model output without intracellular growth: $N = 1$ and $p = 1$. **Right** Model output without extracellular growth: $\alpha_E = 0$. The output of the model in the absence of extracellular growth shows the overwhelming effect that the implicit intracellular growth has on the model. 61

2.12 As in figure 2.10, integration of equations 2.1 to 2.3. Solid lines represent parameters set to their MLEs, while the shaded area is created by varying the parameters within their confidence intervals. Parameter sets were created using LHS with 2000 sets of parameters chosen. The initial conditions for the model are $M_H = 9$, $M_I = 1$ and $E = 1$. The parameter N has been updated to the value of 2, which yields the most qualitatively realistic model output. 63

2.13 Sobol sensitivity analysis measuring how much variance in the model populations can be attributed to variance in the model parameters. Populations are sampled at evenly spaced time points to investigate the sensitivity to parameters throughout the whole time course. . . . 67

2.14 Sobol sensitivity analysis bounding the parameters within their 95% confidence intervals. Measured is the fraction of the variance in the value of the time at which $\frac{\mathrm{d}E}{\mathrm{d}t}$ switches sign from negative to positive, which can be attributed to variance in the given parameter. 68

3.1 A non-homogeneous Erlang process can be well approximated by a homogeneous Erlang process. (**Top**) A four stage process consisting of varied transition rates. (**Centre**) Wait times to progress through all four stages are sampled and plotted. These wait times are used to fit a homogeneous process. (**Bottom**) The three stage process with homogeneous transition rates is a good approximation for the four stage process. 73

3.2 Illustrative example of the macrophage compartments M_k^i where k is the number of intracellular bacteria and i is the Erlang stage of life. As the macrophages progress through Erlang stages of life towards division, they progress through compartments $M_k^i \rightarrow M_k^{i+1}$. Similarly, as the intracellular bacteria grow, the macrophages progress through compartments $M_k^i \rightarrow M_{k+1}^i$. When a macrophage divides, it will be removed from the system and two new daughter cells are born in its place, each with an Erlang age of zero. If the parent cell was infected, the intracellular load is divided between the daughter cells as described in section 3.4.4 79

3.3 Raw intracellular growth time series data. Note the log scale on the y-axis. There are very few data points to consider for the intracellular growth, so all will be included. 80

3.4 Time to division data for both healthy and infected macrophages. Erlang distributions are fitted and 95% confidence intervals are calculated. The plotted confidence interval is obtained by evaluating the maximum and minimum probability for each division time over all parameters within the 95% confidence interval. 82

3.5 Data show exponential growth of intracellular bacteria, to which an exponential growth curve can be fitted using non-linear least squares (**left**). Averaging over the fitted exponential growth rates weighted by their errors results in the kernel density estimate. This distribution is used to find the optimal value for the intracellular growth rate (α_I) (**right**). Note, the orange experiment could be considered an outlier, since the observed growth rate is considerably lower than the others, however, since the data set it small, it will not be removed. The effect of α_I on the outcome of the model will be explored in section 3.5.2. . 83

3.6 Integration of equations 3.1 to 3.6. Solid lines represent parameters set to their MLEs, while the shaded area is created by varying the parameters within their confidence intervals. Parameter sets were created using LHS with 2000 sets of parameters chosen. 87

3.7 Despite the extracellular growth rate being zero, the population of extracellular bacteria can still be maintained through intracellular growth. The resultant population however is an order of magnitude lower than when the bacteria are able to grow extracellularly. 88

3.8 Sobol sensitivity analysis measuring how much variance in the model populations can be attributed to variance in the model parameters. Populations are sampled at evenly spaced time points to investigate the sensitivity to parameters throughout the whole time course. . . . 89

3.9 Sobol sensitivity analysis bounding the parameters within their 95% confidence intervals. Measured is the fraction of the variance in the value of the time at which $\frac{\mathrm{d}E}{\mathrm{d}t}$ switches sign from negative to positive, which can be attributed to variance in the given parameter. 92

3.10 The two models simulated are: **a)** the base model: healthy macrophages that become infected always move to the first Erlang compartment of infected macrophages, **b)** the alternative model: healthy macrophages that become infected are equally distributed across Erlang compartments. The change in model has a negligible effect on the extracellular population, but results in a small increase in the population of infected macrophages. In the base model, the time until division for newly infected macrophages is longer than for those in the alternative model, which results in the observed increase in population of infected macrophages. 94

3.11 Evaluating the time taken for $\frac{\mathrm{d}E}{\mathrm{d}t}$ to change sign from negative to positive for a range of infection rates (β). 96

3.12 By ranging the parameters controlling intracellular growth (α_I) and the probability a macrophage kills an extracellular bacteria, the effect on the resultant population of extracellular bacteria can be measured. Shown is the ratio of bacteria in the alternative model compared to the base model. Note that the colour bar scale is a composite of two linear scales, centered at 1. The value 1 represents that the outcome of the model is the same as the base model. < 1 represents an improvement, and > 1 represents a deterioration in outcome. 99

4.1 Data comprises of individual experiments which track the level of fluorescence from extracellular bacteria over time. Exponential growth curves are fitted (**left**) to the distinct time series data. Each fitted growth rate has an associated error. Taking the fitted value as the mean, and the error as the variance on the mean, a Gaussian kernel densitiy estimate (KDE) can be generated (**right**). The maximum of the KDE is taken as the MLE for the growth rate. 108

4.2 (**left**) Exponential growth curves fitted to time series data of fluorescence from intracellular bacteria. (**right**) Guassian KDE generated from fitted intracellular growth rates and their associated variances. The maximum of the KDE is taken as the MLE for the intracellular growth rate. 109

4.3 Raw data of time from birth to infection for macrophages that become infected during the window of experimentation. 110

4.4 Likelihood function for the rate of infection β calculated using equation 4.6. 111

4.5 An Erlang distribution is fitted to division times and an exponential distribution to death times. Running a stochastic simulation to compare the distribution of birth and death times of the model to the data shows that the model fails to accurately capture death times observed. This is due to the fitting process not considering the dependency between the two mechanisms. 112

4.6 A network representation of the continuous time Markov chain (CTMC) governing the birth death process of a single generation of macrophages. Growth is assumed to follow an Erlang process while the death rate is constant across all stages of life. Macrophages grow by moving through the N_g stages at a fixed rate g and finally transitioning to state G. A cell that reaches state G will divide removing itself from the system to be replaced by two new cells in state 1. Macrophages die by moving from any stage of life to state D. A cell that enters state D is removed from the system and represents a dead cell. 113

4.7 Using a CTMC to capture the joint distribution of Erlang growth and exponential death still fails to fully capture the distribution of death times. 115

4.8 Schematic of the birth death process of macrophages. Macrophages division and death follow two Erlang processes in parallel with parameters (g, N_g) and (d, N_d) for the birth and death processes respectively. 117

4.9 Using a CTMC using Erlang distributions for both growth and death results in a good fit to the data. 118

4.10 95% confidence regions for the parameters governing the birth death processes of healthy and infected macrophages. The 95% confidence regions are calculated by evaluating the likelihood function over a wide range of parameters and identifying the set of parameters which cover 95% of the total likelihood. 119

4.11 Likelihood of observing a infected macrophage dividing into two infected daughter cells. 120

4.12 The core mechanisms modelled in this chapter are depicted in this schematic. 126

4.13 1000 repetitions of the stochastic model with initial conditions of ten macrophages and one bacteria which is initialised intracellularly for 50% of the runs and extracellularly for the other 50%. 127

4.14 Result of running the model presented in chapter 3 with the parameters identified within this chapter. 128

4.15 Sobol sensitivity analysis measuring the fraction of the variance in the modelled populations over time that can be attributed to variance in the given parameters. 130

4.16 Sobol sensitivity analysis measuring the time taken for the extracellular bacteria population to exceed ten. Since stochastic variation will add noise, the time until the gradient of extracellular growth is positive cannot be used. An extracellular population of 10 is large enough that the chance of being significantly reduced is negligible, while being small enough that the simulation is quick, thus a large number of parameters can be sampled. 132

4.17 The relationship between the resultant extracellular bacteria population and the intracellular growth rate is approximately exponential. For each value of α_I, the simulation was run 100 times. The displayed points are the average of those simulations. 133

4.18 A small increase in the intracellular growth rate requires a large increase in the probability to inhibit intracellular growth in order for the outcome not to change. Note the non-linear colour bar centered at 1. 134

5.1 Exponential growth curves fitted to the aggregated extracellular growth experiments. 143

5.2 The individual experiments in the data set for this chapter show a lot of activity regarding infection times of the macrophages. In this plot each coloured group represents an experiment, and each line represents a cell that becomes infected at some point during the experiment. Of the ten experiments that tracked macrophages at the individual level, there were 9 that observed infection events, four of which demonstrate a lot more activity than the other five. Since macrophages move around while infected, more infections early on the experiment will rapidly increase the number of aggregates available, thus further increasing the number of observed infection events. . . . 145

5.3 The number of aggregates in an experiment is not measured, so the infection rate is bootstrapped over the possible number of aggregates. When a macrophage dies, a new aggregate is formed. When a macrophage becomes infected, there is a Bernoulli process which determines whether the macrophage consumes the entire aggregate, resulting in the number of aggregates going down, or the macrophage only partially consumes the aggregate, resulting in no change to the number of aggregates. This plot shows the an example experiment and how the possible number of aggregates changes according to events within a single experiment. Along the x axis an I represents a healthy macrophage internalises either all or part of an aggregate, and a D represents the death of a macrophage. 146

5.4 Likelihood landscape for the rate of infection (β) and the probability of a macrophage partially internalising an extracellular aggregate (p_{partial}). 147

5.5 Scatter plot of the number of initially internalised bacteria compared to whether that cell died within the next eight hours. The data is binned into 21 bins of equal width order in better to visualise the trend. Using least mean squared difference fitting results in an MLE of $a = 64.03$. 149

5.6 The maximum log likelihood as a function of each of the parameters governing the infected macrophage division process. $p_{\text{ratio}} = 0.86$ shows that most division events will be of the clumped form. $p_{\text{clump}} = 0.008$ shows that clumped divisions are dominated by completely infected-healthy divisions. Finally, $p_{\text{spread}} = 0.4941$ shows that spread divisions divide the intracellular load on average equally between the daughter cells. 151

5.7 Data showing the relationship between the time spent infected and the intracellular load at the time of the cells death. There are a significant number of cells that contain more than the assumed maximum load. Cells infected from birth tend to be born with a small number of bacteria, resulting in higher loads more commonly being observed in cells that were infected for a long time. Cells that become infected have a more dispersed distribution, many dying with a high load only shortly after they became infected. 152

5.8 Measuring the effect of the maximum intracellular load on the resultant extracellular bacteria population (see equation 5.3). There is a very small decrease of approximately 2% when increasing the maximum intracellular load from the assumed 50 to the observed 600. The baseline is calculated from the average of 5000 runs, while the compared values are averaged from 1000 runs due to computational limitations. The error bars are a bootstrapped 95% confidence interval and the fitted line is a simple linear regression, again shown with bootstrapped 95% confidence interval. (**left**) Linear x scale up to the observed 600. (**right**) Log x scale increasing the limit beyond feasibility to observe change. 154

5.9 6000 realisations of the model with parameters kept at their MLEs. Solid lines show the mean of the simulations and the shaded regions show the standard deviation. At 200 simulated hours, the healthy macrophages have sustained their population, but are unable to control the infection and extracellular bacteria are growing exponentially. 160

5.10 Comparison of distribution of wait times for macrophage birth death processes between the model and the data. The division process fits well, while the death process has a flatter distribution in the data than the model, especially for infected macrophages. 161

5.11 Distribution of wait times from a macrophage being born to that macrophage becoming infected. The model skews to the left when compared to the flatter distribution observed in the data. This is likely a result of the homogeneous mixing assumption of a non-spatial modelling. Since all agents are always able to interact, interactions are modelled as a Poisson process thus having exponential wait times. 163

5.12 The intracellular load at the time of infection and the time of death is measured and compared. (**left**) The distribution of loads at the time of infection are close, with the model slightly underestimating, but still able to capture the tail. (**right**) The distribution of the intracellular load at the time of death is significantly heavier in the tail than the data. 164

5.13 Sobol sensitivity analysis measuring the fraction of the variance in the modelled populations over time that can be attributed to variance in the given parameters. 166

5.14 Sobol sensitivity analysis tracking the variation in the number of simulated hours required for the extracellular population to exceed 1000 cells. The vast majority of variation is attributed to the extracellular growth rate, with the intracellular growth rate and the infected macrophage death rate having the next highest contributions. 168

5.15 The intracellular growth rate and the rate of infection are allowed to vary within approximately ±50% of their MLE values. The resultant extracellular population after 200 simulated hours is compared to the base model as a ratio of logs (see equation 5.4). An increase to the intracellular growth rate requires a larger proportional increase in the rate of infection to offset the increased rate of bacteria growth. 170
5.16 A new mechanism is introduced to allow macrophages to inhibit any intracellular growth according to a fixed probability. This probability and the rate of infection are allowed to vary: the probability of growth inhibition varies completely between 0 and 1, while the rate of infection varies within ±50% of its MLE. Note the y-axis is flipped to mirror figure 5.15. 172

C.1 The result of fitting d_a, d_l and g when d_b is fixed to some value between 0 and 1. 198
C.2 Result of multiple optimisations forcing $d_b = 1$, i.e the death rate is linearly dependent on $I(t)$. 199
C.3 Fitting Gompertz model with $d_b < 0.01$. 199
C.4 Fitting Weibull over a range of values for d_b. 200
C.5 Output of the optimisation for a range of values of d_l where the death rate $d(I) = d_l$ is a constant. 201
C.6 Scatter plot of the data compared to example samples data drawn from a binomial distribution with $p = 0.2371$. Data is shown in blue, the proposed model is shown in red. 202
C.7 The uniform distribution overestimates the number of bacteria internalised when the extracellular population gets large. 203
C.8 The truncated uniform distribution provides a better qualitative fit to the data than the uniform distribution. 204

List of Tables

1.1 An individual experiment is recorded as a series of snapshots of the Petri dish stitched together to produce a movie of the interactions between the macrophages and the bacteria. The bacteria are genetically tagged with a green phosphorescent gene, and the images are passed through a green light filter in order to enhance the experimenters ability to identify bacteria. Over the course of the experiment individual macrophages are tracked and the time of important events is noted, as well as whether the macrophage died or divided. When a macrophage divides, its daughter cells are tracked under new names, which are defined as the name of the parent cell with a '0' or a '1' concatenated to the end. 13

1.2 Example data showing the time series of the intracellular load for a single cell in a single experiment. 14

1.3 Estimated growth and death rates of *Mycobacterium tuberculosis* (Mtb) measured experimentally from the biological literature. 28

1.4 Estimated growth rates of intracellular Mtb measured experimentally from the biological literature. 33

2.1 Variable and parameter symbols and their descriptions for equations 2.1 to 2.3. 46

2.2 Descriptive statistics for data used in this chapter. The values used here are the wait times, thus the time from when a macrophage was born to when it either divided or died. 48

2.3 Contingency table for division of infected macrophages, investigating the relationship between the form of a cells division and that of their parents. Fishers exact test results is $\chi = 0.5962$ and thus not significant. 56

2.4 Table of parameter values and confidence intervals. Parameter N was taken from literature [151] so no confidence interval is calculated. Note that since the values of g_H and d_H (respectively g_I and d_I) are close, the population of macrophages will remain approximately constant. . 59

3.1 Variable and parameter symbols and their descriptions for equations 3.1 to 3.6. 76

3.2 Model parameters, their maximum likelihood estimate (MLE) values, and 95% confidence intervals. Note confidence intervals for parameters governing Erlang growth of healthy and infected macrophages are shown separately as parameters spaces in table 3.3. 85

3.3 (**Left**) 95% confidence intervals for Erlang growth parameters for healthy macrophages. (**Right**) 95% confidence intervals for Erlang growth parameters for infected macrophages. 85

4.1 Summary of the data relating to intra- and extracellular bacteria growth. There is significantly more data relating to the growth of intracellular bacteria, however the central tendency of the extracellular bacteria suggests a higher growth rate over the intracellular. 106

4.2 Summary of the data relating to the macrophages. Macrophages exhibit very high variance for all three measures shown as well as ranges that cover almost the full timespan of the experiments. 107

4.3 Model parameters, descriptions, maximum likelihood values and 95% confidence intervals. Note that the confidence intervals for the birth death processes are parameter spaces, and are better captured in a contour plot. These are shown in figure 4.10. 122

4.4 List of possible events within the model. Each event has a rate at which it occurs, and a resultant action if it occurs. $E \to E + 1$ denotes the extracellular bacteria population increasing by 1. $M(a_g)$ represents a macrophage with Erlang division parameter a_g. $M_I(I)$ represents an infected macrophage with I intracellular bacteria. . . . 125

5.1 Summary of data relating to the aggregated bacteria growth. In total 46 populations of bacteria were independently tracked and the total area of fluorescence was measured. 140
5.2 (On previous page) Complete and final list of model parameters along with descriptions, their fitted values, and 95% confidence intervals. All fitted values are fitted using maximum likelihood methods, and 95% confidence intervals are estimated using the likelihood surface generated. 156
5.3 Test statistics and the corresponding critical values at the 5% significance level for the Kolmogorov-Smirnov (KS) test and the Anderson-Darling (AD) test for the range of distributions presented in this section. 165

A.1 Each colour in equations A.1 to A.6 refers to a separate mechanism being modelled in chapter 3. 192

Acknowledgments

Many thanks go to my supervisors, Matt Keeling and Ian Hall, for their continued guidance over the last four years. I am very grateful for all the support they have provided. I would also like to thank Chiara Toniolo for her useful insights into the biological mechanisms and for sharing her experimental data.

To the staff and students of MathSys I would like to say thank you so much for providing such an incredible environment. My time here has been an unforgettable experience.

Finally, I would like thank all of my amazing friends and family, for their continued encouragement and support. In particular to my housemates, Chris, Sami and Sophie. And, of course, to Victoria, without whom this would not have been possible.

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

The work presented was carried out by the author except in the cases outlined below:

- All data was collected and provided by Chiara Toniolo at the École polytechnique fédérale de Lausanne.

No parts of this thesis have previously been published.

Abstract

Tuberculosis (TB) kills over one million people annually, ranking it the leading cause of death from a single infectious agent. Generally a pulmonary infection, the primary route of transmission for the disease is respiratory. After being deposited in the lungs, *Mycobacterium tuberculosis* (Mtb) are met by the host's first line of defence: the alveolar macrophage. The early interaction between the macrophages and the bacteria set the stage for how the disease develops in the lungs. Understanding these dynamics is a crucial step in understanding the variety of host responses observed, ranging from early clearance, to latent infection, to active disease. This thesis explores the development of mathematical models of these early interactions, and identifies key mechanisms contributing to the proliferation of the bacteria.

A literature search was used to identify the current biological and mathematical understanding of the processes and interactions involved in Mtb infections. This was followed by the development of a range of iterative mathematical models. The models were parameterised using maximum likelihood methods on a rich data set, which tracks, at the individual cell level, the interactions between macrophages and bacteria. A mechanistic approach was used during the model development process, with a focus on balancing biological realism with mathematical complexity. Each iteration of the model developed and built upon the previous, which resulted in a robust model that was able to capture the dynamics of macrophages and bacteria over the 200 hours tracked in the experiments.

Using sensitivity analysis and a range of computational experiments, the mechanisms involved in the system were analysed to identify factors contributing to the proliferation of bacteria. Naturally, the bacteria growth rate was the most dominant. While investigating the methods of control available to the macrophages, it was shown that a small reduction to the intracellular growth rate for all macrophages results in a more beneficial outcome than complete inhibition of growth in only some macrophages. It was further demonstrated that an increase to the rate of phagocytosis has a beneficial effect, but only up to a point, after which it becomes detrimental.

Chapter 1

Introduction

Tuberculosis (TB) is a globally endemic disease, the causative agent of which is *Mycobacterium tuberculosis* (Mtb). Approximately 1.7 billion people are infected with Mtb and are thus at risk of becoming one of 10 million people that develop active TB annually [1]. TB is one of the top 10 causes of death worldwide as of 2016 [2]. The only infectious diseases that rank higher fall under the umbrella terms ‘lower respiratory infections’ and ‘diarrhoeal diseases’ [2]. It has been the leading cause of death from a single infectious agent since 2007 [2]. While there is no broadly efficacious vaccine, almost 100 years after its initial development, the Bacillus Calmette–Guérin (BCG) vaccine still provides an approximate 50% reduction of risk of vaccinated adults developing active TB [3]. Treatment of the disease requires a regimen of antibiotic and chemotherapeutic drugs over a period of at least 6 months due to the ability of a population of Mtb to exhibit varying phenotypic resistance through adapting their metabolic functions [4].

The classical view of TB infection has been a dichotomy of infectious states: latent or active [5]. Those with latent TB infection (LTBI) are neither symptomatic nor infectious and may go their entire lives without knowing they are infected. Individuals with LTBI have a roughly 1 in 10 chance of developing active disease later in life, but this process can take decades [6]. In contrast to this view, TB can be thought of as a spectrum of disease outcomes ranging from natural clearance of the bacteria by the innate immune response to active TB disease [7] with LTBI being a dynamic process.

The predominant route of infection for humans is the inhalation of droplets containing minute numbers of bacilli which have been exhaled by infected and infectious individuals [8]. Upon entering the lungs, the bacilli are engulfed by innate alveolar macrophages whose role it is to eliminate the invading organism and begin the signalling process for the adaptive immune response [9]. The Mtb bacteria however, have a repertoire of defence mechanisms to enable their survival [10].

1.1 Global epidemiology

The Global TB report 2019 [1] recorded the estimated incidence of TB per 100,000 population by country in 2018 and found almost every country in the world to be affected. In fact, only 7 countries had an estimated incidence of 0, the sum of whose populations totalled less than 250,000. At the other end of the spectrum, the 5 countries with the highest burden account for 55% of the global burden of TB, while accounting for 45% of the global population, as shown in figure 1.1.

Figure 1.2 (top) shows the estimated burden of TB per 100,000 population. It is clear that the distribution of cases is not evenly spread globally and that developing countries are more strongly affected, in particular South-East Asia and sub-Saharan Africa. Of note is the evident relationship between TB and the Human Immunodeficiency Virus (HIV), shown by the strong correlation of incidence for the two diseases (see figure 1.2 bottom). HIV is a serious risk factor for developing TB, with an approximate 20 fold increase over those who are HIV negative [11]. Additional risk factors include air pollution, smoking, malnutrition and immunosuppressive drugs [11].

In low incidence countries, import of infection is a more significant factor than community transmission, for example rates of TB in the UK are approximately 14 times higher for people born outside of the UK [1] compared to those born within the UK. Conversely in high incidence countries, the role of super spreaders remains a critical issue for the design of intervention strategies [13]. Handel et al. [14] assume duration of cough as a proxy for infectious duration and demonstrate that 20% of

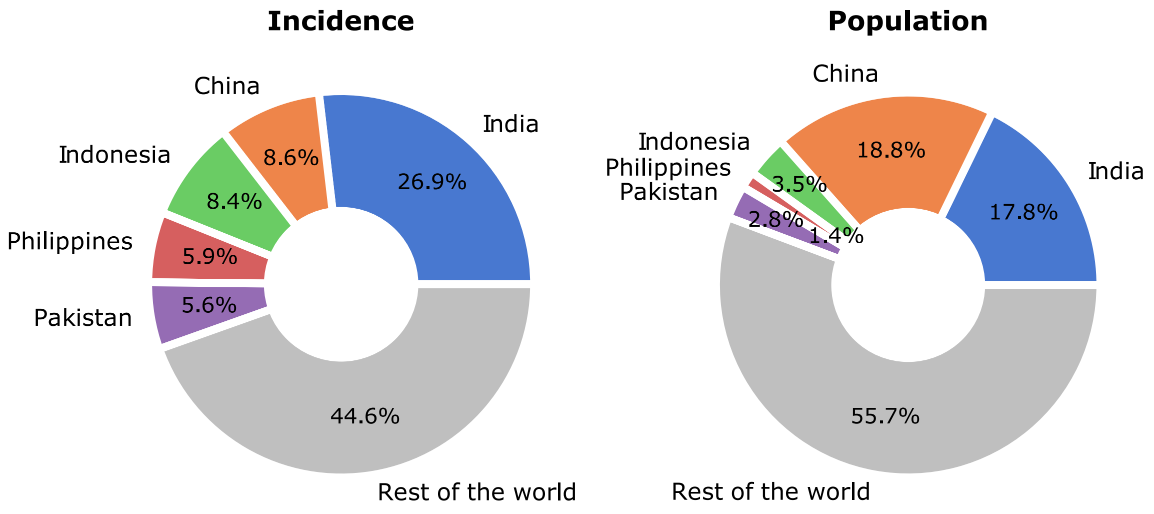

Figure 1.1: The distribution of the estimated global incidence of TB (**left**) compared to the distribution of the global population (**right**) as reported in the Global Tuberculosis Report 2019 [1]. The 5 countries with the highest absolute incidence account for roughly 55% of the the global incidence, while only accounting for 45% of the human population. The incidence is estimated per 100,000 population, predominantly from case notification data combined with either expert opinion or standard adjustment. With the exception of China, all of these countries are significantly over-represented for burden of TB demonstrating the imbalance in how the disease is managed globally.

patients over 6 cohorts from China, Peru, The Gambia, and Uganda contributed to 50% of the total cough days.

]

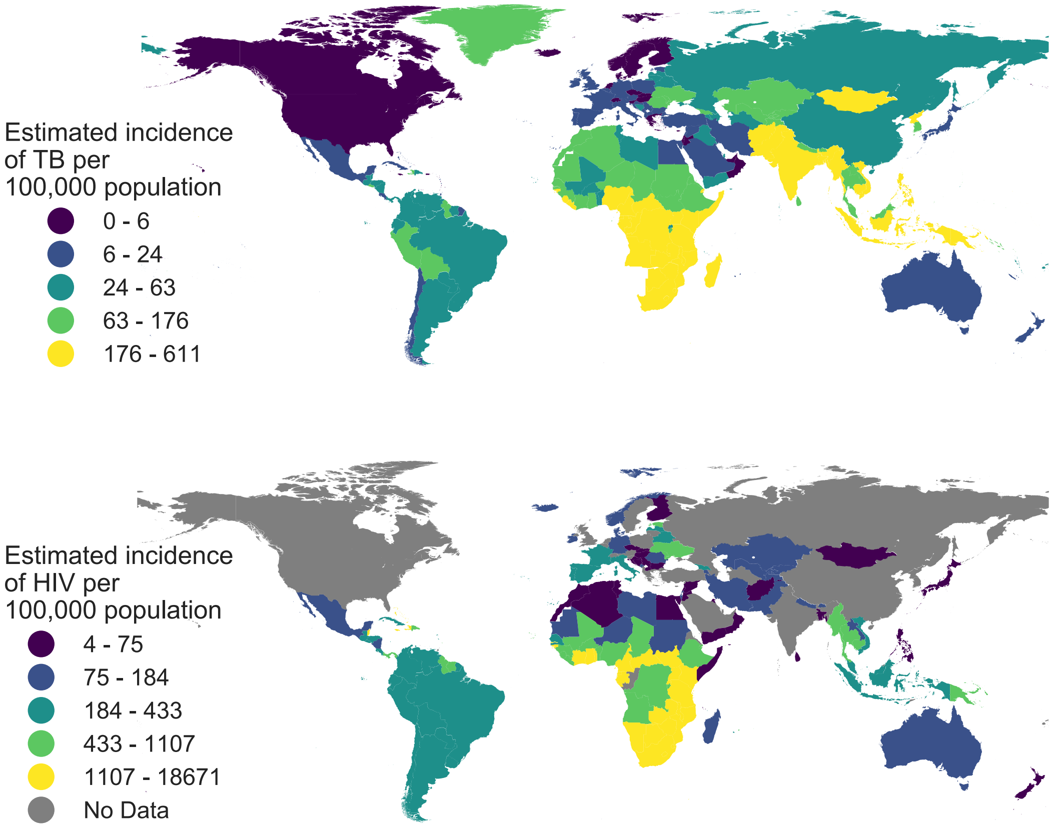
Estimated incidence of TB per 100,000 population
0 - 6
6 - 24
24 - 63
63 - 176
176 - 611
Estimated incidence of HIV per 100,000 population
4 - 75
75 - 184
184 - 433
433 - 1107
1107 - 18671
No Data

Figure 1.2: The estimated incidence of TB is not uniformly distributed globally. South East Asia and sub Saharan Africa demonstrate the highest levels of active disease (**top**) [1]. As coinfection with HIV is a significant risk factor, the global incidence of HIV per 100,000 population is also shown (**bottom**) [12]. This risk is particularly evident in southern Africa where HIV prevalence in new and relapsed TB cases can exceed 50% [1]. Figures produced using data from references [1, 12].

1.2 Understanding the within-host dynamics

Developing the understanding of the within-host dynamics of a disease is a crucial step in developing the understanding of the disease as a whole [15]. Two natural goals for within-host modelling of disease are firstly, to understand the mechanisms involved in sustaining the infection and how to control them, and secondly to predict the bacterial load and therefore infer the transmission potential of the host. By developing mechanistic models of the in vitro interactions of immune cells and bacteria, this thesis will explore the former.

The within-host dynamics of TB infection roughly follows three stages, shown in figure 1.3. The first stage is the initial infection, during which time the immune response is non-specific. The infected individual has inhaled air-borne droplets containing minute numbers of bacilli, and the bacilli have been deposited in the lungs [8, 16]. Despite providing an effective initial barrier against most bacterial infections, the innate alveolar macrophages are the primary target of Mtb, which is able to reside within them and proliferate [9, 17, 18].

The second stage is the formation of the granuloma. At this point, the adaptive immune response has been initiated and the hallmark granuloma forms to contain the infection. If the individual is immunocompromised, the infection will typically result in active disease characterised by uncontrolled growth of bacteria within the lungs and potentially other organs. For immunocompetent hosts, the granuloma will contain the infection and prevent further growth of the bacteria, but it is often not able to fully clear the infection. This stage is commonly referred to as latent TB infection (LTBI) and the likelihood of developing active disease is roughly 5 to 10% with half of the risk being concentrated in the first two years of infection [5].

The third and final stage is reactivation of the infection. The host immune response is no longer able to contain the infection and secondary active disease is established. There are a number of risk factors related to reactivation, highest among them being coinfection with HIV and administration of immunosuppressive drugs such as for organ-transplantation recipients [22].

Since the classical dichotomy of active vs latent TB is being reconsidered in favour of a spectrum of disease [7], the three stages described above can be thought of as having blurred boundaries with LTBI being a dynamic process of constant endogenous reinfection and control [23]. As a dynamic process, LTBI can be thought of as a quasi-steady state, opening an interesting direction of research in model construction.

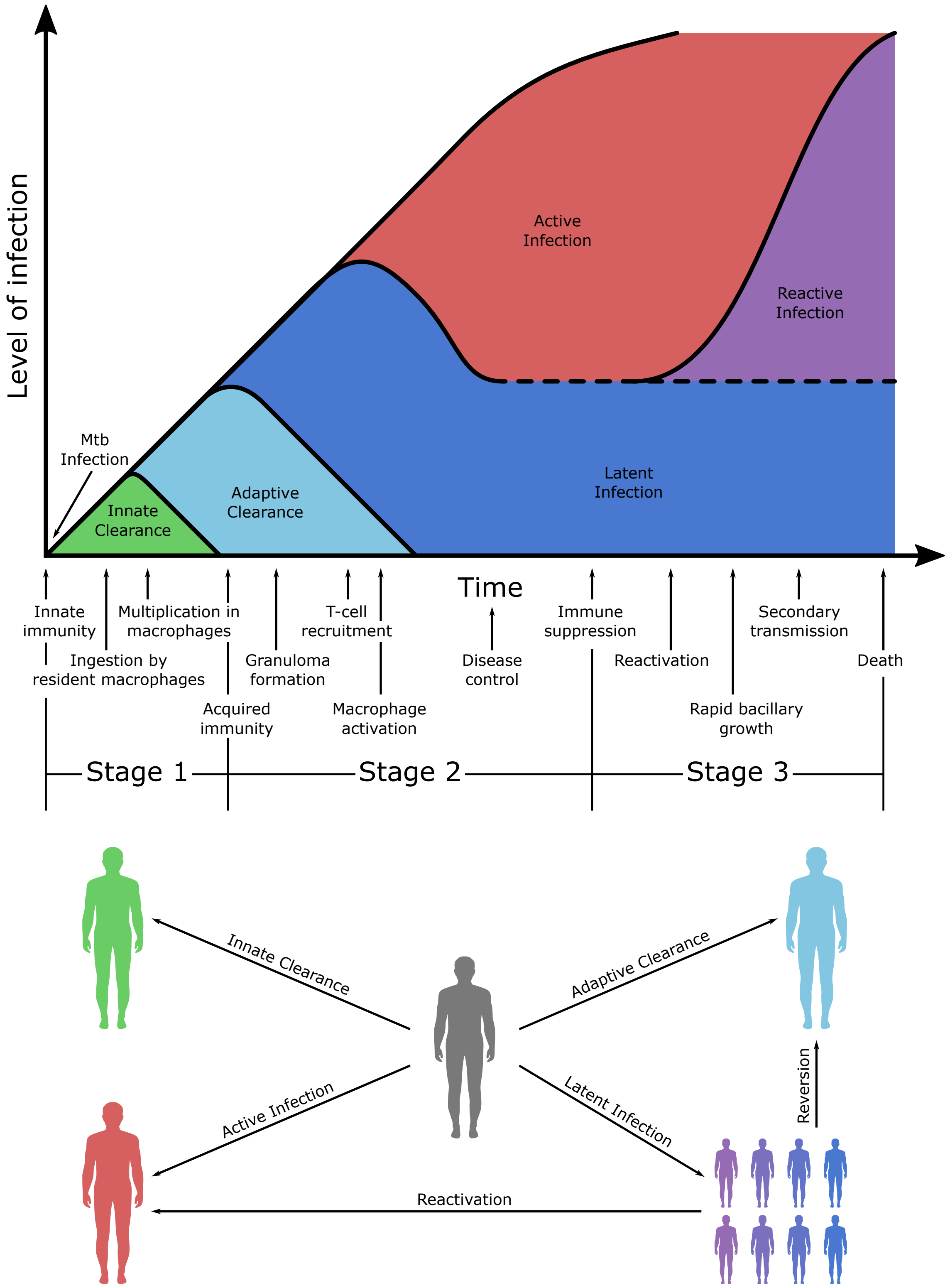
Level of infection
Active Infection
Reactive Infection
Mtb Infection
Adaptive Clearance
Latent Infection
Innate Clearance
Time
Innate immunity
Ingestion by resident macrophages
Multiplication in macrophages
Acquired immunity
Granuloma formation
T-cell recruitment
Macrophage activation
Disease control
Immune suppression
Reactivation
Rapid bacillary growth
Secondary transmission
Death
Stage 1
Stage 2
Stage 3
Innate Clearance
Adaptive Clearance
Active Infection
Latent Infection
Reversion
Reactivation

Figure 1.3: (On previous page) TB exhibits a wide variety of responses in the host, as well as heterogeneity in its route from exposure to active infection. The exposure process can be broadly categorised into three stages: natural clearance, adaptive immune control, and active disease (**top**). An interferon-γ release assay (IGRA) test is able to differentiate between innate clearers and adaptive clearers [19] however an exposed individual that does not manage to clear the infection can progress to active disease either directly (primary TB infection) or after containing the infection for a period as LTBI (secondary TB infection) (**bottom**). Figure adapted from references [5, 20, 21].

1.3 Importance of early dynamics

It is often stated that the early dynamics of TB infection is instrumental to the outcome of the disease [24–29], whether the host is able to clear the infection, or becomes either latently or actively infected. In a study of TB case contacts by Verrall et al. [19], early clearers were defined as testing persistently negative over the three month follow up, compared to converters whose test converted from negative to positive. The test used is the interferon-γ release assay (IGRA) which is a blood test to measure the immune response to the TB bacteria. Comparing the immune response in the two groups to alternative bacterial stimuli showed that early clearers had in an increased production of pro-inflammatory cytokine production. This demonstrates a critical difference in outcome as a result of heterogeneous host response.

Alveolar macrophages play an important role in the establishment of infection and the initial growth of bacteria. Leemans et al. [30] show that, since the bacteria depend on infiltrating the host cells, it is possible for a depletion of macrophages to carry some protective effects for the host. However, the effects of activated macrophages are still shown to be beneficial [31] demonstrating the dual harmful and protective role of macrophages.

1.4 Disease progression

The progression of infection varies dramatically from person to person. For the majority of people, infection will result in LTBI which can remain undetected for years [32]. For the roughly 10% of people that go on to develop active disease, TB initially presents most commonly as a persistent cough [21]. If left untreated, the infection continues to develop which can lead to coughing blood and extensive scarring in the lungs [33]. As the bacteria continue to degrade the lung tissue, the space fills with liquid, resulting in a wet, hacking cough [34]. The afflicted is now highly contagious and also severely weakened, opening the doors for additional pathogenic infections. As more and more liquid fills the lungs, oxygen levels drop and they will eventually succumb to respiratory failure [35].

1.5 Description of experiment

The work presented in this thesis has been developed in collaboration with experiment. Initially a small trial data set was collected to which an initial mathematical model was fitted. Through an iterative process of model construction, additional complexity was built into the models. In turn, additional data was required to parametrise these new mechanics, thus informing the experimentalists which data were required to be collected.

All experiments were performed by Chiara Toniolo at the École polytechnique fédérale de Lausanne. The primary focus of the experiments was to investigate the heterogeneity of the response exhibited by a homogeneous population of macrophages to the Mtb bacteria. The models presented in the subsequent chapters of this work are a secondary outcome of the experiments, although additional data was collected as a result of this work. The experimental data allows detailed models of the early interaction between these two cellular populations to be fitted using data entirely collected from a single set of experiments, reducing the reliance on parameter estimation and qualitative evaluation. The available data was used to inform how the models developed, fitting mechanisms as they were observed in the data, and introducing additional mechanisms as that data became available.

The macrophages used were bone marrow derived macrophages (BMDMs) from C57BL/6 mice. While BMDMs are distinct from alveolar macrophages, they infiltrate most tissue types and contribute to the maintenance of the macrophage population there [36, 37]. BMDMs are also observed to settle and persist in the lungs during disease [38]. There are important differences between mouse and human cells [39]. In the experiments mouse cells were used as they are significantly easier to work with. This means that while the qualitative findings are important, the precise quantitative findings may not directly translate into within-human models.

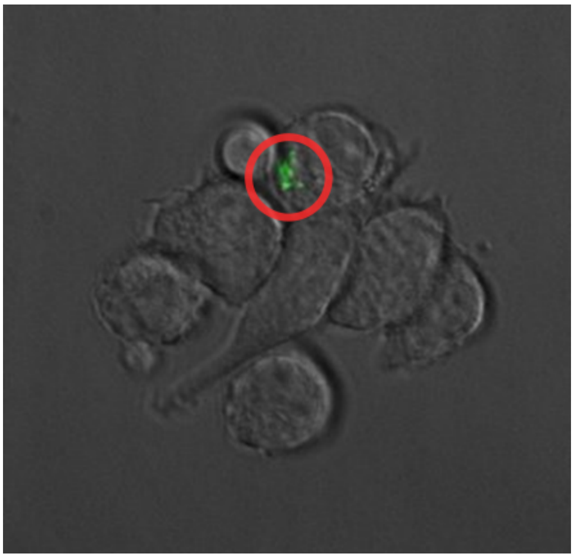

Figure 1.4: A snapshot of some macrophages, one of which contains intracellular bacteria (circled in red). The level of bacteria is measured as area of fluorescence since individual bacteria cannot be visually resolved.

The bacteria used were a green fluorescence protein (GFP) expressing Mtb Erdman strain. This strain is a highly virulent strain of Mtb and is widely used as a standard virulent laboratory strain. The GFP protein allows for easier discrimination between bacteria and macrophage and has no other effect on the host [40].

Macrophages and bacteria were cultured independently and allowed to develop. A total of 10^4 healthy macrophages were left to adhere to a micro-dish for 24 hours. Bacteria were filtered to eliminate aggregates and ensure single cell suspensions. These were then used to infect the macrophages with a multiplicity of infection (MOI) 1:1 and left over a period of four hours to become infected. Infected macrophages were washed to eliminate extracellular bacteria prior to time lapse imaging.

Samples were photographically imaged through a microscope and were taken every two hours. These images were then filtered with a green fluorescence protein high quality (GFP HQ) dichroic filter which selectively passes light in a specific frequency range, allowing for easy identification of bacteria. Figure 1.4 shows the resultant image to be processed into tabular data. The images were then manually segmented and analysed using ImageJ software by Chiara Toniolo. Manual thresholding of

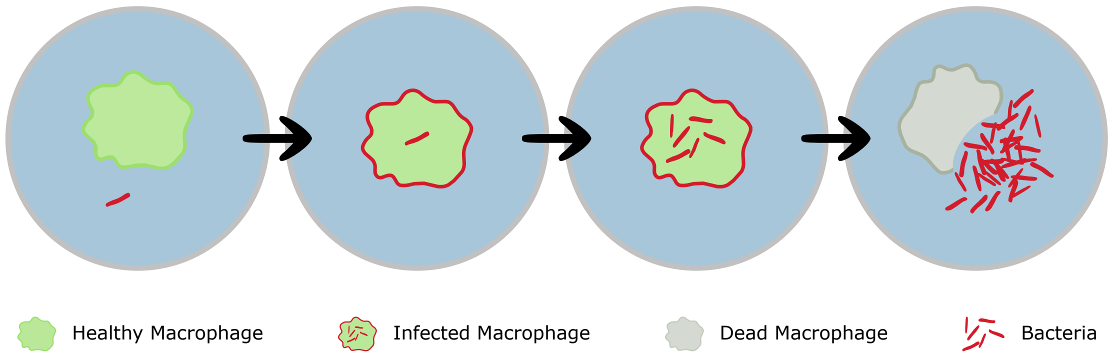

Figure 1.5: Macrophages and bacteria are introduced on a Petri dish and the interactions between them are tracked. Of particular interest is the infection of macrophages, intracellular growth of bacteria, and eventual release of bacteria back into the extracellular space.

fluorescence intensity was applied to discriminate pixels corresponding to the bacteria from the background. The area of fluorescence was measured as a proxy for the bacterial load. Figure 1.5 shows a schematic representing the stages through the experiment.

The experiments yielded a rich data set tracking individual macrophages from healthy, to infected, to division, to death. Table 1.1 shows an example of a few rows of the data. Movie numbers refer to a set of time lapse images localised to the same area of the dish with each movie being taken independently of the others. Cells are named numerically, with daughter cells appending either a 0 or 1 to the name of the parent cell. Thus when cell ‘1’ divides, two new cells are formed: ‘10’ and ‘11’. Cells that were not infected were not tracked in the initial experiments, however in subsequent experiments they were included in the data. The data was tracked hourly, thus for each cell the time of birth, death or division, and time of infection is recorded. This time is measured relative to the experiment, not the the cells, thus a time of birth of 40 represents a cell that was born 40 hours after the start of the experiment. The fate of the macrophage defines whether the cell divided (D), or was killed (K). Lastly, the ‘infected from’ column denotes the time at which a macrophage became infected.

	Movie	Cell	Born	Dead / Div	Infected from	Fate
0	02	1	0	28	0	D
1	02	10	29	68	29	D
2	02	100	69	79	69	K
3	10	21	57	78	67	D
4	10	210	79	92	79	D
5	10	211	79	123	79	D

Table 1.1: An individual experiment is recorded as a series of snapshots of the Petri dish stitched together to produce a movie of the interactions between the macrophages and the bacteria. The bacteria are genetically tagged with a green phosphorescent gene, and the images are passed through a green light filter in order to enhance the experimenters ability to identify bacteria. Over the course of the experiment individual macrophages are tracked and the time of important events is noted, as well as whether the macrophage died or divided. When a macrophage divides, its daughter cells are tracked under new names, which are defined as the name of the parent cell with a '0' or a '1' concatenated to the end.

If the cell remained healthy throughout the experiment this value will remain empty. If the macrophage was infected from birth, this value will equal the value for 'born'.

In addition to tracking macrophages at the individual level, there are time series data of the intracellular bacterial fluorescence within a macrophage, measured as area of fluorescence (see table 1.2). When an infected macrophage divides, it is possible to explore the division of the intracellular load by considering the initial loads of the two daughter cells. One consideration for this analysis is that the sum of the fluorescence of the daughter cells does not necessarily equal the load of the parent cell immediately before division due to fluctuations in fluorescence captured and the bacteria population shapes pre and post division.

Some experiments demonstrated extracellular bacteria in isolation of macrophages,

	Movie	Cell	Time	fluorescence
0	25	17	24	1.479
1	25	17	32	1.479
2	25	17	40	1.584
3	25	17	48	4.542
4	25	17	56	5.809
5	25	17	64	7.077
6	25	17	72	10.563

Table 1.2: Example data showing the time series of the intracellular load for a single cell in a single experiment.

measuring the population growth free from macrophage interactions. Since bacteria are tracked at the macro population level, individual births and deaths of bacteria are not specifically recorded. The data do however note the starting number of bacteria and corresponding fluorescence making it easier to estimate the numbers of bacteria from the total fluorescence. Figure 1.5 shows a schematic representing the stages through the experiment, and figure 1.4 shows a snapshot taken directly from the experiment. The maximum duration of the movies is roughly 140 hours for earlier experiments and extended to 200 hours for subsequent experiments. The limit on time owes to the increase in population of bacteria making it difficult to continue processing images beyond this point.

1.6 Methodology

As mentioned in the previous section, the development of the models presented in this thesis grew predominantly based on the observable mechanisms in the data. This section will outline the various methodologies used to design, parametrise, and evaluate each model.

1.6.1 Mechanistic modelling

Mechanistic models make the assumption that a complex system can be understood by examining the workings of individual parts, and how they are coupled. A mechanistic model can lie anywhere on the "continuum between *mechanism sketch* and *ideally complete description of the mechanism*" [41]. Since mechanistic plausibility must also be taken into consideration, it is not sufficient to simply reproduce the data. Mechanistic model construction should have a basis in the real-world understanding of how a system evolves and there should be a clear relationship between observations, assumptions and equations. Throughout this thesis a series of models of increasing complexity are presented. The models were developed organically through discussion with biological experimentalists and evaluating how well they are able to capture the dynamics of an experimentally observed biological system.

1.6.2 Ordinary differential equations

Ordinary differential equations (ODEs) form the foundation for a wide variety of mathematical models and are especially well used in population dynamics [42]. A wealth of research in numerical methods has resulted in very efficient and accurate algorithms for solving ODEs [43]. Combined with only requiring a single integration due to their determinism, this makes exploring large parameter spaces very computationally simple. Further, ODEs are often mathematically tractable and thus some analytical analysis can be performed prior to computational.

There are some drawbacks to the use of ODEs however. When modelling small populations, ODEs may suffer what is known as the "atto-fox" problem, whereby the model predicts that a population reduces then recovers from numbers that, in reality, would render the species extinct as a result of chance fluctuations in the discrete number of individuals [44]. Due to this, ODEs are often more suited to large populations, where stochastic effects are dampened due to the law of large numbers.

Since ODEs very quickly become non-analytically solvable, evaluating systems of ODEs will be performed through numerical integration. Numerically evaluating a system of ODEs involves discretising the time element of the model then iterating through time to estimate the state of the system. An initial value problem is defined by some function $f : [t_0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d$ and an initial condition $y_0 \in \mathbb{R}^d$ such that

$$y'(t) = f(t, y(t)), \quad y(t_0) = y_0. \tag{1.1}$$

The simplest method of numerically solving this system is the Euler method. The derivative y' can be approximated by

$$y'(t) \approx \frac{y(t+h) - y(t)}{h}. \tag{1.2}$$

Rearranging this gives

$$y(t+h) \approx y(t) + hy'(t) = y(t) + hf(t, y(t)). \tag{1.3}$$

Thus the sequence $t_n = t_0 + nh$ can be used to create a sequence y_n which numerically approximates $y(t_n)$.

The Euler method is an example of a explicit method, since the numerical sequence can be evaluated immediately. An example of an implicit method is the backwards Euler method, where the approximation of y' is defined as

$$y'(t) \approx \frac{y(t) - t(t-h)}{h}. \tag{1.4}$$

Which results in

$$y_{n+1} = y_n + hf(t_{n+1}, y_{n+1}). \tag{1.5}$$

In order to define the sequence (y_n), one must solve an equation for y_{n+1}. The implicit methods may be more computationally expensive than the explicit method, however it can also be more stable when considering stiff equations. Intuitively, stability is the rate at which the error $|y_n - y(t_n)|$ grows, thus the increase in stability allows for a larger time step h to be used while maintaining the same level of accuracy. A larger value of h results in fewer calculations and thus higher computational efficiency. Adaptive time stepping allows a numerical integrator to estimate the expected error from a given time step and then adjust the size of the time step to maximise computational efficiency within the constraint of a given tolerance on the error.

The implementation of these solvers that will be used in this thesis is provided in SciPy package [45] and uses the VODE integrator [46] which itself implements the Adams methods. This method uses polynomials to approximate y_{n+1}. For a detailed construction of the multistep method see [47] .

1.6.3 Agent based modelling

When modelling small populations, it is often more suitable to use a stochastic model, so that the effects of variance can be measured [48]. Agent-based models (ABMs) provide a flexible framework within which it is possible to simulate highly complex and potentially heterogeneous interactions between elements of the system. Since the models are inherently stateful, it is trivial to apply ad hoc rules to the different agents and track how this affects the evolution of the system. It is for these reasons that an ABM was developed when transitioning to a stochastic framework. An important feature of ABMs is the ability to investigate low-probability events which might have a large effect to the system as a whole.

The most common method of evaluating these models is an implementation of the Gillespie algorithm [49] which makes use of the Poisson distribution to determine which events occur when. Suppose there are two possible events in a system, event A and event B, each modelled by a Poisson process with rates α and β respectively which depend only on the current state of the system. In order to sample which event occurs next, one could sample two times $t_A \sim \text{Expon}\left({}^1\!/\!_\alpha\right)$ and $t_B \sim \text{Expon}\left({}^1\!/\!_\beta\right)$. If $t_A < t_B$ then event A is enacted, the system is updated and the process repeats. However as the number of events increases, this process becomes very computationally inefficient. Instead, two uniform random numbers are drawn $r_1, r_2 \sim \text{Uniform}\left[0, 1\right]$. The first is used to select the event based on the rates α, β, thus if $r_1 < {}^\alpha\!/\!_{\alpha+\beta}$ then event A is enacted. The second random number r_2 is used to define the time step. Since each event is occurring at rate α and β respectively, the rate of any event occurring is $\alpha + \beta$. Thus inverting the cumulative distribution function of the exponential distribution gives the time step as $-\log(r_2)/(\alpha+\beta)$. The system can now be updated and the process repeated until an end condition is met. With this construction, no matter how many events are being modelled individually, each update to the system will always only require the generation of two random numbers.

The simple model described above is an example of a continuous time Markov chain (CTMC). Consider the event A to be toggling a switch S_A between states $\{0, 1\}$, and event B toggles the switch S_B between states $\{0, 1\}$. This describes a four state process on $S = \{0, 1\}^2$ with the transition matrix

$$Q = \begin{pmatrix} -\alpha-\beta & \beta & \alpha & 0 \\ \beta & -\alpha-\beta & 0 & \alpha \\ \alpha & 0 & -\alpha-\beta & \beta \\ 0 & \alpha & \beta & -\alpha-\beta \end{pmatrix}. \tag{1.6}$$

A CTMC [50] is defined by a finite or countable state space S and a tranisition matrix Q with dimension equal to S. The elements of $Q = (q_{i,j})$ are such that for $i \neq j$, $q_{i,j} \geq 0$ and $q_{i,i} = -\sum_{i \neq j} q_{i,j}$. $q_{i,j}$ defines the rate at which the state changes from state i to state j.

1.6.4 Parameter inference

Parameter inference is the process of determining the most likely set of parameters (θ) for a given model and a given set of data (D). Three common methods of parameter inference are maximum likelihood estimate (MLE), maximum a posteriori (MAP), and Bayesian inference. MLE maximises the likelihood function $\mathcal{L}(D|\theta)$ and results in a point estimate for the parameters. It is possible to then use the likelihood function to approximate the uncertainty of this estimate.

MAP is a Bayesian approach that still results in a point estimate rather than a full distribution of θ and is sometimes called the ‘Poor Man’s Bayes’. In order to determine the posterior distribution of θ, a prior distribution $P(\theta = x)$ is updated with the likelihood function $P(D|\theta = x)$ according to $P(\theta = x|D) = \frac{P(D|\theta=x)P(\theta=x)}{P(D)}$.

Here $P(D)$ is simply a normalising constant, and since it is independent of θ can be ignored for the purpose of finding the mode of the posterior distribution. The MAP is then the set of parameters θ that maximise the posterior distribution. When nothing is known about the distribution of θ, or when avoiding adding bias to the model, a uniform distribution on θ can be used. This is called a flat prior. Using a flat prior for MAP will reduce the posterior to the likelihood function and thus MLE is a special case of MAP.

Bayesian inference results in the full posterior distribution $P(\theta = x|D)$ as defined for MAP but requires the calculation of $P(D)$ such that $P(\theta = x|D)$ integrates to 1. Again, with a flat prior on the distribution of θ, the posterior distribution reduces to the normalised likelihood function.

The literature search presented in section 1.7 has identified a range of analogous parameters along with a range of estimates. These could be used to define prior distributions for the models developed throughout this thesis, however doing so may add bias into the parameter inference. Since differing experimental design naturally results in differing dynamics, for this study no prior distribution will be defined and thus parameters will be inferred using MLE.

1.6.5 Sobol sensitivity analysis

Through the process of perturbing the values of input parameters to a model and observing the effect on the output, it is possible to numerically quantify the uncertainty of the output of a model with respect to its input parameters. As a variance based method, Sobol sensitivity analysis decomposes the variance of the output to be divided into fractions which can be attributed to individual parameters or sets of parameters [51]. Sobol sensitivity analysis is suitable for both continuous and stochastic models where non-monotonous relationships are expected between the input parameters and the output measure [52], making it a good choice for this work. The fundamental assumption of the Sobol method is that a model f with n input parameters $\mathbf{X}$ can itself be decomposed in the following way,

$$f(\mathbf{X}) = f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{i<j}^{n} f_{i,j}(x_i, x_j) + \cdots + f_{1,2,\ldots,n}(x_1, x_2, \ldots, x_n). \tag{1.7}$$

When these functional decompositions are orthogonal, and f is suitably integrable, then since orthogonal functions are uncorrelated, the variance of the model f can be written as

$$\begin{aligned} D = \mathrm{Var}(f) = &\int f^2(\mathbf{X})\,\mathrm{d}\mathbf{X} - f_0^2 \\ = &\sum_{s=1}^{n} \sum_{i_1<\cdots<i_s}^{n} \int f_{i_1\ldots i_s}^2 \,\mathrm{d}x_{i_1} \ldots \mathrm{d}x_{i_s} = \sum_{s=1}^{n} \sum_{i_1<\cdots<i_s}^{n} D_{i_1\ldots i_s}. \end{aligned} \tag{1.8}$$

Finally the sensitivity index for each parameter and each set of parameters can be defined

$$S_{i_1,\ldots,i_s} = \frac{D_{i_1,\ldots,i_s}}{D}. \tag{1.9}$$

In practice these indices are calculated numerically, predominantly being limited to the first order indices (S_i) and the second order indices ($S_{i,j}$). Saltelli [53, 54] presents a robust and efficient algorithm for estimating first, second and total order sensitivity index. In particular, an efficient method of sampling high-dimensional parameter spaces is defined: the Saltelli sampler. This quasi-random low-discrepancy sequence results in a smaller number of parameter sets needing to be evaluated while still allowing the sensitivity indices to converge. A heuristic definition of a low-discrepency sequence (x_n) is that it should fill the parameter space, minimising holes, and that projections to lower dimensions should also leave few holes. Formally, discrepency of a set $P = \{x_1, \ldots, x_N\}$ with respect to the interval $[a, b]$ is defined as

$$\sup_{a \le c \le d \le b} \left| \frac{A(P; [c, d])}{N} - \frac{d-c}{b-a} \right| \tag{1.10}$$

where $A(P; [c, d])$ is the number of points in P that are contained in $[c, d]$. A low-discrepency sequence is therefore a sequence for which the discrepency is low. A sample of uniformly random distributed points does not qualify, as when projecting to lower dimensions, many points will overlap.

Figure 1.6 shows the process of creating this sequence of parameters for a model with two parameters (without loss of generality they can be assumed to be contained with the unit hypercube). Each additional frame shows how the next iteration of the sequence introduces more parameters. The final two frames show the comparison of the Saltelli sampler with the Latin-Hypercube sampler.

To generate N samples with the Latin-Hypercube sampler, the parameter space is divided into $M = 2^m < N$ even hypercubes, and each sample is uniformly chosen from within these smaller spaces. For example, given a parameter space of the unit square $[0, 1]^2$, to generate 4 samples, the parameter space is first divided into the four quadrants $[0, 0.5] \times [0, 0.5]$, $[0, 0.5] \times [0.5, 1]$, $[0.5, 1] \times [0, 0.5]$, and $[0.5, 1] \times [0.5, 1]$. Each sample is then drawn from each subset of the parameters space independently. This is an improvement over purely sampling the parameter space uniformly as it ensures equal dispersion throughout the parameter space without resorting to the regularity of a grid.

The classic example function for demonstrating global sensitivity analyses is the Ishigami function,

$$f(\mathbf{X}) = \sin(x_1) + a\sin^2(x_2) + bx_3^4\sin(x_1), \tag{1.11}$$

where $a = 7$ and $b = 0.1$. This is because it is a smooth function which exhibits strong non-linearity and non-monotonicity. The input parameters are each bound by the intervals $[-\pi, \pi]$. Figure 1.7 shows the sampled parameter space of the three parameters in the top three frames, and how they each relate to the model output in the bottom three frames.

Using Python package SALib [55] to perform the numerical analysis of the Ishigami function results in the sensitivity indices shown in figure 1.8. While parameter x_1 has the largest effect on the model output, some of this effect is dependent on the value of x_3, whose only effect on the output is via x_1. Parameter x_2 is completely independent of the other two, but has a slightly contribution on the variation of the output than that of x_1.

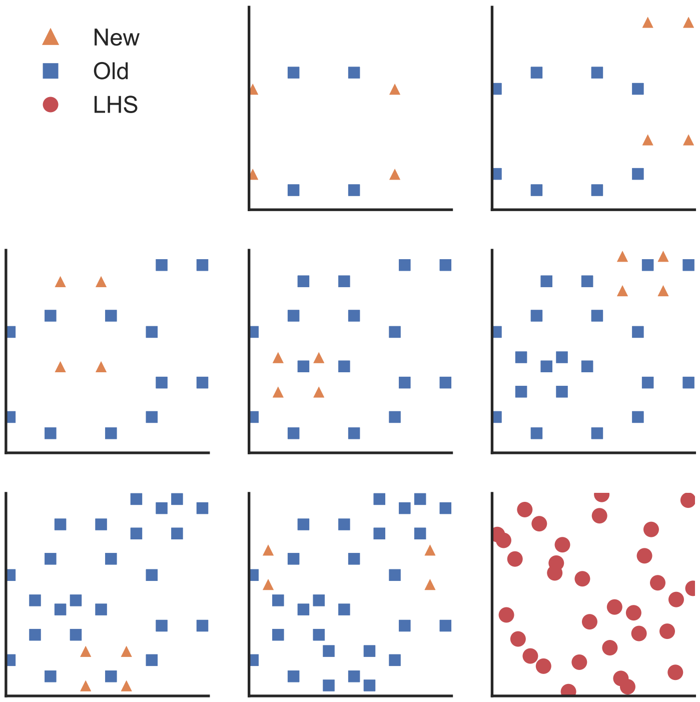

Figure 1.6: Progressive iterations of the Saltelli sampler fill in the parameter space with the goal of covering as much of the space as possible with as few parameter sets as possible, allowing the calculation of the global sensitivity indices to converge quickly. Shown is the coverage of the parameter space for a two parameter model. The last two frames (bottom middle and bottom right) show how the Saltelli sampler compares with the Latin Hypercube sampler.

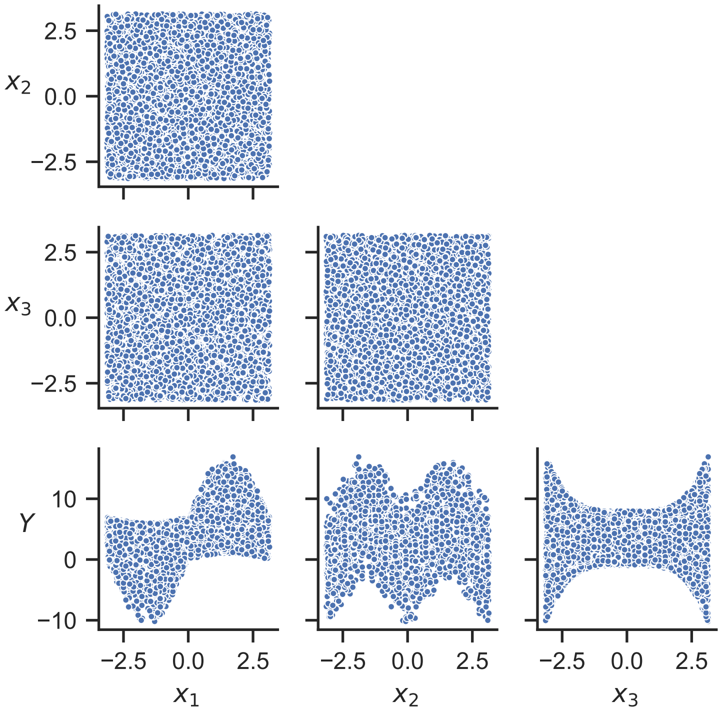

Figure 1.7: Sampling the hypercube $[-\pi, \pi]^3$ using the Saltelli sampler and evaluating the Ishigami function over the resultant samples. The displayed output is passed to the Sobol sensitivity analysis along with the output parameters to determine the sensitivity of the model to each of the parameters. (**Top**) The Saltelli sampler evenly samples all three dimensions with good coverage even when projecting on to lower dimensions. (**Bottom**) The Ishigami function displays strong non-linearity and non-monotonicity with respect to the input parameters.

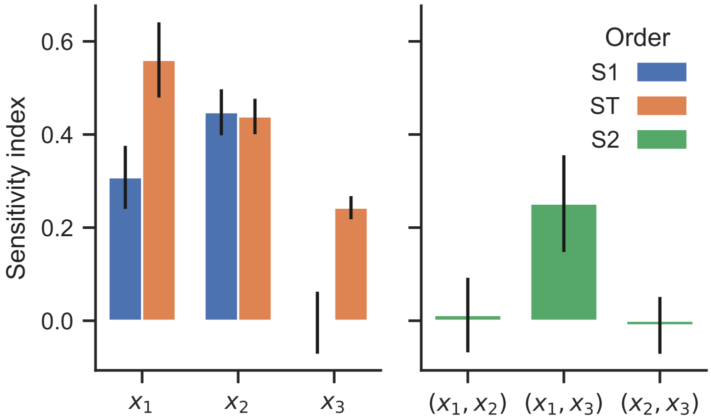

Figure 1.8: Sobol sensitivity indices calculated for the Ishigami function. Parameter x_1 has the most effect on the model output. Parameter x_2 is naturally independent of x_1 and x_3. Parameter x_3 only affects the model output in conjunction with x_1.

1.6.6 Goodness of fit

Goodness of fit is a method of quantitatively analysing how well a statistical model fits a set of observations. Two commonly used methods for evaluating goodness of fit are the Kolmogorov-Smirnov (KS) test [56] and the K-sample Anderson-Darling (AD) test [57]. Both methods are used to determine whether two sample sets are drawn from the same distribution, the null hypothesis being that they are drawn from the same distribution. Both tests are non-parametric and use the empirical distribution defined by the given set of observations to calculate their respective test statistics.

For a given set of n independent and identically distributed observations $(X_1, \ldots, X_n)$ the empirical distribution function is defined as

$$F(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{X_i < x} \tag{1.12}$$

where $\mathbf{1}_{X_A}$ is the indicator of event A.

Given two sets of observations $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_m)$ with two corresponding empirical distributions $F(x)$ and $G(y)$, the KS test quantifies the distance between these two empirical distributions by finding the maximum distance between them. Thus the KS test statistic is defined as

$$D_{\mathrm{KS}} = \sup_z \left| F(z) - G(z) \right| . \tag{1.13}$$

For a given significance level α, the null hypothesis can be rejected at level α when

$$D_{KS} > \sqrt{- \ln\left(\frac{\alpha}{2}\right) \frac{n+m}{2nm}} . \tag{1.14}$$

The KS test provides an intuitive measure for determining how similar two sets of observations are, however since it only looks for the largest difference between the two distributions, it can often fail to capture the tail of the distribution.

The AD test extends the KS test by calculating not just the largest difference but the sum of the squares of the differences. Then by weighting those differences by weights inversely proportional to the uncertainty of the empirical distribution, the tail of the distribution can be accounted for. Extending the above definitions to include $H(z)$, the empirical distribution of $(X_1, \dots, X_n, Y_1, \dots, Y_m)$, the test statistic for the AD test is given by

$$D_{AD} = \frac{n+m}{nm} \int_{-\infty}^{\infty} \frac{(F(z) - G(z))^2}{H(z)(1 - H(z))} \, \mathrm{d}H(z). \tag{1.15}$$

For a given significance level α, the null hypothesis can be rejected at level α when

$$D_{AD} > t(\alpha) = b_0 + b_1 + b_2. \tag{1.16}$$

where b_0, b_1 and b_2 can be found in table 2 of [57].

The implementations for both hypothesis tests are available in Python's SciPy package [45]. These will be used in chapter 5 to investigate how well the model captures the data.

1.6.7 Identifiability

When fitting mathematical models to experimental data, the goodness of fits tests outlined above are not sufficient to prove the model was chosen correctly. For example if the experimental data has only a small number of observations, it is possible that a large change in parameter values do not significantly influence the goodness of fit. Analysing the identifiability of the model can help to address this issue. A model that results in the same outcome for different sets of parameters is said to be structurally unidentifiable [58]. That is, for a given model $\mathcal{M} = \{M(\theta) : \theta \in \Theta\}$, $\mathcal{M}$ is identifiable if and only if $M(\theta_1) = M(\theta_2) \implies \theta_1 = \theta_2$. Since a model $\mathcal{M}$ can often be highly complex and mathematically intractable, proving this statement is often not feasible. Instead, identifiability can be demonstrated through computation [59].

As discussed in section 1.6.4, the method of parameter inference that will be used is maximum likelihood. Thus, the definition of identifiability can be translated to the framework of maximum likelihood. Given a model $\mathcal{M} = \{M(\theta) : \theta \in \Theta\}$ with likelihood function $\mathcal{L}(\theta|X)$, then $\mathcal{M}$ is said to be non-identifiable if $\mathcal{L}$ does not have a unique maximum over the parameter space Θ [60].

1.7 Literature review

There are two aspects to developing a model of complex biological systems: the current understanding of the biological mechanisms involved, developed through experimentation in vitro and in vivo; and the current methodologies and practices involved in modelling such a system. In this section, the literature is reviewed and summarised to understand firstly, the biology and pathogenesis of tuberculosis infections, and secondly the range of modelling methodologies being used.

Initial literature sources were identified using a systematic approach by searching Google Scholar and PubMed using the search term ((‘within-host’ OR ‘macrophage’) AND ‘tuberculosis’ AND (‘mathematical’ OR ‘model’)). Results were divided into biological and mathematical groups. The biological results were filtered to results containing descriptions, experiments, or data of mechanisms involved in Mtb interactions. Mathematical results were grouped by whether they modelled the innate or adaptive immune response, and whether or not the model is multiscale. A number of reviews were identified during this search which were used to identify further sources.

1.7.1 Biology and pathogenesis of tuberculosis infections

Mycobacterium tuberculosis (Mtb) infections are highly complex and heterogeneous resulting in a large range of observable mechanisms employed by both host and pathogen. The scope of this thesis is the initial interactions of Mtb with innate alveolar macrophages. Since these interactions are being measured in vitro, it is natural to prioritise those studies focussing on the times prior to the influence of the adaptive immune response. As such, while there is a wide range of literature detailing the more complicated interactions of the adaptive immune response and granuloma formation, this review will focus on the subset pertaining to the early dynamics, where early dynamics is considered to be the time prior to the adaptive immune response.

Ref	**Growth rate** (hour^{-1})
James, Williams, and Marsh [65]	0.0289 – 0.0472
Cox [66]	0.0289
Lipsitch and Levin [67] (data from [68])	0.0083
North and Izzo [69]	0.0111
Sershen, Plimpton, and May [70] (data from [71])	0.00095 – 0.06
Lin et al. [72]	0.0132 – 0.0182
Ref	**Death rate** (hour^{-1})
Lipsitch and Levin [67] (data from [68])	0.0063
Sershen, Plimpton, and May [70] (data from [71])	0.0001 – 0.0009

Table 1.3: Estimated growth and death rates of Mtb measured experimentally from the biological literature.

Life cycle of *Mycobacterium tuberculosis*

Experiments investigating the life cycle of Mtb often result in varied results. This is due to varying strains of the bacteria as well as gene expression variability under different conditions. The two most studied strains are the H37Rv and the Mtb Erdman strains [61, 62]. H37Rv was first isolated in the early 1900's and has since been maintained in a range of different laboratories around the world [63]. It has been shown that these strains derived from H37Rv have evolved differently and it is noted that the "notion of H37Rv as a standard reference strain should be used with some caution" [64]. In comparison to H37Rv, the Mtb Erdman strain used in the experiments described in section 1.5 is used as a standard virulent strain as it has a slightly higher doubling time than H37Rv [62]. Through various studies involving experimentation, the literature demonstrates a wide range of growth rates of Mtb as shown in table 1.3.

At the higher end of the spectrum for extracellular growth rates, the H37Rv strain of Mtb used by James, Williams, and Marsh [65] retained their infectious capabilities, demonstrating a correlation between growth rate and virulence. Cox [66] agrees with the faster replication rate, demonstrated using *Mycobacterium bovis*, and even hypothesises that Mtb could divide much faster depending on which genes are being expressed by the bacterium. At the lower end of the growth rates, Lipsitch and Levin [67] and Sershen, Plimpton, and May [70] (using data from [68] and [71] respectively) both provide similar estimates of bacteria growth rates which are several times lower than the other sources. The impact of this range will be investigated later in this work (see section 2.4.2).

North and Izzo [69] demonstrate that despite its ability to disseminate around the body, Mtb only grows progressively in the lungs, suggesting the lungs are the crucial site of resistance against infection. Since this study aims to model the interactions between bacteria and alveolar macrophages, of particular interest is the specific measurement of extracellular bacteria growth within pulmonary lesions, performed by Lin et al. [72]. While the growth rate fell substantially over the experiment, the rate during the first few days, during which initial interactions would occur, was between 0.0132 and 0.0182.

Phenotypic heterogeneity in a genetically homogeneous population describes the variation that exists between individual cells [73] and can be beneficial for a favourable outcome for the bacteria, especially over long term and persistent infections such as TB [74]. It is well known that populations of Mtb are able to exhibit phenotypic heterogeneity under stress conditions [75], in particular they are able to adapt their metabolic functions to slow growth [76]. A significant result of this heterogeneity is phenotypic resistance to drugs, which rely on metabolic functions, in an otherwise sensitive population [77, 78]. Separate is the growth of genotypic resistance, in the form of multi-drug resistant TB (MDR-TB) and extensively drug resitant TB (XDR-TB) [79, 80].

Investigating the relationship between virulence and several other characteristics of various strains of Mtb, Ordway et al. [81] identified a lack of correlation between virulence and growth or drug-resistance rates in vitro. Since it is generally accepted that more virulent strains have higher growth rates in vivo, this could imply that growth rates in vitro do not correlate with growth rates in vivo.

Mtb is well regarded as a highly infectious pathogen, for which the minimum infectious dose is very low. The Center for Disease Control and Prevention (CDC) state a 50% infectious dose of less than 10 [82, 83], implying the minimum infectious dose to be lower. To investigate the minimum infectious dose of TB, Walsh et al. [84] inoculated two cynomolgus macaques, shown to closely resemble human infection [85], with doses of 10 colony forming units (CFUs)[1]. They found that, despite not exhibiting symptoms of disease, both monkeys exhibited immunological and culture evidence of Mtb suggesting dose may correlate with disease progression and outcome. In a similar study, by infecting calves with a range of doses, Dean et al. [87] demonstrated that 1 CFU is sufficient to result in pulmonary pathology typical of bovine tuberculosis. And in a murine model of Mtb, Plumlee et al. [88] demonstrated that ultra low doses of 1 to 3 CFUs recapitulated key features of human TB.

Life cycle of alveolar macrophages

The alveolar macrophage is the first line of defence against foreign particles and pathogens. It is a long lived cell that resides within the alveolar sacs of the lung. The population of macrophages in the lungs is maintained by two mechanisms: recruitment of new cells from circulation, and production of new cells by cellular division within the lungs [89–92]. Although Van Furth, Diesselhoff-den Dulk, and Mattie [93] show that the predominant method is through recruitment rather than local proliferation.

Macrophage recruitment is predominantly governed by two mechanisms [94]. Firstly, pathogen-associated molecular patterns excite a receptor found on macrophages known as a toll-like receptor, attracting the macrophages along the concentration gradient. Secondly, macrophages themselves release chemokines to which additional macrophages are attracted. Cambier et al. [94] demonstrate the manipulation of these recruitment pathways by the Mtb bacteria to result in preferential conditions for the bacteria.

---

[1] A CFU is the unit of viable bacteria in microbiology, where it is assumed that each colony of bacteria is separate and founded by a single cell [86, p. 18].

The average life span of the alveolar macrophage is said to be 81 days [95], however more recent studies have shown it to be much longer. Murphy et al. [96] found negligible replacement of alveolar macrophages after eight months in healthy mice. Joshi, Walter, and Misharin [97] remark that “While it is certain that alveolar macrophages are long-lived, current fate mapping studies do not cover the entire lifespan of the animal.”

In a model of the interactions between pathogenic yeast and macrophages it is suggested that resident macrophages are induced into mitosis in reaction to the presence of a pathogen [98, 99]. Dühring et al. [100] use this to categorically estimate the division rate of mammalian alveolar macrophages to be 0.0176 per hour. Mtb infection models often ignore explicit macrophage division in favour of recruitment, as recruitment is the more dominant form of macrophage population maintenance. Antony et al. [101] demonstrate the host’s quick response to a drop in macrophage population with increased recruitment of more cells to the site of infection. This work will investigate the role macrophage division has in limiting the population of bacteria.

Process of phagocytosis

Phagocytosis is the process of internalising foreign particles and pathogens by host immune cells with the goal of eliminating them and thus avoiding any harm they could cause to the host. The resident alveolar macrophage population present in the lungs at the time of infection is orders of magnitude larger than that of the invading pathogen. Macrophages are highly specialised phagocytic immune cells [102] thus phagocytosis is almost an inevitable occurrence. Some pathogens, including Mtb, have adapted to survive and even depend on the process [103].

The mechanisms of phagocytosis are extensive and complex, to the extent that it is possible that the exact route of phagocytic entry into the macrophage may have a major impact on the survival chance of Mtb [104]. For a detailed review of the mechanisms see Aderem and Underhill [105]. For the scope of this work it is sufficient to understand that a healthy macrophage that is unaffected by external interferences will immediately internalise a bacteria that it comes into contact with.

Interferon-γ (IFNγ) is a critical messaging protein for the innate and adaptive immune response, released by macrophages during phagocytosis. Its exact function during early Mtb infections is clouded by heterogeneity. Macrophages' core function is to internalise foreign particulates, however without the presence of IFNγ, tumor necrosis factor (TNF), and other cytokines, macrophages lack the ability to fully kill off the Mtb bacteria [106–109]. In the presence of IFNγ, macrophages may transition from their current 'resting' state into an 'activated' state, enhancing their ability to kill pathogens. When studying Mtb specifically however, studies of the effect of IFNγ range from its being inhibitory [110] to enhancing [111, 112] of resultant mycobacterial growth.

Macrophages are freely able to ingest bacteria they come in to contact with, although there is some evidence showing that only a limited number of macrophages have the capacity to phagocytose viable mycobacteria [113]. Innate resting macrophages pre-activation have very low to no chance to kill bacteria [20]. Infected macrophages may undergo phenotypic changes diminishing its phagocytic capabilities [114] and there is evidence of phenotypic heterogeneity within the healthy macrophage population regarding their phagocytic abilities and their reactions to cytokine cues [115].

Macrophages have a unique roll in the infectious process, serving as the primary effector cell against Mtb, but also providing a safe environment within which the bacilli can replicate [116]. If a macrophage does not eliminate the bacilli, there is a chance it can inhibit the intracellular growth through exposure to IFNγ [117]. While not eliminating the bacteria this can serve to control the infection for a long time, a precursor to LTBI.

Ref	**Growth rate** (hour^{-1})
Paul, Laochumroonvorapong, and Kaplan [118]	0.0037 – 0.0144
Silver, Li, and Ellner [119]	0.0191
Zhang et al. [120]	0.0181 – 0.0211

Table 1.4: Estimated growth rates of intracellular Mtb measured experimentally from the biological literature.

Intracellular processeses

One of the core effector functions for eliminating the Mtb bacteria is autophagy [121]. This is the process of phagolysosome fusion [122]. By targeting regulatory lipids and enzymes of the host macrophage [123], Mtb is able to interrupt this process and consequently survive and grow within the macrophage [10, 18, 124].

Macrophages infected with Mtb may also attempt to undergo apoptosis (cell death) to eliminate the possibility of providing a safe environment for the bacilli to grow [29, 125]. Apoptotic macrophages are then engulfed by uninfected macrophages to eliminate the bacteria in a process called efferocytosis [121]. Evidence that this apoptosis is beneficial to the host is difficult to obtain, but it is supported by the observation that many bacteria have evolved mechanisms to facilitate survival within macrophages [126]. Intracellular Mtb modify the life cycle of macrophages, either limiting apoptosis [127] or in favourable conditions promoting it [128, 129]. Virulence of Mtb strain correlates with the ability of the pathogen to suppress apoptosis of infected cells [130]. When Mtb inhibits apoptosis of macrophages, the result is necrosis, a preferable environment for the bacilli [131], while successful apoptosis of the macrophage results in reduced viability of the surviving bacilli [132].

Once Mtb has gained entry to the macrophage it must establish residence therein. Mtb has evolved a host of specific mechanisms for acquiring the material required for growth and out-competing the host cell from sequestering resources to interfering with cell mechanisms [133]. See [121] for an in depth review of the mechanisms Mtb has at its disposal to interfere with effector functions of macrophages.

Similar to experiments involving extracellular bacteria, results for intracellular bacteria growth vary. Paul, Laochumroonvorapong, and Kaplan [118] measured the intracellular growth rate between 0.0037 and 0.0144 per hour. An alternative experiment involving the same H37Rv strain of bacteria measured a growth rate of 0.0191 per hour [119]. Zhang et al. [120] found the intracellular growth rate of seven isolates to be bimodal between 0.0181 and 0.0211 per hour with a mean of 0.0203. This shows heterogeneity of interactions between macrophages and Mtb as well as the potential sensitivity of the system to environmental factors.

Paul, Laochumroonvorapong, and Kaplan [118] also measure the number of intracellular bacteria over several days of infection of in vitro macrophages. The distribution of intracellular bacteria after six days drops sharply beyond 128 cells, suggesting a maximum intracellular load. Similar estimations from Hirsch et al. [134] yields a maximum load of about 50 cells after seven days of infection, and from Repasy et al. [135] yield 40 cells after eight weeks. These studies demonstrate the need for modelling work to consider a maximum intracellular load.

Later processes and granuloma formation

A side effect of using macrophages as a safe haven for bacterial growth is that infections have a tendency to remain local to the original deposition and it is here that the granuloma will form [136]. It has been demonstrated in a macaques model that the majority of granulomas begin with a single bacterium [72].

The rate at which the adaptive immune response is initiated is critical to the ultimate success of the response [137]. T Cells are a type of white blood cell and form an important part of the adaptive immune response and require antigen presentation from macrophages to become activated [138, p. 1309]. This process occurs first within the draining lymph nodes once dendritic cells have transported bacilli there

[139] and has been shown to take up to two weeks in mice [140]. In macaques it was shown that granuloma formation begins at the earliest two weeks post infection [85].

The hallmark of TB infection is the formation of the granuloma in the lungs. The granuloma is a complex structure facilitated by the host's adaptive immune response which contains the initial infection, however the fight between host immunity and bacterial survival is controlled by the latter allowing the bacilli to persist, contributing to early Mtb proliferation and dissemination [141–143].

1.7.2 Mathematical modelling of tuberculosis infections

There is a continuum of time scales and complexity for modelling the immune response to tuberculosis. Broadly, modelling can be divided into two categories, pre-adaptive immune response and post-adaptive immune response. The latter can be further divided by whether or not the model is multiscale. Since the scope of this thesis is the innate immune response as considered in vitro, without the influence of an adaptive immune response, the literature of mathematical models will be divided to mirror this focus: firstly, models that only consider the innate response, secondly, models that consider the adaptive response.

Innate immune response models

The importance of the initial innate immune response to tuberculosis infections is somewhat overlooked in the modelling literature, however heterogeneity in the host immune response at this early stage can have a large impact on the resultant outcome of the infection [19]. An important aspect of the innate immune response is the rate at which foreign particulates and organisms are phagocytosed by the alveolar macrophages. Tran, Jones, and Donaldson [144] present a non-TB specific mathematical model of phagocytosis of particles post inhalations. A key element is the use of mass action for the rate at which monocytes interact with foreign particles. This is a simple construction for both ODE models and stochastic models. Some experimentation has shown phagocytic rates to follow Michaelis-Menten dynamics

[145], demonstrating a saturation in the rate of phagocytosis dependent on the level of extracellular bacteria.

Models of this stage of infection often take the form of either ordinary or partial differential equations, however Carvalho et al. [146] take a systems biology approach and implement a petri net (PN) model. This results in a high level of abstraction and allows the authors to focus on the mechanisms involved. It is important to note that one level of abstraction is the removal of time dependence in the model.

Marino et al. [147] parametrise an ODE model to a set of novel mice experiments in order to investigate the impact of the age of the host on the infection outcome. Fitting the full model to a single set of experiments is unusual in the realm of within-host TB modelling as the majority of models use parameters taken from a range of sources and use sensitivity analysis to understand the effects. For example Gammack, Doering, and Kirschner [148] extend other models in the literature [144, 149] and make use of parameters therein.

Adaptive immune response models

A significant proportion of the modelling literature regrading the adaptive immune response to TB has been produced by the Kirschner Lab [150], implementing a range of model structures and increasing complexity. Since the biology of the adaptive immune response is so complicated, the detail of the modelling quickly rises, making it infeasible to investigate everything with a single model. As such, each model seeks to understand the impact of a specific mechanism or type of cell.

The earlier modelling efforts constitute ODE formulations involving varying numbers of cellular populations. Initially Wigginton and Kirschner [149] present a system of nonlinear ODEs modelling two bacterial populations, six cellular populations and four cytokines. This model is then extended to a two compartment model, incorporating the need to transport bacilli to the lymph nodes before a TB specific immune response can be mounted [151–154]. A parallel extension of the Wigginton model extends the populations of macrophages and T-cells [155] and later the effect of TNF [156] culminating to a model of 32 equations and 210 parameters [147]. These variations and extensions are used to investigate a range of hypotheses, such as the cost of achieving latency and the specific role of the lymph nodes.

Two common limitations of ODE models are capturing heterogeneity and modelling small populations. To evaluate these while investigating the interplay between mechanisms, Pienaar and Lerm [157] develop a Markov-chain model.

Separate to the Kirschner group, Hao, Schlesinger, and Friedman [158] develop a partial differential equation (PDE) model similar to the multiscale models discussed in the next section from Gammack, Doering, and Kirschner [148] and Gammack et al. [159]. Integrating the PDEs over time and space, Hao, Schlesinger, and Friedman [158] demonstrate how the radius of the granuloma develops over time.

Multiscale models

Although this thesis is not directly concerned with the adaptive immune response present in all the multiscale models, the literature provides detailed mechanistic constructions and parameter estimation useful for the model construction process discussed later. Multiscale models of TB infection almost unanimously take the same form: an ABM at the tissue scale with cytokines and sometimes bacteria being modelled continuously with either ODEs or PDEs. The ABM may be in a continuous space, or a lattice, and may also be either two or three dimensional. Since these models require significant computational power, there is also literature specific to the optimisation of model fitting and validation [160]. Once again the majority of the models are from the Kirschner Lab and present either a precursor to, or an implementation of GranSim, the modelling software used to implement multiscale models.

The modelling framework begins with Segovia-Juarez, Ganguli, and Kirschner [136] who present a novel ABM on a lattice representing a small section of lung space.

The model consists of macrophage and T-cell agents interacting with bacteria and continuous chemokines. Many of the parameters within the model are estimated from biological literature and the rest are explored using latin hypercube sampling (LHS) and partial rank correlation coefficient (PRCC). This is a theme which continues throughout the development of later models. Gammack et al. [159] demonstrate the need for a range of approaches from continuos deterministic models to discrete stochastic ones to fully understand the underlying biology. As GranSim increases in complexity, the individual pieces of literature begin exploring specific mechanisms within the infection process, such as the effect of TNF [161, 162] and interleukin-10 (IL-10) [163] or how the ratio of macrophage states can be predictive of the outcome of the granuloma [164].

In parallel to the GranSim modelling, Warrender, Forrest, and Koster [26] developed and used an early version of the simulator CyCells [165]. The model presented is also an ABM, however it is in a continuous three dimensional space. The authors conclude that the processes and heterogeneity involved in the formation of granulomas are crucial to the eventual outcome of the infection. It is therefore important to properly characterise the early dynamics such as these.

Explicitly including the effects of oxygen levels in the within host environment, Sershen, Plimpton, and May [70] demonstrated high correlation between these levels and the level of bacteria at the end of the simulation. Bowness et al. [166] combine these effects with the differentiated bacterial populations presented by Pienaar and Lerm [157] to demonstrate that when bacteria are deposited further away from blood vessels in the lungs, less favourable outcomes for the host are more likely.

In a unique modelling framework, Pitcher et al. [167] develop a network based meta-population model of the whole lung and associated lymph system incorporating oxygen availability as well as adaptive immune responses mechanisms. All parameters of the model are taken from the Kirschner group. The model demonstrates the large effect heterogeneity can have, even within simple models and with small heterogeneity.

This will also be explored in this study through sensitivity analyses and computational experimentation.

1.8 Thesis Summary

In this thesis, a series of iterative models of the in vitro interactions between macrophages and *Mycobacterium tuberculosis* (Mtb) has been developed based on current understanding of the biological literature and observable mechanisms in a rich experimental data set. The models were then analysed to investigate the dominant mechanisms affecting how the system evolves.

The model structure presented in this body of work has been built mechanistically from first principles based on the underlying biology and has been developed independently within this thesis. A novel experimental dataset which tracked the interactions between macrophages and Mtb was used to inform on and parametrise the included mechanisms. During the model development, limitations of the available data were identified and fed back to the experimentalist. This lead to additional data being captured in subsequent experiments enriching the data and allowing for more detailed modelling to be undertaken.

The key results of this study are firstly that a simple, novel, and robust model, which is able to capture the dynamics of a complex biological system, has been developed. Through mathematical, computational, and sensitivity analyses, the key mechanisms driving the system trajectory have been investigated. It has been shown that the dominant factor in the proliferation of bacteria is the bacteria growth rate, both extracellularly and intracellularly. While investigating the possible mechanisms of macrophage control, two methods were contrasted. First, each macrophage had a small probability to completely inhibit the intracellular growth of bacteria. Second, all macrophages reduced the intracellular growth rate, but did not inhibit it completely. It was demonstrated that the latter had a larger impact on reducing the resultant number of extracellular bacteria compared to the former. It was also shown that while increasing the rate of phagocytosis by a small amount had a protective effect in delaying the bacteria growth, increasing it further reversed this effect.

Chapters 2 to 5 present the development of each iteration of the model, the data they are parametrised with, and their analyses. A discussion of the results and the limitations of this study is presented in chapter 6.

Chapter 2

Model 1: Differential Equations With Implicit Intracellular Bacteria

In this chapter, a simple ordinary differential equation (ODE) model will be fitted to a small 'proof of concept' data set. The model will consider the interactions of extracellular bacteria and macrophages, with intracellular bacteria being modelled implicitly through the role of infected macrophages. Parameters will be fitted using maximum likelihood methods, the likelihood function of which will be used to define confidence intervals on those parameters. These ranges will then be used to explore the sensitivity of the model to each parameter. Finally, by investigating the method of implicitly modelling intracellular bacteria, the model will show that explicitly modelling intracellular bacteria is necessary to fully capture the system dynamics.

2.1 Model description

In this section the various components and mechanisms that make up this preliminary model will be described. The purpose of this model is to capture the basic life cycles of the two types of cells as observed in the in vitro experiments. The two cells to be modelled are the innate alveolar macrophages and *Mycobacterium tuberculosis* (Mtb), as well as a simple version of their interactions; bacteria uptake, macrophage death, and bacteria release.

2.1.1 Mechanisms

There are four fundamental parts to the model presented in this chapter, each of which has their own complex biology. Throughout this thesis this biology will be considerably simplified in order to give a high level overview of the cellular population dynamics. Macrophages and bacteria will follow simple birth death processes and interaction will be limited to the taking up and release of bacteria by macrophages.

Macrophages

The macrophages used in the experiment described in section 1.5 are murine cells, thus it is difficult to draw direct conclusions in how the system would react within a human host. Aston et al. [168] demonstrate a “clear difference between murine and human responses.” Direct comparison of murine and human macrophage models is difficult because of a range of factors such differences in experimental design, varying strains of bacteria, and alternative multiplicities of infection. The focus of this study is to model and understand the system that was observed in 1.5, and thus conclusions will not be directly drawn about the effect on human cells.

In order to replicate the experimental results, macrophage growth is incorporated into the model. Growth is observed in the data at the individual level, measuring the time of birth and the time of division of individual cells. Macrophage death is also directly observed and tracked, so macrophage death can be explicitly modelled, independent of growth. Deaths are observed analogously with individual birth and death times recorded.

Bacteria

The observations of bacteria are at the population level, rather than the single cell level, so it is not possible to directly observe division and death events of the bacteria. Since each experiment demonstrates exponential growth of the bacteria population, a simple exponential model can be fitted to the observations. The growth of bacteria is measured as the total area within an image demonstrating the green fluorescence of the green fluorescence protein (GFP), which itself grows exponentially. Since the parameter of interest is the rate of growth, this area can be used as a proxy value for population count assuming that the number of bacteria is proportional to the area of green fluorescence. There is a possibility that bacteria become stacked on top of each other which would result in an underestimation of the growth rate, however since the time scales and growth rates are small, the variance as a result of this is likely to be smaller than the variance between experiments and so is ignored.

Phagocytosis

The role of the macrophage is to ingest foreign particulates and pathogens and eliminate them from the system. Mtb are well adapted to survive this process and remain alive intracellularly. The data measure the time at which healthy macrophages become infected, however since the bacteria are not tracked at the single cell level, there are no records of macrophages killing the bacteria. The time to infection data are used to fit a force of infection model assuming homogeneous mixing of the macrophages and the bacteria. If β is the rate of infection, then the force of infection is proportional to β and the also the number of infective agents, i.e. the number of bacteria [169].

In the biological literature macrophages are often classified into three groups, namely, resting, activated, and suppressed [37]. For the purposes of this study the state of the macrophages will largely be ignored. Macrophage activation often occurs through interactions with chemical signals from additional immune cells, such as T and B lymphocytes [110]. Since these cells are not present in the experimentation, the state of the macrophage will be assumed to be resting. It is, however, noted that this status can be thought of as having blurred boundaries [170]. Through computational experiments, the role of macrophage suppression will be explored in later chapters by considering the impact of reducing the macrophages ability to internalise bacteria.

Infected macrophage division

When an infected macrophage divides, there are two possibilities for the infection status of the daughter cells, either they are both infected, or one is infected and the other healthy. This process is observed in the data through either 1 or 2 infected daughter cells being recorded post division.

2.1.2 Model definition

As a preliminary model, the construction presented in this chapter will be a simple ODE model wherein the interactions of entities will follow the law of mass action: the

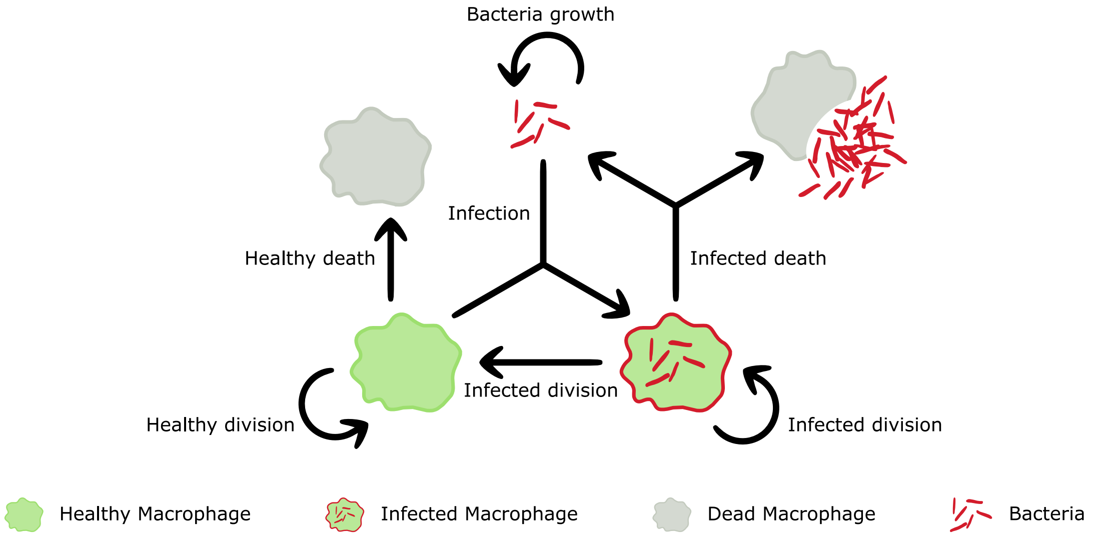
Bacteria growth
Infection
Healthy death
Infected death
Infected division
Healthy division
Infected division
Healthy Macrophage
Infected Macrophage
Dead Macrophage
Bacteria

Figure 2.1: Schematic of the model showing the mechanisms included and the possible interactions between cell types. Healthy macrophages get infected by bacteria and become infected macrophages. Macrophages follow a birth (division) death cycle, while bacteria grow exponentially. Infected macrophages divide to produce both infected and healthy macrophages. And finally, when infected macrophages die, they release their intracellular load back into the extracellular space.

rate of interaction is directly proportional to the product of the size of the populations. The structure of the resultant model is a non-linear system of three ODEs, which tracks the populations of extracellular bacteria (E), healthy macrophages (M_H) and infected macrophages (M_I). Since the data capturing bacteria growth is 1 dimensional, bacteria death is incorporated into the parameter α_E and not modelled explicitly. Note that the only non-linear term in the equations describes the infection of a macrophage by a bacteria. The goal of this preliminary model was to be as simple as possible, and so most terms were kept linear. The interactions of the macrophages and bacteria are reminiscent of predator-prey dynamics [157] and so an adaptation of the Lotka–Volterra equations [171] was used. A schematic of the model is shown in figure 2.1. The equations are defined as follows:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \alpha_E E - \beta E M_H + d_I N M_I \tag{2.1}$$

$$\frac{\mathrm{d}M_H}{\mathrm{d}t} = g_H M_H + g_I(1-p)M_I - \beta E M_H - d_H M_H \tag{2.2}$$

$$\frac{\mathrm{d}M_I}{\mathrm{d}t} = g_I p M_I + \beta E M_H - d_I M_I \tag{2.3}$$

Extracellular bacteria grow exponentially and are internalised by healthy macrophages at a per capita rate proportional to the populations of bacteria and healthy macrophages, resulting in an infected macrophage. The units of β in equation 2.1 are $(\text{bacteria} \times \text{day})^{-1}$ however in equations 2.2 and 2.3 the units of β are $(\text{macrophage} \times \text{day})^{-1}$. It is assumed that a single bacteria interacts with a single healthy macrophage to form a single infected macrophage and thus the magnitude of β is the same in all 3 equations.

Both healthy and infected macrophages also replicate exponentially, however when an infected macrophages replicates it results in both healthy and infected macrophages according to some proportion, seen in the $g_I p$ and $g_I(1-p)$ terms in equations 2.2 and 2.3. This is a result of the intracellular load either being divided between both daughter cells, or all of the load being passed to a single daughter cell.

Healthy and infected macrophages both have an exponential death rate. The death

Symbol	**Description**	**Unit**
E	Extracellular bacteria	cells
M_H	Healthy macrophages	cells
M_I	Infected macrophages	cells
α_E	Extracellular growth rate	hour^{-1}
N	Average number of intracellular bacteria	cells
g_H	Healthy macrophage growth rate	hour^{-1}
d_H	Healthy macrophage death rate	hour^{-1}
g_I	Infected macrophage growth rate	hour^{-1}
d_I	Infected macrophage death rate	hour^{-1}
p	Probability an infected macrophage divides into 2 infected macrophages	scalar
β	Infection rate of healthy macrophage by extracellular bacteria	cell^{-1}hour^{-1}

Table 2.1: Variable and parameter symbols and their descriptions for equations 2.1 to 2.3.

of an infected macrophage results in its intracellular load (N) being returned to the extracellular population, seen in the $d_I N M_I$ term in equation 2.1.

By setting $\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\mathrm{d}M_H}{\mathrm{d}t} = \frac{\mathrm{d}M_I}{\mathrm{d}t} = 0$, the steady state of the system can be analysed. Eliminating the $\beta E M_H$ term from equations 2.1 and 2.3 results in

$$0 = \alpha_E E + \left[(N-1)\, d_I + g_I p\right] M_I. \tag{2.4}$$

Since $E \geq 0$, $M_I \geq 0$ and $[(N-1)\, d_I + g_I p] > 0$ for $N > 1$, it is trivial to conclude that the only steady state is $E = M_I = M_H = 0$. The assumption that $N > 1$ implies that the intracellular population of an infected macrophage at the time it dies is always greater than 1. Since the macrophages involved are not activated, it is a reasonable assumption that intracellular growth occurs within all macrophages that become infected.

Computing the Jacobian matrix for the system at the steady state results in the matrix

$$J = \begin{pmatrix} \alpha_E & 0 & d_I N \\ 0 & g_h - d_h & g_I(1-p) \\ 0 & 0 & d_I(p-1) \end{pmatrix} \tag{2.5}$$

for which the eigenvalues are $\lambda_1 = \alpha_E > 0$, $\lambda_2 = g_h - d_h$, and $\lambda_3 = d_I(p-1) > 0$. Since 2 eigenvalues are strictly positive, the fixed point of the system is not stable. If $g_H < d_H$ then the fixed point would be a saddle, potentially resulting in fluctuations in the cellular populations.

2.2 Data

Here, the data to be used in this chapter will be briefly described. First, the data relating to the extracellular bacteria growth is shown in its raw state in figure 2.2. Note the log scale on the y-axis. It is clear from this figure that the time series data here is non-stationary with a constant trend in the log-plot, thus the data are growing exponentially.

The remaining data are captured at the individual level and thus can be summarised as a population. Table 2.2 shows the summary statistics for these data. By far the most observed event in the data is the division of a healthy macrophage and overall approximately 150 more healthy macrophages are observed than infected macrophages. The time to death for both populations has a similar mean, however the time to death for infected macrophages has much higher variance. The same is true for division, however infected macrophages also have a longer average time to division.

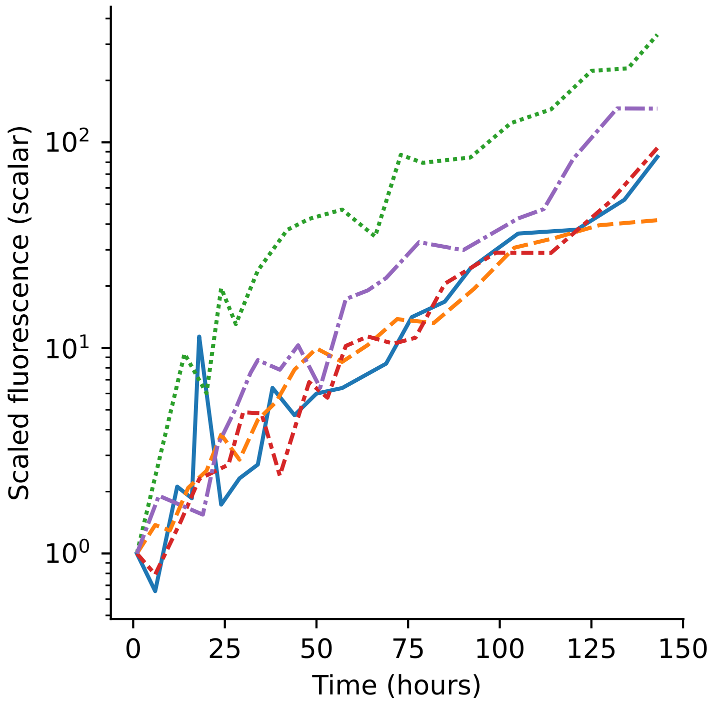

Figure 2.2: Raw data of extracellular bacteria growth. Note the log scale on the y-axis. Clearly these data are non-stationary and the constant trend in the log scale demonstrates exponential growth.

Mechanism	Count	Min (hours)	Mean (hours)	Max (hours)	Std (hours)
Healthy death	67	2.0000	22.0000	82.0000	15.2236
Healthy division	240	6.0000	18.9333	53.0000	6.9868
Infected death	65	1.0000	24.9385	126.0000	23.5524
Infected division	93	6.0000	23.3978	59.0000	13.7597
Infection	13	1.0000	17.7692	45.0000	11.4029

Table 2.2: Descriptive statistics for data used in this chapter. The values used here are the wait times, thus the time from when a macrophage was born to when it either divided or died.

2.3 Parameter Fitting

In this section the process of fitting parameters to the data will be described. The data is structured by experiment and macrophage. Key information is measured, such as time of division, time of infection as described in section 1.5. This structure of data lends itself to extracting distributions describing each mechanism, to which the parameter governing that mechanism can be fitted.

2.3.1 Macrophage division

Figure 2.3 shows the time from birth to division data for macrophages. The data is divided by the infection status of the macrophage at the time of division. The raw data is recorded at the cell level measuring the time of birth, the time of division (for those cells that divided) and the infection status of the cell. Data was captured over 11 experiments involving 47 macrophages. Healthy macrophages demonstrate a clear delay followed by a peak, whereas the infected macrophages seem to exhibit a flatter distribution. Exponential growth is a commonly used functional form for modelling populations [172], thus has been fitted here. While it is clear that neither of these are distributed according to an exponential distribution, since this chapter is presenting as simple a model as possible, exponential growth will continue to be assumed, allowing the ODEs to remain three dimensional. In future chapters more appropriate distributions will be considered.

The fitted distribution is an exponential distribution. Figure 2.3 shows the optimally fitted exponential distribution to each data set and the 95% confidence intervals for the fitted growth parameter. The optimal parameters were inferred using maximum likelihood estimate (MLE).

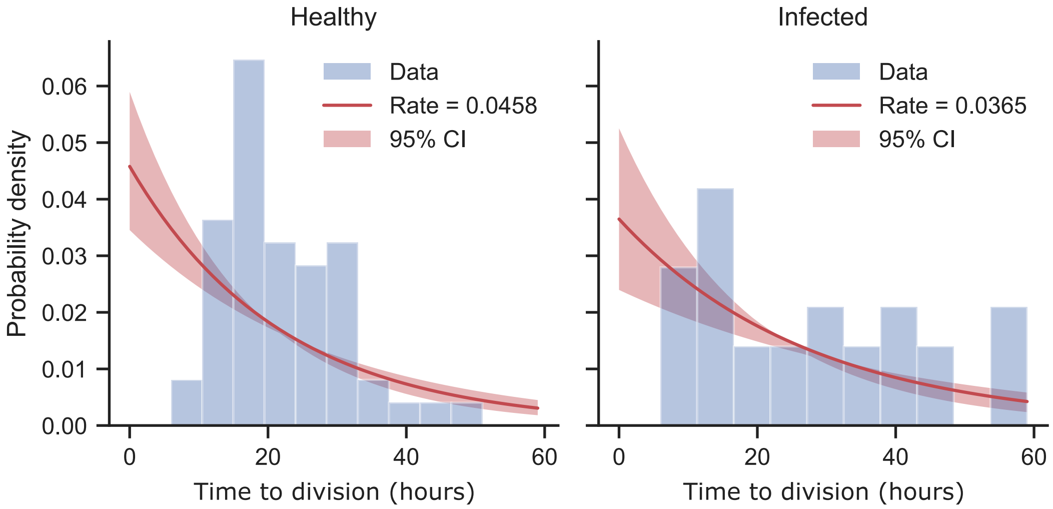

Figure 2.3: Division wait times for both healthy and infected macrophages are recorded, and demonstrate distinct distributions. Despite a poor fit, for simplicity of the model an exponential distribution is assumed and fitted using maximum likelihood and shown against the data along with its 95% confidence interval. Better fitting distributions will be investigated in future chapters.

2.3.2 Macrophage death

The available data for macrophage deaths is similar to that of macrophage division: birth and death times at the individual level. Plotting the data shows a clear exponential disappearance rate as shown in figure 2.4. Since birth rates are being modelled independently for both healthy and infected macrophages, the distributions for death times are also fitted independently. The fitted distributions are also shown in figure 2.4 along with the 95% confidence intervals.

On average, the time to death is very close to the time to division within each type of cell (infected or healthy), however the time to death has higher variance. This can be seen by the similar rate values that are fitted compared to the difference in maximum wait times - no division events occur over 60 hours post birth whereas there are death events observed up to 100 hours post birth. In subsequent chapters the model for macrophage division will be improved to account for these differences.

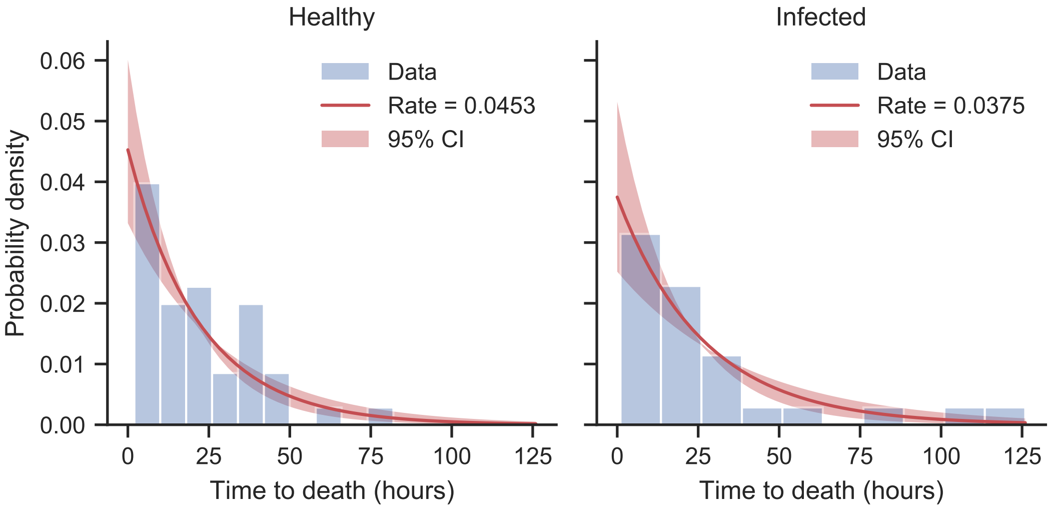

Figure 2.4: Birth to death wait times for both healthy and infected macrophages are recorded, and demonstrate similar distributions. Exponential distributions fit the data well. Since birth rates are distinct for both healthy and infected macrophages, death times will be kept distinct.

2.3.3 Bacteria growth

The available data for fitting the extracellular bacteria growth rate is a one-dimensional time series of the total area of fluorescence. There are five experiments available in which the extracellular bacteria are observed to grow without interaction with macrophages, these will be used to determine the model of growth for extracellular bacteria. Plotting the data (figure 2.5 (left)) shows that the level of fluorescence grows exponentially in time, thus a simple exponential growth model will be assumed. Assuming that the level of fluorescence is proportional to the total bacteria population, the exponential growth parameter governing the rate of growth of the fluorescence will be the same as that governing the growth of the bacteria.

Since the data is only available at the population level it is not possible to fit separate growth and death models for the bacteria, instead the exponential growth parameter will define the difference in growth rate and death rate. While additional experiments could have been performed with the purpose of separately measuring the independent rates of growth and death of extracellular bacteria, it is ultimately

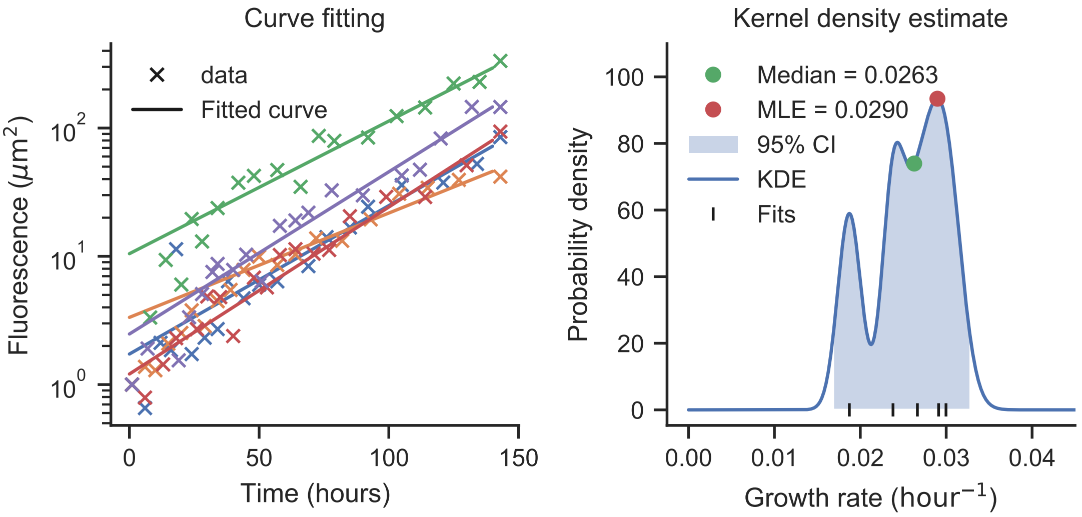

Figure 2.5: Data show exponential growth of extracellular bacteria, to which an exponential growth curve can be fitted. Averaging over the fitted exponential growth rates weighted by their errors results in the kernel density estimate. This distribution is used to find the optimal value for the extracellular growth rate.

the resultant growth rate of the difference between the two that is of interest. The model of growth used to fit the data is

$$E(t) = E_0 \exp(\alpha_E t) \,. \tag{2.6}$$

Allowing for uncertainty in the initial number of bacteria, both E_0 and α_E are estimated for each experiment. Figure 2.5 (left) shows the raw data as well as the fitted growth curves. For each exponential curve from the data, the fitting process yields a MLE as well as a standard error. Assuming a normal distribution on the errors, these can be used to define Gaussian kernels, $N(\mu_i, \sigma_i^2)$[1], which can be summed resulting in figure 2.5 (right). The MLE of the resultant distribution is used to define the value of α_E used in the model.

---

[1] μ_i represents the MLE for the growth rate, and σ_i^2 represents the normal distributed errors on the fitted growth curves.

2.3.4 Infection

In order to fit the rate of infection, the time of birth and the time of infection of macrophages is used. The movement of bacteria from the extracellular space to the intracellular space is not directly observed and is instead inferred from the observation of a healthy macrophage becoming infected. It is therefore not possible to distinguish between the intracellular load increasing due to intracellular growth and infected macrophages internalising additional bacteria. For this reason the model will assume only healthy macrophages internalise bacteria, and for simplicity only one bacteria is internalised at a time. This second assumption will be explored more in chapter 5. The force of infection, the rate at which healthy macrophages become infected, is assumed to be proportional to the population of extracellular bacteria, however this value is not tracked in parallel to infection times.

The experimental set up described in section 1.5 states that macrophages were washed to remove extracellular bacteria, thus it should be assumed that there are no extracellular bacteria at the beginning of the experiment. However, there are records of macrophages becoming infected prior to observing an infected macrophage dying, implying the existence of extracellular bacteria. This is likely a result of the camera that was used in the experiments not capturing the entire petri dish, thus it is possible for the death event to have occurred out of frame. The assumption of a bacteria population of zero result in the likelihood function evaluating to zero when including data that exhibits an infection event before an infected macrophage death event. Since the number of observed infection events is already limited, for the purpose of fitting the rate of infection, it will be assumed that there is a single extracellular bacteria at the start of the experiment. The bacteria population is then assumed to grow exponentially using the exponential growth rate already fitted in this chapter. The possible implications of this assumption are that the fitted infection rate will be too low, since the simulated macrophages are experiencing a higher force of infection than reality, but becoming infected in the same amount of time. The impact of this will be investigated in the sensitivity analysis.

At any given time the force of infection on a healthy macrophage is dependent on the infection rate β and the number of extracellular bacteria $E(t)$. Interactions are modelled using mass action, and so the probability of observing a healthy macrophage born at time t_0 which remains healthy until time t_1 can be written as

$$\mathbb{P}\left[M_H(t_0) \to M_H(t_1)\right] = \exp\left(-\int_{t_0}^{t_1} \beta E(s)\,\mathrm{d}s\right). \tag{2.7}$$

The data is recorded hourly, so when a macrophage is recorded as having changed state from healthy to infected, the infection event must have occurred sometime in the previous hour. Thus, the probability that a healthy macrophage born at time t_0 becomes infected between times $t_1 - 1$ and t_1 can be written as

$$\mathbb{P}\left[M_H(t_0) \to M_I(t_1)\right] = \mathbb{P}\left[M_H(t_0) \to M_H(t_1 - 1)\right]\Big(1 - \mathbb{P}\left[M_H(t_1 - 1) \to M_H(t_1)\right]\Big). \tag{2.8}$$

By iterating over the instances of a healthy macrophage becoming infected and evaluating the probability of each event according to the product of the above equations, the log of these probabilities can be summed to define a log-likelihood function of the parameter β. Since the level of extracellular bacteria is not recorded for these data, $E(s) = E_0 \exp\left(\alpha_E s\right)$ is assumed to be exponential growth with the growth rate fitted in section 2.3.3. For macrophages that are observed to become infected (not including those that were born infected) the time of birth is subtracted from the time of infection to give the wait time. Figure 2.6 shows the distribution of times from birth to infection and figure 2.7 shows the likelihood function of β along with its MLE and 95% confidence interval.

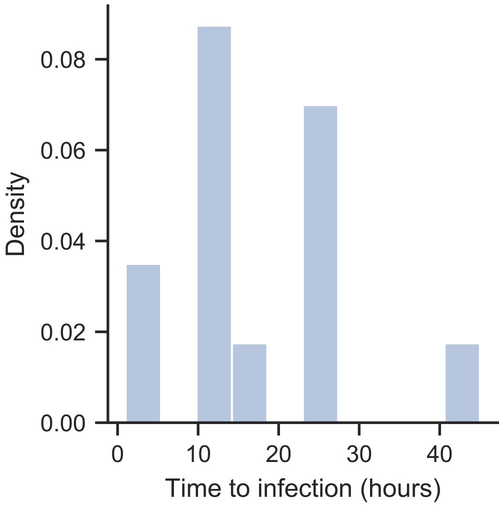

Figure 2.6: Distribution of time to infection wait times for healthy macrophages that were observed to become infected. Time to infection is the time of infection subtract time of birth. Macrophages that were born infected are not included.

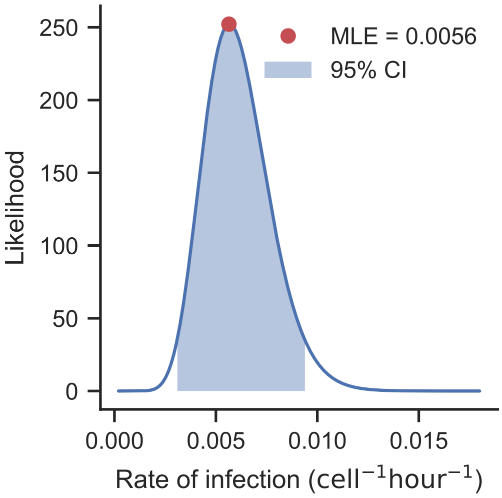

Figure 2.7: Likelihood function of the rate of infection β given birth and infection times of macrophages.

2.3.5 Infected macrophage division

When an infected macrophage divides there are two possible outcomes for the intracellular bacteria. Either the macrophage divides into an infected-healthy (IH) macrophage pair, or the macrophage divides into an infected-infected macrophage pair. The frequency of each outcome is shown in figure 2.8. Clearly the infected-healthy pair is the more likely outcome with a ratio of approximately 2:1. However it may be possible that the type of division a cell undergoes is influenced by the type of division its parent underwent, that is, the infected daughter resulting from an IH type division, may be more likely to divide into one infected and one healthy macrophage. By performing a Fishers exact test, the hypothesis of there being a correlation can be statistically evaluated. The contingency table is shown in table 2.3. The result of this test ($\chi = 0.5962$) is that the forms of division are not significantly dependent. Since there is no correlation in division type from one generation to the next, for simplicity in the model, a Bernoulli parameter ($p = 0.3$) is calculated using

$$p = \frac{\text{Number of infected-infected division events}}{\text{Total number of division events}}. \tag{2.9}$$

p is then used to determine which form of division takes place at the time of division (see figure 2.9).

	Self IH	**Self II**	**Total**
Parent IH	7	2	9
Parent HH	4	3	7
Total	11	5	16

Table 2.3: Contingency table for division of infected macrophages, investigating the relationship between the form of a cells division and that of their parents. Fishers exact test results is $\chi = 0.5962$ and thus not significant.

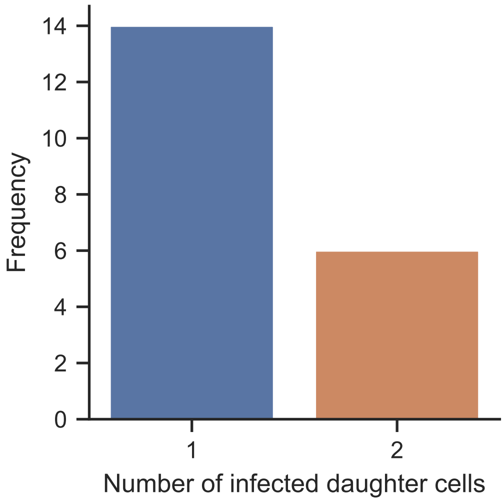

Figure 2.8: Frequency of infected cell division which results in 1 or 2 infected daughter cells.

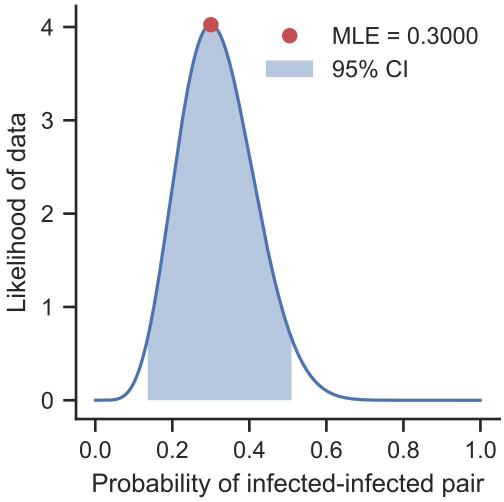

Figure 2.9: Likelihood and 95% confidence interval for probability of infected-infected division.

2.3.6 Parameter values

The fitted values and related confidence intervals are shown in table 2.4. The parameter N denotes the number of intracellular bacteria that are released when an infected macrophage dies. Since this value is not measured in the data, it cannot be fitted. A common value used in the literature for the maximal intracellular load is $N = 50$ [151, 157, 173]. To begin with, this value will be used in this model, but the effect of N will be analysed in section 2.4.1.

The majority of mathematical and biological literature focus on recruitment as the method of macrophage population replenishment. Studies investigating the interactions of macrophages and bacteria cells in vitro, without concern for recruitment [9, 174] do not explicitly measure the growth rate of the macrophages being studied. It is therefore not possible to compare the values regarding macrophage growth fitted here with the literature. The rate of growth of bacteria is very close to that of the literature (see table 1.3). Finally, the rate of infection of healthy macrophage features the most uncertainty, both here and in the literature. Pienaar and Lerm state an uptake range of between 0.0083 to 1.1667 bacteria per hour, and settle on 0.3333 [157], while Marino and Kirschner use a lower 0.01667 [151].

Parameter	Value	Confidence Interval	Unit
α_E	0.0290	(0.0168, 0.0328)	hour^{-1}
g_H	0.0458	(0.0401, 0.0521)	hour^{-1}
d_H	0.0453	(0.0332, 0.0601)	hour^{-1}
g_I	0.0365	(0.0271, 0.0472)	hour^{-1}
d_I	0.0375	(0.0252, 0.0532)	hour^{-1}
p	0.3000	(0.1361, 0.5095)	scalar
β	0.0056	(0.0031, 0.0094)	$\text{cell}^{-1}\text{hour}^{-1}$
N	50		cells

Table 2.4: Table of parameter values and confidence intervals. Parameter N was taken from literature [151] so no confidence interval is calculated. Note that since the values of g_H and d_H (respectively g_I and d_I) are close, the population of macrophages will remain approximately constant.

2.4 Model results

In this section the output of the model will be analysed. Since the model is a simple ODE model, it is computationally cheap to simulate. First the model is run with the parameters fitted in the previous section. The sensitivity of the model to these parameters will then be analysed, before investigating the direct relationships between parameters and what these relationships say about the system being modelled.

2.4.1 Model output

The model described in this section is simulated using an implementation of the VODE [46] integrator in the Scipy.integrate Python package [45]. First, parameter values are assumed to be their MLE values defined in the previous sections, then the parameters are allowed to vary within their 95% confidence intervals using Latin Hyper-Cube [175] sampling to efficiently cover the parameter space. The model output in figure 2.10 shows the simulated output of the model. After 140 hours the population of extracellular bacteria are growing exponentially, while the population

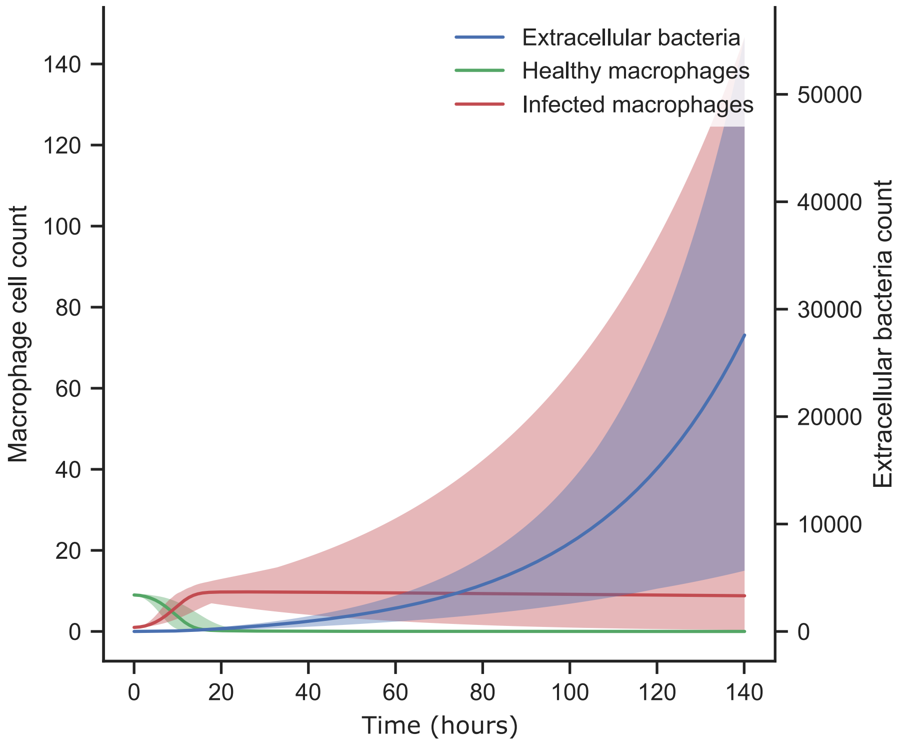

Figure 2.10: Integration of equations 2.1 to 2.3. Solid lines represent parameters set to their MLEs, while the shaded area is created by varying the parameters within their confidence intervals. Parameter sets were created using latin hypercube sampling (LHS) with 2000 sets of parameters chosen. The initial conditions for the model are $M_H = 9$, $M_I = 1$ and $E = 1$, this is in contrast to the experimental initial conditions in which macrophages were infected with a multiplicity of infection of 1.

of healthy macrophages has long been depleted due to the force of infection applied by the significantly larger population of bacteria. The fitted growth rate of the infected macrophages (g_I) is greater than the fitted death rate (d_I). Thus, the population of infected macrophages is able to persist. When the infected macrophages divided, a portion of them are born healthy, however due to the high force of infection from the large population of extracellular bacteria, these healthy macrophages are immediately converted back to infected macrophages.

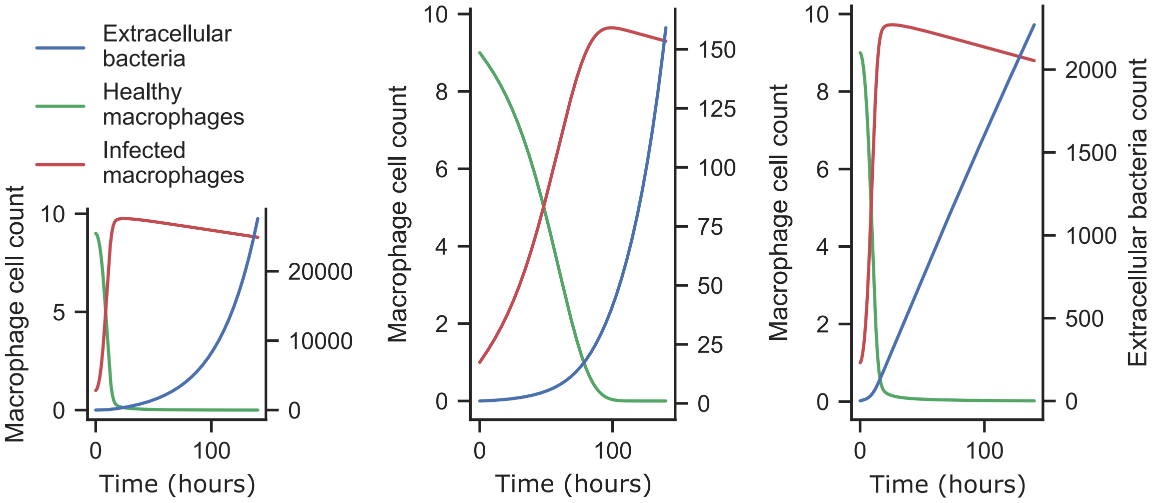

Figure 2.11: Left Base model as seen in figure 2.10 without the confidence intervals. **Centre** Model output without intracellular growth: $N = 1$ and $p = 1$. **Right** Model output without extracellular growth: $\alpha_E = 0$. The output of the model in the absence of extracellular growth shows the overwhelming effect that the implicit intracellular growth has on the model.

There is a very large overestimation of the number of extracellular bacteria. This is due to the assumption that the number of bacteria released when an infected macrophages dies is the average number of intracellular bacteria: $N = 50$ [151]. Since the model is an ODE it can be thought of as a mean field approximation of the true stochastic events. An infected macrophage dying with a high load would be a low probability event, however in the ODE construction, there is a constant positive pressure on the population of extracellular bacteria. This in turn increases the rate at which healthy macrophages are becoming infected resulting in a positive feedback loop that pushes the extracellular population up. To investigate the effect of the value of N, the model is run

1. In the absence of intracellular growth: set $N = 1$ and $p = 1$
2. In the absence of extracellular growth: set $\alpha_E = 0$.

In the absence of extracellular growth, the extracellular population grows approximately linearly. This is a result of the first two terms of equation 2.1 being zero (the healthy macrophage population is immediately depleted and does not recover) and the population of infected macrophages approximately constant. By reducing the value of N, the output of the model begins to look more reasonable and more comparable to what would be expected given the data. Under the naive assumption that intracellular bacteria grow at the same rate as extracellular bacteria, as roughly seen in the literature review (see tables 1.3 and 1.4) then at the end of the 140 hours, the total expected number of bacteria in the system will be $E_0 \exp(0.0290 \times 140) \approx 116$. Figure 2.12 shows the output of the model, with the parameter N changed to 2.

The assumption that $N = 2$ implies that when the average infected macrophage dies, it contained two intracellular bacteria. This is significantly below the values used in the mathematical literature. Since the model being presented here is such a simple model, there is only one mechanism promoting cellular death: the macrophage birth death cycle. This is in contrast to, for example, Marino and Kirschner [151] in which there are a host of factors from the adaptive immune response resulting in the death of infected macrophages. Thus despite the intracellular load being limited by N, there will often be intracellular loads significantly smaller than N at the time of the host cell's death. Another contributing factor is the higher than usual rate of division observed in the cells used in this experiment. While the mathematical literature generally assumes that macrophage populations are maintained by recruitment [93] experimentation does demonstrate macrophage division [92], as observed in this data set. The process of infected macrophage cell division increases the infected macrophage population by p at a rate of g_I, thus the intracellular population grows by pN at a rate of g_I. When $N = 50$ this is significantly higher than is realistic. If, instead, the macrophage population was maintained by recruitment, the population of intracellular bacteria would no longer grow with the macrophage population. For the rest of this section, N will be assumed to be 2.

While it is not possible to write down an analytical solution to equations 2.1 to

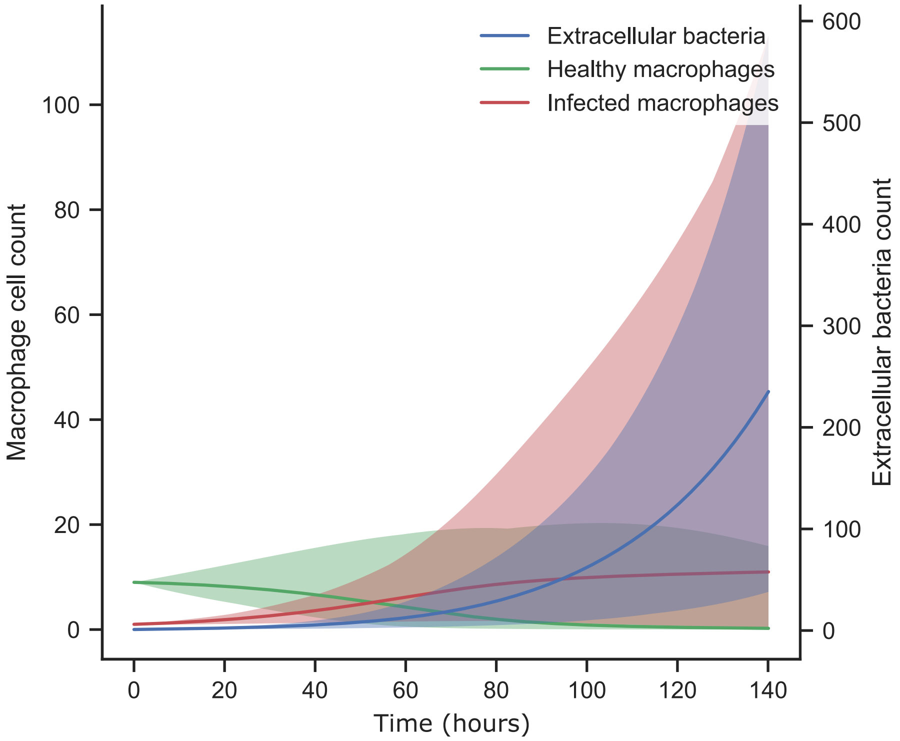

Figure 2.12: As in figure 2.10, integration of equations 2.1 to 2.3. Solid lines represent parameters set to their MLEs, while the shaded area is created by varying the parameters within their confidence intervals. Parameter sets were created using LHS with 2000 sets of parameters chosen. The initial conditions for the model are $M_H = 9$, $M_I = 1$ and $E = 1$. The parameter N has been updated to the value of 2, which yields the most qualitatively realistic model output.

2.3, the late time dynamics seen in 2.12 can be investigated. Assuming that E is large and M_H is small, some terms from equations 2.1 can be ignored. Since E is significantly larger than M_H and M_I, the growth of E is dominated by the first term.

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \alpha_E E - \beta E M_H + d_I N M_I \approx \alpha_E E. \tag{2.10}$$

Similarly since the population of healthy macrophages M_H has been depleted, the rate of change of M_H becomes

$$\tfrac{\mathrm{d}M_H}{\mathrm{d}t} = g_H M_H + g_I(1-p)M_I - \beta E M_H - d_H \tag{2.11}$$

$$\approx g_I(1-p)M_I - \beta E M_H.$$

Here the $\beta E M_H$ term cannot be ignored as while M_H is small, E is large, thus how the product EM_H acts depends on how fast E is diverging to infinity and M_H is converging to zero. By adding equation 2.3 to 2.11 an approximation for the total macrophage population ($M = M_H + M_I$) can be determined.

$$\begin{aligned} \tfrac{\mathrm{d}M}{\mathrm{d}t} &= \tfrac{\mathrm{d}M_H}{\mathrm{d}t} + \tfrac{\mathrm{d}M_I}{\mathrm{d}t} \\ &\approx g_I(1-p)M_I - \beta E M_H + (g_I p - d_I)M_I + \beta E M_H \\ &= (g_I - d_I)M_I. \end{aligned} \tag{2.12}$$

Thus whether the population of macrophages ultimately dies out is dependent on the sign of $g_I - d_I$.

One artefact of this model seen in figure 2.12 is the upper boundary on the 95% confidence interval of infected macrophages continuously increasing. This results from areas in the parameter space where $g_I > d_I$ and the lack of stable equilibrium away from zero, however it is not a result that would translate from in vitro to in vivo. Once the infection gets out of the control of the innate macrophages, additional factors will come into play, such as necrosis [119], requiring the adaptive immune response to counter [107]. This overestimate of the 95% confidence intervals for the birth death process of the macrophages is likely a result of the small amounts of data used in this study. This problem will be rectified in later chapters as the in vitro experimentation progress from its preliminary stages.

2.4.2 Sensitivity

Performing a Sobol sensitivity analysis as described in section 1.6.5 yields the sensitivity indices shown in figure 2.13. Variation in the total population of each type of cell is measured at uniformly sampled times of every 5 hours and attributed to variation in the input parameters. A limit of 140 hours was chosen as this was the limit of the original experiments. Later chapters will extend this limit to 200 hours. The system progresses through three stages, initially the system is driven by healthy macrophages internalising extracellular bacteria and becoming infected. Once the healthy macrophage population has been depleted, the system is driven by infected macrophages dying and releasing their intracellular loads back into the extracellular space. Finally, once the extracellular bacteria population has grown sufficiently large, the exponential growth of the bacteria dominates the system for the remaining time.

The dynamics of the extracellular bacteria are dominated by the death rate of infected macrophage (d_I) and the extracellular growth rate (α_E). However, initially there is a significant contribution from the rate of infection (β). This contribution is sustained only as long as there are no infected macrophages in the system. Once the population of infected macrophages begins to increase, the rate of death of infected macrophages quickly suppresses the effects of infection of the extracellular population. This is due to the rate of death of infected macrophages being significantly higher than the rate of infection. After approximately 75 hours, the extracellular population has grown large enough that the exponential growth governed by α_E overtakes d_I as the dominant parameter and the system enters its final stage of exponential and self sustained extracellular bacteria growth.

There is a larger amount of uncertainty in the value fitted for the healthy macrophage death rate (d_H) than the healthy macrophage growth rate (g_H). The consequence of this uncertainty is the larger contribution to the variance in the macrophage population resulting from variance in d_H. As the population of infected macrophages grows, the number of healthy macrophages being born as a result of infected-healthy division events steadily increases. This can be seen in the steady increase in the significance of the death rate of infected macrophages d_I.

Naturally, the initial parameter contributing to the population of infected macrophages is the rate of infection (β). The significant effect of the rate of death of healthy macrophages (d_H) on the healthy macrophage population has a knock on effect onto the infected macrophage population. As the extracellular population grows large, healthy macrophage are converted to infected macrophage almost immediately. Thus a larger death rate for healthy macrophages will result in a lower population of infected macrophages. As the population of healthy macrophages begins to die out, the death rate of infected macrophages (d_I) takes over as the dominant parameter. Throughout the simulation, the effect of the growth rate of infected macrophage (g_I) steadily increases, however this is more of a result of the impacts of β and d_H reducing rather than those of g_I growing. Similarly, the impact of g_H steadily grows, until the healthy macrophages die out. This relationship stems back to infection, when there are more healthy macrophages to infect, there will be more healthy macrophages converting to infected macrophages.

In order for there to be a reduction in the number of extracellular bacteria ($\frac{\mathrm{d}E}{\mathrm{d}t} < 0$) immediately post infection ($M_I \approx 0$), M_H must be greater than ${}^{\alpha_E}/_{\beta}$. The immediate effect of this is the reduction in the number of healthy macrophages and the increase in number of infected macrophages, in turn increasing the number of extracellular bacteria. Eventually, E will be sufficiently large and M_H sufficiently small that $\frac{\mathrm{d}E}{\mathrm{d}t} > 0$. In figure 2.12 the initial conditions include an initially infected macrophage, so this process is not observed. Measuring the time taken until $\frac{\mathrm{d}E}{\mathrm{d}t}$ switches sign from negative to positive, figure 2.14 shows how variation in the parameters affects this value.

In order of size of contribution, the infection rate β, the infected macrophage death rate d_I, and the extracellular growth rate α_E tell the story of which mechanisms are involved in limiting or increasing the growth of extracellular bacteria. By increasing the rate at which macrophages internalise bacteria, the extracellular population is diminished, halting its growth. Conversely, an increase in the death rate of infected macrophages and of the extracellular growth rates, push bacteria back into the extracellular space and allow them to grow faster than they can be internalised, resulting in the rapid switch back to $\frac{\mathrm{d}E}{\mathrm{d}t} > 0$.

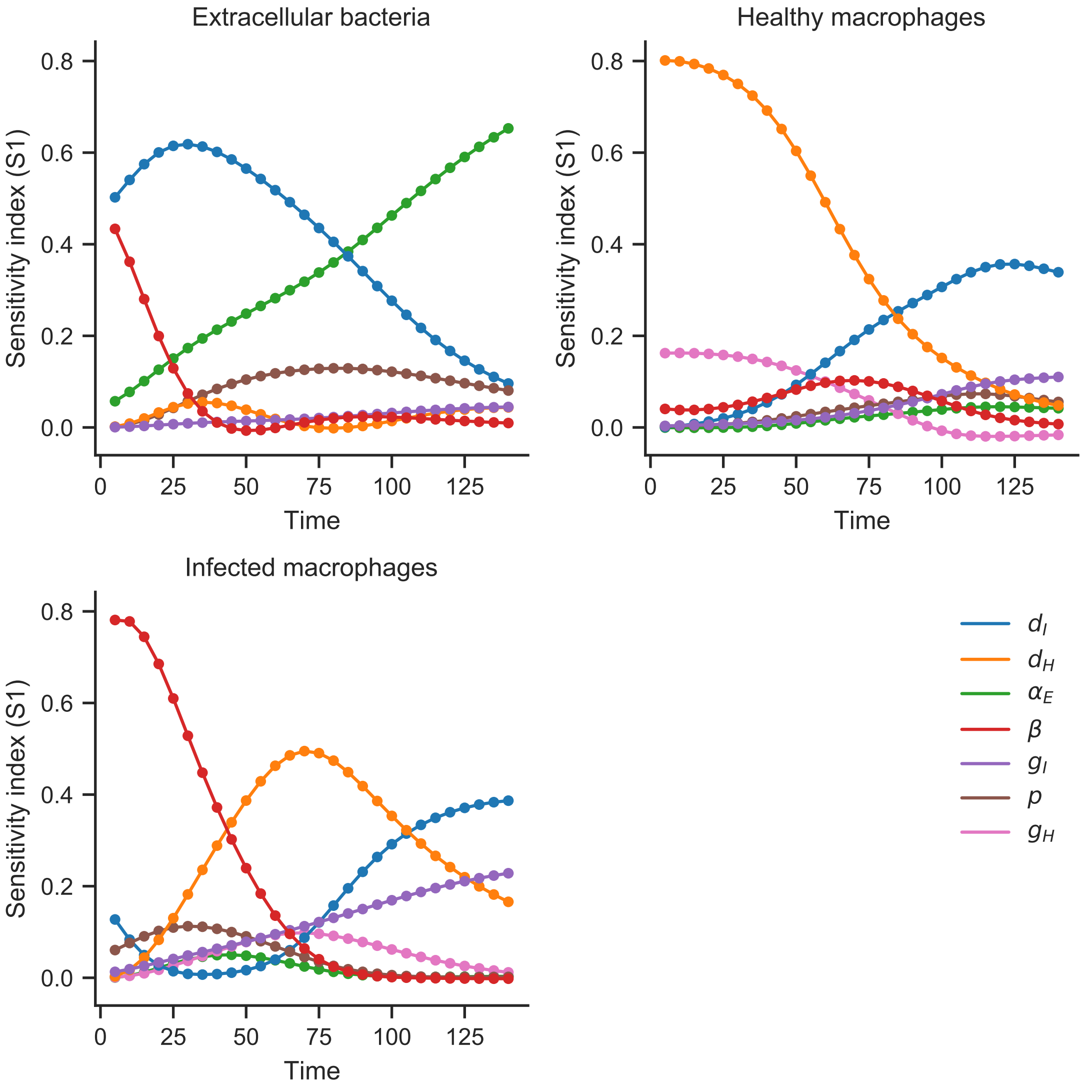

Figure 2.13: Sobol sensitivity analysis measuring how much variance in the model populations can be attributed to variance in the model parameters. Populations are sampled at evenly spaced time points to investigate the sensitivity to parameters throughout the whole time course.

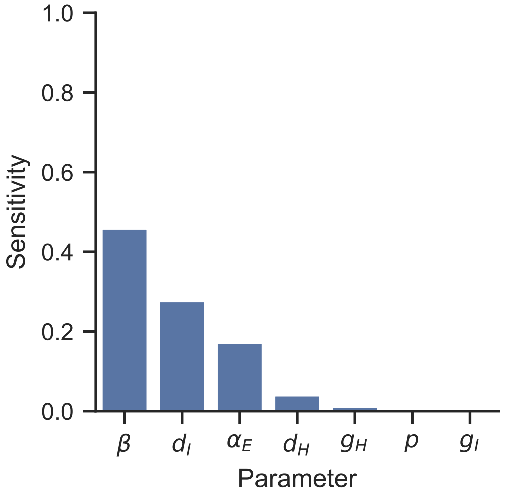

Figure 2.14: Sobol sensitivity analysis bounding the parameters within their 95% confidence intervals. Measured is the fraction of the variance in the value of the time at which $\frac{\mathrm{d}E}{\mathrm{d}t}$ switches sign from negative to positive, which can be attributed to variance in the given parameter.

2.4.3 Identifiability

Analysing the identifiability as described in section 1.6.7 amounts to determining whether the MLEs are unique for each mechanism. Macrophage growth and macrophage death are simple one parameter probability distributions fitted to data, thus it is trivially true that the MLEs are unique. For the kernel densitiy estimate (KDE) defined for the extracellular growth, the likelihood function is plotted in figure 2.5 and shows a unique maximum. The same is true of the infection rate in figure 2.7 and the fraction of infected-infected divisions in figure 2.8. Thus, since all MLEs are unique it can be concluded that the model is identifiable.

2.5 Conclusions

The goal of this chapter was to explore the possibilities for which mechanisms could be modelled and reasonably fitted using data collected from the experiment described in section 1.5. To that end, a range of simple mechanisms for macrophage and bacterial growth and their interactions have been fitted, and a simple three equation ODE model was developed. A sensitivity analysis was performed evaluating how variance in the model parameters affects variance in the state of the system. In the first 70 hours of simulation the model is more sensitive to the macrophage dynamics and the rate of infection, however as the bacteria population grows, the bacteria growth rate takes over as the dominant mechanism. A second analysis was performed measuring the time until the rate of change of extracellular bacteria is positive, for which the rate of infection was the dominant mechanism. Delaying the growth of the extracellular bacteria would give a host system more time to develop any required adaptive response, demonstrating the importance of early capture. Finally a steady state analysis showed there are no non-zero steady states in the model, and the long term dynamics of the model depend on the ratio of g_I to d_I.

Intracellular bacteria are not explicitly modelled, instead the parameter N was used to define the number of bacteria that is released when an infected macrophage dies. The interaction of this parameter with the birth death process of the macrophages highlighted the importance of explicitly modelling intracellular bacteria. This is a similar approach taken to that of Wood, Egan, and Hall [176]. The model presented in [176] is a model of Francisella tularensis wherein intracellular bacteria are also modelled implicitly and a fixed number of bacteria are released at cell death, this number being the solution to the expected number of intracellular bacteria at the expected time of rupture. There are two key differences between the models. Firstly, the population of healthy phagocytes in [176] is infinite, an assumption of the within-host setting of the model where phagocytes vastly outnumber the bacteria. This contrasts to the finite population of macrophages modelled here where by the population of healthy macrophages gets depleted and the bacteria are able to persist within the infected macrophage population. Secondly, the phagocytes in [176] are readily able to kill the bacteria with a death rate significantly higher than the bacteria survival rate. The base model developed here does not include any mechanism to kill the bacteria, and this process is only explored through computational simulation.

The model presented here will be improved over the following chapters. First and foremost, the intracellular bacteria will be modelled explicitly, and the macrophage division process will be modelled using a more suitable distribution than exponential. Fundamentally, the model presented in this chapter performed as expected. As a simple ODE model with only 1 non-linear term and no non-trivial steady states, the resulting dynamics are simple to deduce from the model definition. However there is value in the development process as those parameters that are fitted and can be compared to the literature, such as the rate of infection, are in line with other estimates [100, 118–120, 144].

Chapter 3

Model 2: Differential Equations With Explicit Intracellular Bacteria and Erlang Growth

In this chapter the model from chapter 2 will be further developed and extended. Most notably intracellular bacteria will be modelled explicitly. Additionally the macrophage birth death process will be improved from simple exponential growth. While exponential growth is often a good choice for modelling population growth, it may not be appropriate at the individual level. Since the population of macrophages is small, it is important to accurately capture the individual dynamics.

Traditionally the log-normal distribution has been used to model the cell cycle [177, 178], however more recently the Erlang distribution has been introduced as an effective model for cell cycle times [179, 180] and it has been shown to accurately capture human cell cycle dynamics [181]. Since the Erlang distribution can be thought of as a series of exponential distributions, it fits well mechanistically with the current biological understanding of the cell cycle: cells move through a series of stages before undergoing cell division. Additionally, its relationship to the exponential distribution makes the Erlang distribution a convenient model for implementing in both ordinary differential equations (ODEs) and agent-based models (ABMs). For these reasons the Erlang distribution will be used over the log-normal.

3.1 Erlang growth

The Erlang distribution is a two parameter family of distributions over the positive real numbers, whose parameters are a positive integer k and a positive real number λ. It is a special case of the Gamma distribution: having the same functional form with the restriction that k is an integer. When $k = 1$ the Erlang distribution is simply the Exponential distribution, and for $k > 1$, it is the sum of k Exponential distributions each with mean $1/\lambda$ [182, p. 152]. A Poisson process describes events that occur independently at some fixed rate, the classic example being buses arriving at a bus stop. The Exponential distribution is used to describe the wait times between events being modelled by a Poisson process [182, p. 339], thus the Erlang distribution can be used to describe wait times between k occurrences of these events. By considering a macrophage to have a series of stages of life which it must pass through in order to divide, and by assuming that the rate at which the macrophage traverses these stages is fixed and the same, then the time from birth to division will be an Erlang distributed random variable.

The validity of the above assumptions is founded in the biological literature. The cell cycle is described in biological literature as a series of four consecutive stages: cell growth (G1), DNA synthesis (S), interphase (G2), and mitotic phase (M) [181, 183], however the transition rates between these stages is not necessarily be the same. Despite this, the whole process can be well approximated using the Erlang distribution. Figure 3.1 shows how an Erlang distribution can be used to model a here-named inhomogeneous Erlang process, with the assumption that while the rates between the four stages may not be the same, they are fixed. Testing the goodness of fit of the 100,000 samples of the inhomogeneous Erlang process against the homogeneous Erlang process using the Kolmogorov-Smirnov test, as described in section 1.6.6, results in no rejection of the hypothesis that they are the same distribution.

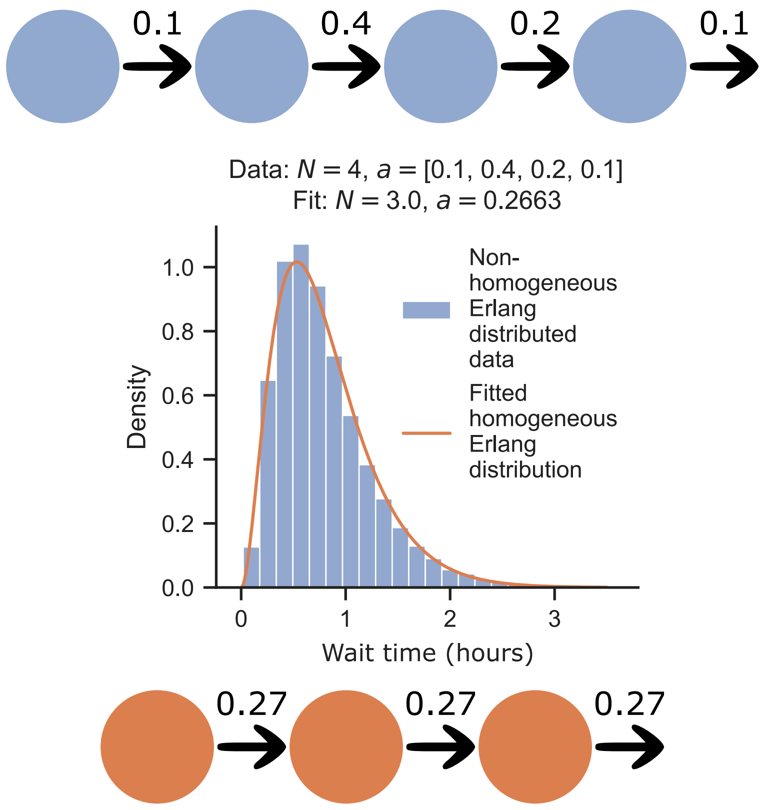

Figure 3.1: A non-homogeneous Erlang process can be well approximated by a homogeneous Erlang process. (**Top**) A four stage process consisting of varied transition rates. (**Centre**) Wait times to progress through all four stages are sampled and plotted. These wait times are used to fit a homogeneous process. (**Bottom**) The three stage process with homogeneous transition rates is a good approximation for the four stage process.

3.2 Model description

In this section, the additional mechanisms integrated into the model will be explained, along with how they affect the mathematical construction of the model. Fundamentally, the model remains a system of ODEs implementing mass action mixing for modelling the interaction between cells. A notable difference will be the variable number of equations in the system, which will be driven by the number of Erlang stages the macrophages must pass through and the maximum number of intracellular bacteria.

3.2.1 Mechanisms

The additional mechanisms included in this chapter are explicit intracellular growth, and Erlang growth for the macrophages. Both of these processes can be fitted to data directly observed in the experiments.

Macrophage birth death process

Both healthy and infected macrophages are observed to undergo division in the experiments. For healthy macrophages, this results in an additional healthy macrophage being added to the system. For infected macrophages, there is a chance that the new macrophage will be either healthy or infected. As modelled in the previous chapter, it will be assumed that a fixed proportion of infected cells divide to form two infected daughter cells. Additionally, since intracellular bacteria are now being modelled explicitly, when an infected macrophage divides into two infected daughter cells, the number of intracellular bacteria will be divided between the two new cells evenly.

Intracellular bacteria growth

Intracellular bacteria are observed in the data through fluorescence levels over time in an infected macrophage. They are observed to grow intracellularly according to exponential growth. While the literature suggests there is a limit to the number of bacteria that are able to remain contained within a single macrophage [118, 134, 135, 184], the data do not show any sign of logistic growth. This suggests that the bacteria continue to grow until the macrophage cannot contain them any longer, at which point it bursts. The exact cause of death of the macrophage is difficult to determine with the methodologies used in the experiment, since this is not the target of the study. In some experiments it is observed that the bacteria continue to proliferate on the remains of the dead macrophage. Since macrophage apoptosis is considered to be a defence mechanism [127], the survival and proliferation of the bacteria supports the assumption that the macrophage dies as a result of the intracellular bacteria growth.

3.2.2 Model definition

The structure of the model is a non-linear system of ODEs. The model tracks the populations of extracellular bacteria (E) and macrophages (M_k^i) where k denotes the number of intracellular bacteria the macrophage contains, and i denotes the Erlang stage of life the macrophage is at. The equations are defined as follows:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \alpha_E E - \beta E \sum_{i=0}^{N_H} M_0^i + d_I \sum_{i=0}^{N_I} \sum_{k=1}^{N} k M_k^i + \alpha_I N^2 \sum_{i=0}^{N_I} M_N^i \tag{3.1}$$

$$\frac{\mathrm{d}M_0^0}{\mathrm{d}t} = 2 g_H M_o^{N_H} + g_I M_1^{N_I} + (1-p) g_I \sum_{k=2}^{N} M_k^{N_I} - \beta E M_0^0 - d_H M_0^0 - g_H M_0^0 \tag{3.2}$$

$$\frac{\mathrm{d}M_0^i}{\mathrm{d}t} = g_H M_0^{i-1} - \beta E M_0^i - d_H M_0^i - g_H M_0^i \tag{3.3}$$

$$\begin{aligned} \frac{\mathrm{d}M_1^0}{\mathrm{d}t} &= p g_I \left(2 M_2^{N_I} + M_3^{N_I} \right) + g_I M_1^{N_I} + \beta E \sum_{i=0}^{N_H} M_0^i \\ &\quad - d_I M_1^0 - g_I M_1^0 - \alpha_I M_1^0 \end{aligned} \tag{3.4}$$

$$\begin{aligned} \frac{\mathrm{d}M_k^0}{\mathrm{d}t} &= p g_I \left(M_{2k-1}^{N_I} + 2 M_{2k}^{N_I} + M_{2k+1}^{N_I} \right) + (1-p) g_I M_k^{N_I} + \alpha_I (k-1) M_{k-1}^0 \\ &\quad - d_I M_k^0 - g_I M_k^0 - \alpha_I k M_k^0 \end{aligned} \tag{3.5}$$

$$\frac{\mathrm{d}M_k^i}{\mathrm{d}t} = g_I M_k^{i-1} + \alpha_I (k-1) M_{k-1}^i - d_I M_k^i - g_I M_k^i - \alpha_I k M_k^i \tag{3.6}$$

$$\begin{aligned} &i \in \{1, \ldots, N_H\} \quad \text{if } k = 0 \\ &i \in \{1, \ldots, N_I\} \quad \text{if } k \in \{1, \ldots, N\} \end{aligned}$$

Symbol	Unit	Description
E	cells	Extracellular bacteria
M_k^i	cells	Macrophages with k intracellular bacteria and in stage i of life
α_E	hour^{-1}	Extracellular growth rate
α_I	hour^{-1}	Intracellular growth rate
N	cells	Maximum number of intracellular bacteria
g_H	hour^{-1}	Healthy macrophage Erlang growth rate
N_H	scalar	Healthy macrophage Erlang growth stages
d_H	hour^{-1}	Healthy macrophage death rate
g_I	hour^{-1}	Infected macrophage Erlang growth rate
N_I	scalar	Infected macrophage Erlang growth stages
d_I	hour^{-1}	Infected macrophage death rate
β	$\text{cell}^{-1}\text{hour}^{-1}$	Infection rate of healthy macrophage by extracellular bacteria
p	scalar	Proportion of infected macrophages that divide to produce two infected daughter cells. The intracellular bacteria are assumed to divide evenly between the daughter cells.

Table 3.1: Variable and parameter symbols and their descriptions for equations 3.1 to 3.6.

There are eight individual mechanisms being modelled in equations 3.1 to 3.6. In Appendix A each of these mechanisms have been colour coded and the relevant terms have been coloured to match the mechanisms. Table 3.1 defines the variables and parameters used.

Extracellular growth The extracellular bacteria remain being modelled by exponential growth. The $\alpha_E E$ term in equation 3.1 controls this growth.

Infection Infection events occur between healthy macrophages and extracellular bacteria. Any healthy macrophage can become infected, regardless of what stage of Erlang growth they are at. The number of Erlang stages is not the same between healthy and infected macrophages, thus it is not well defined which class a healthy macrophage should move to post infection. For simplicity in the model, when a healthy macrophage does become infected it is always moved to the M_1^0 class of macrophage. The effect of this assumption is investigated in section 3.5.4. Healthy macrophages and extracellular bacteria interact at a density dependent rate under a homogeneous mixing assumption. The rate parameter β is assumed to be the same for all equations despite having different units, as in the previous chapter.

Healthy macrophage growth Healthy macrophages grow by transitioning at a fixed rate (g_H) through a fixed number (N_H) of Erlang growth stages before dividing. When a healthy macrophage $M_0^{N_H}$ divides it is removed from the system and two new macrophages M_0^0 are added.

Healthy macrophage death Healthy macrophages die at a fixed rate d_H. This rate is independent of the age of the macrophage, as there was not enough evidence in the data to support modelling the death rate as age dependent.

Infected macrophage growth Infected macrophages also grow by transitioning at a fixed rate (g_I) through a fixed number of Erlang growth stages (N_I) before dividing. When an infected macrophage $M_k^{N_I}$ divides it is removed from the system and two new macrophage $M_{k_1}^0$ and $M_{k_2}^0$ are added. k_1 and k_2 are chosen such that $k_1 + k_2 = k$. To understand how k_1 and k_2 are chosen it is easiest to understand by breaking the possibilities into cases.

Case 1 When $k = 1$ there is only one option for the two daughter cell. The first is healthy M_0^0 and the second has a single intracellular bacteria M_1^0. This case refers to the $g_I M_1^{N_I}$ terms in equations 3.2 and 3.4.

Case 2 When $k > 1$, with rate $1 - p$, the division process results in one healthy macrophage M_0^0 and one infected macrophage M_k^0. This can be seen in the terms involving $(1 - p)$ in equations 3.2 and 3.5.

Case 3 When $k > 1$, with rate p, the division process results in two infected macrophages. If $k = 2l + 1$ is odd, then the division is as even as possible such that $k_1 = l$ and $k_2 = l + 1$. Otherwise if $k = 2l$ is even then the division is perfectly even such that $k_1 = k_2 = l$. This can be seen in the first term of equations 3.5, where new cells are produced from three different parent compartments, corresponding to the three possible initial loads for daughter cells.

The assumption that the intracellular load divides either perfectly symmetric or perfectly asymmetric is for simplicity within this chapter. Future chapters will investigate allowing the division to vary.

Infected macrophage death When an infected macrophage dies, it releases its intracellular load back in to the extracellular space, thus the rate at which the extracellular population grows due to macrophage death is proportional to the number of infected macrophages and the load of those macrophages, as seen in the $d_I \sum_{i=0}^{N_I} \sum_{k=1}^{N} k M_k^i$ term in equation 3.1.

Intracellular growth The intracellular bacteria growth is modelled as exponential growth with rate parameter α_I. When an intracellular bacteria divides, its host macrophage transitions from class M_k^i to class M_{k+1}^i. The maximum intracellular load for a macrophage is N, so when an intracellular bacteria divides within a macrophage of class M_N^i, the macrophage bursts and the intracellular bacteria are released back in to the extracellular population. This can be seen in the $\alpha_I N^2 \sum_{i=0}^{N_I} M_N^i$ term in equation 3.1. The N^2 term is a result of the intracellular growth being proportional to the number of intracellular bacteria N, and the number of intracellular bacteria being released back to the intracellular space being N.

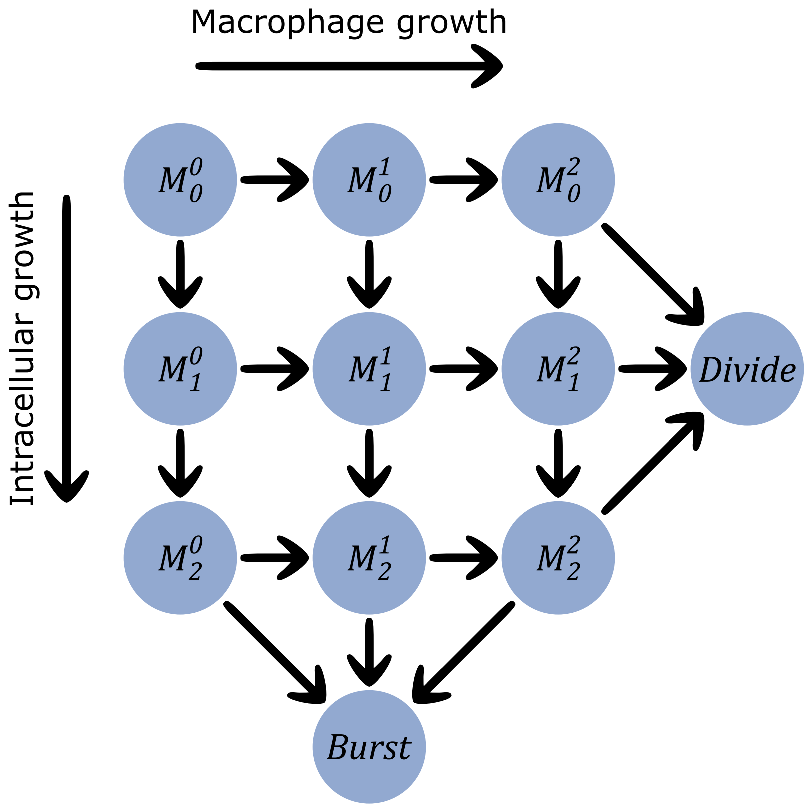

Figure 3.2: Illustrative example of the macrophage compartments M_k^i where k is the number of intracellular bacteria and i is the Erlang stage of life. As the macrophages progress through Erlang stages of life towards division, they progress through compartments $M_k^i \rightarrow M_k^{i+1}$. Similarly, as the intracellular bacteria grow, the macrophages progress through compartments $M_k^i \rightarrow M_{k+1}^i$. When a macrophage divides, it will be removed from the system and two new daughter cells are born in its place, each with an Erlang age of zero. If the parent cell was infected, the intracellular load is divided between the daughter cells as described in section 3.4.4

Figure 3.2 shows the structure of the model compartment M_k^i. As the macrophage progresses through stages of Erlang growth, it will progress through the compartments $M_k^i \rightarrow M_k^{i+1}$. As the intracellular bacteria grow, the macrophage will progress through the compartments $M_k^i \rightarrow M_{k+1}^i$.

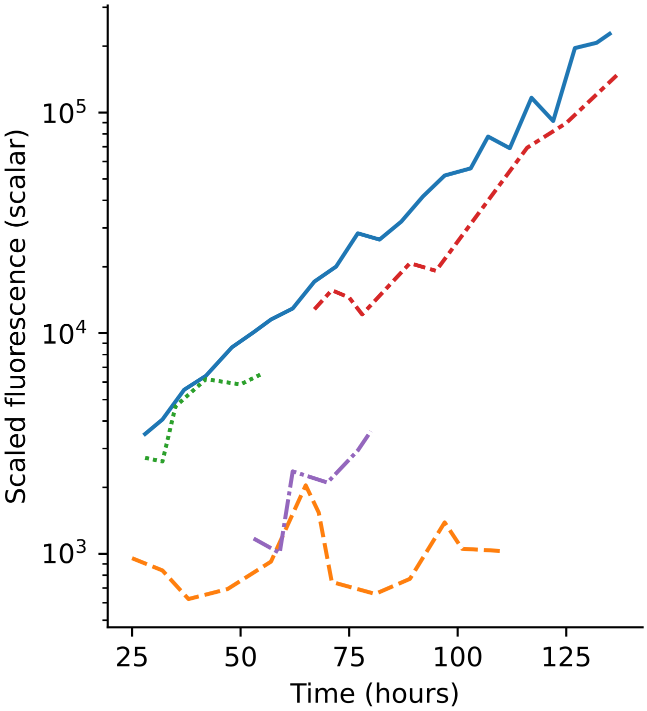

Figure 3.3: Raw intracellular growth time series data. Note the log scale on the y-axis. There are very few data points to consider for the intracellular growth, so all will be included.

3.3 Data

The only change in data used for this chapter compared to the previous chapter is the inclusion of the intracellular growth. There are relatively few data points available for this, however this is greatly improved in the subsequent chapters. Since there are so few data points, they will all be considered for this chapter. Similar to the exponential growth in the previous chapter, most of the data exhibit strong non-stationarity as well as clear exponential growth shown in figure 3.3.

3.4 Parameter Fitting

The model presented in this chapter is fitted to the same data as used in chapter 2. Therefore, only the parameters that are affected by the changes to how mechanisms are being modelled require fitting or refitting. Those parameters govern the macrophage birth-death process, the intracellular growth rate, and how intracellular bacteria are divided when an infected macrophage divides.

3.4.1 Macrophage birth death process

The model explicitly measures the number of intracellular bacteria, as well as applying an Erlang growth model to the macrophages. An infected macrophage is represented by M_k^i where k denotes the number of intracellular bacteria and i denotes the stage in the Erlang growth the macrophage is at. The rate at which the macrophage progresses through growth stages is dependent on the infection status of the macrophage. When $k = 0$ the macrophage is healthy and has growth parameters $g = g_H$ and $N = N_H$. When $k > 0$, the macrophage is infected and has growth parameters $g = g_I$ and $N = N_I$. The macrophage progresses through these stages at the fixed rate g until it reaches the final stage M_k^N. From stage N the macrophage divides at a constant rate resulting in two new macrophages both in stage 0 of life.

The data available for fitting the birth death process is the time that a macrophage was born and the time that the macrophage divided. The time to division is defined as the difference between these two times, this distributions for which are shown in figure 3.4. Maximum likelihood estimate (MLE) is used to identify the optimal parameters as well as the 95% confidence intervals, where the likelihood function is simply the sum of the probability density function over the data point for a given set of parameters. These are plotted over the data in figure 3.4. Since the Erlang distribution is a two parameter distribution, with one of those parameters being an integer, the 95% confidence interval is actually a set of intervals, each representing a range of values of g for a given value of N. These ranges are shown in table 3.3.

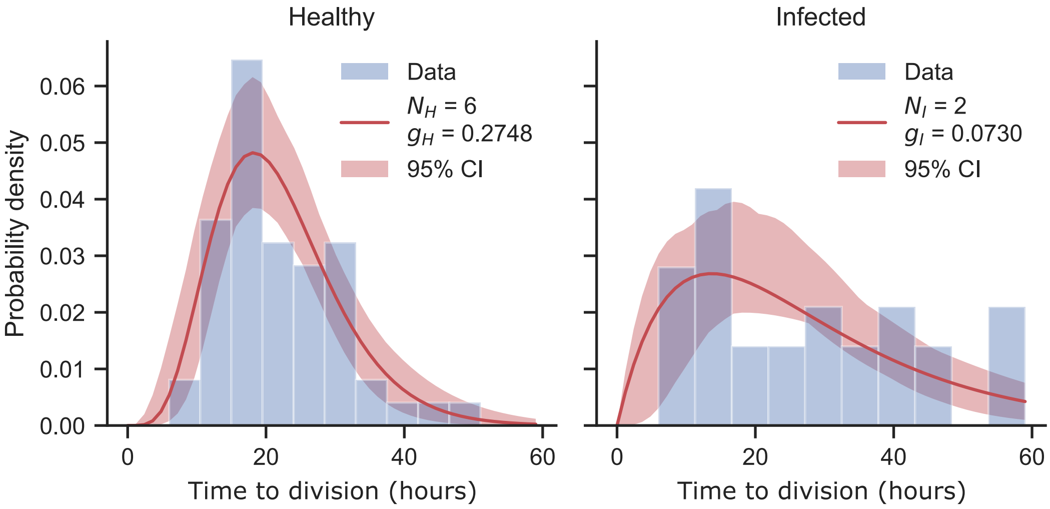

Figure 3.4: Time to division data for both healthy and infected macrophages. Erlang distributions are fitted and 95% confidence intervals are calculated. The plotted confidence interval is obtained by evaluating the maximum and minimum probability for each division time over all parameters within the 95% confidence interval.

3.4.2 Intracellular bacteria growth

Intracellular bacteria are observed to grow exponentially, analogously to their extracellular counterparts. Despite the space within a macrophage being limited and potentially limiting the growth of the bacteria, there is no observation of logistic growth. It is assumed that the bacteria continue to grow until the macrophage can no longer sustain the intracellular load, at which point the macrophage bursts, and the bacteria are deposited back into the extracellular space.

The data available for fitting the intracellular growth rate is repeated experiments measuring the total fluorescence of a single population of intracellular bacteria, each experiment resulting in a time series for that population. In order to calculate the intracellular growth rate, each population's time series is fitted individually with an exponential curve. This curve fitting process results in a MLE as well as an error. By assuming Normally distributed errors, these can be summed in a Gaussian kernel, resulting in the final distribution of the intracellular growth rate. From this, the

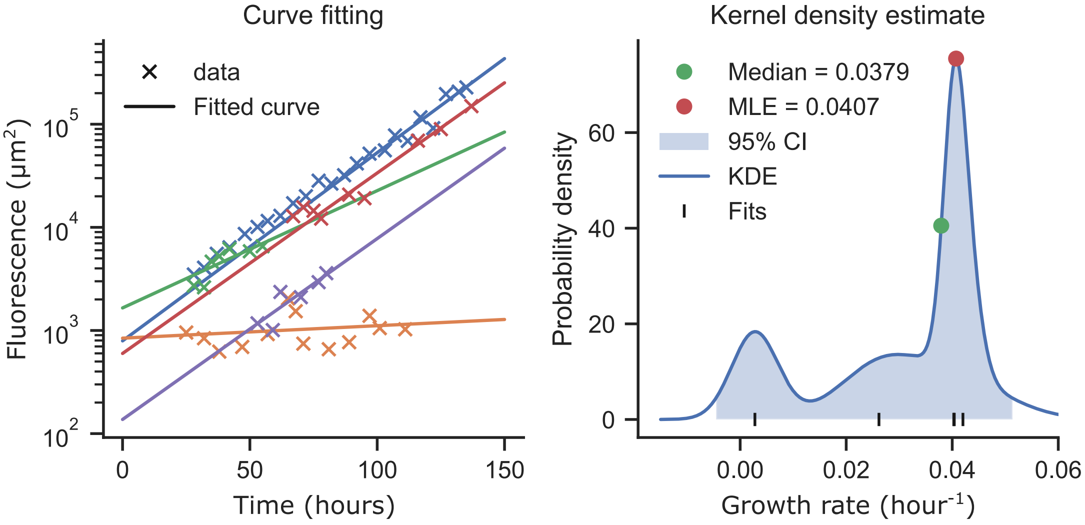

Figure 3.5: Data show exponential growth of intracellular bacteria, to which an exponential growth curve can be fitted using non-linear least squares (**left**). Averaging over the fitted exponential growth rates weighted by their errors results in the kernel density estimate. This distribution is used to find the optimal value for the intracellular growth rate (α_I) (**right**). Note, the orange experiment could be considered an outlier, since the observed growth rate is considerably lower than the others, however, since the data set it small, it will not be removed. The effect of α_I on the outcome of the model will be explored in section 3.5.2.

MLE is used. This is shown in figure 3.5. The orange experiment shown could be considered an outlier, as its fitted growth rate is almost zero, however since the data set is so small it will not be removed. The effect of the intracellular growth rate will be explored in section 3.5.2.

To model the number of intracellular bacteria, macrophages are divided into components M_k^i where k denotes the number of intracellular bacteria. As the bacteria grow, the macrophages move through components $M_k^i \to M_{k+1}^i$ at a constant rate $\alpha_I k$.

Contrasting against the extracellular growth rate shown in figure 2.5, the distribution of intercellular growth rates shown in figure 3.5 has higher variance. The intracellular growth rate also also results in a higher MLE than that of the extracellular growth rate.

3.4.3 Infection

There is no change in this model to how the rate of infection would be calculated, however it is possible that the best fitting number of Erlang stages differs between infected macrophages and healthy macrophages. This complicates the compartment choice a macrophage should move to upon becoming infected, as there may not be a one-to-one mapping. In the interest of simplicity, a healthy macrophage (M_0^i) which becomes infected always moves to compartment M_1^0. It will be demonstrated in section 3.5.2 that the choice of compartment has very little effect on the overall outcome of the system.

3.4.4 Infected macrophage division

The information in the current data set is insufficient to identify how the intracellular bacteria are divided between daughter cells post division, however as the detail in the data improves, this will be possible as seen in chapter 5. For this chapter, as done previously, the model will assume that there is some chance that a macrophage divides to create either one or two infected daughter cells. Since the intracellular load is now being measured explicitly, it is assumed that when an infected macrophage divides into two infected daughter cells, the intracellular load of the parent is divided equally between the two daughters. Since the intracellular load is required to be an integer, if the initial load is an odd integer $2k + 1$, the daughter cells are of equal proportion k and $k + 1$.

3.5 Model results

In this section the underlying ODEs of the model will be solved computationally and the output will be analysed. The final model parameters along with their

Parameter	**Value**	**95% CI**	**Unit**
α_E	0.0290	$(0.0168, 0.0328)$	hour^{-1}
α_I	0.0407	$(-0.0045, 0.0514)$	hour^{-1}
N	50		cells
g_H	0.2748		hour^{-1}
N_H	6		cells
d_H	0.0453	$(0.0332, 0.0601)$	hour^{-1}
g_I	0.0730		hour^{-1}
N_I	2		cells
d_I	0.0375	$(0.0252, 0.0532)$	hour^{-1}
β	0.0056	$(0.0031, 0.0094)$	cell^{-1}hour^{-1}
p	0.3000	$(0.1361, 0.5095)$	scalar

Table 3.2: Model parameters, their MLE values, and 95% confidence intervals. Note confidence intervals for parameters governing Erlang growth of healthy and infected macrophages are shown separately as parameters spaces in table 3.3.

N	g_H min	g_H max	N	g_I min	g_I max
4.0	0.174514	0.192557	2.0	0.054229	0.094324
5.0	0.200576	0.258714	3.0	0.082295	0.140433
6.0	0.240671	0.312843	4.0	0.120386	0.176519
7.0	0.282771	0.360957	5.0	0.174514	0.190552
8.0	0.330886	0.405062			
9.0	0.385014	0.441148			

Table 3.3: (**Left**) 95% confidence intervals for Erlang growth parameters for healthy macrophages. (**Right**) 95% confidence intervals for Erlang growth parameters for infected macrophages.

confidence intervals are shown in tables 3.2 and 3.3. A sensitivity analysis will be performed identifying the dominant factors in the outcome of the model. Finally, some computational experiments will be performed investigating the relationships between key mechanisms.

3.5.1 Model output

The system of ODEs is integrated over 140 hours with all parameters set at their MLEs. Then the parameters are allowed to vary within their 95% confidence intervals and the model is integrated multiple times in order to capture the variation in model output as a result of variation in the parameters. This is shown in figure 3.6. Stochasticity can be a major driving force of the resultant dynamics within a system. It is generally agreed that ODE models should be reserved for modelling large populations, as the stochastic effects of the individual have a smaller effect as the population size grows. Despite this, the ODE model presented here will continue to be studied before studying the effects of stochasticity in later chapters. This is discussed further in section 3.6.

Compared to the previous model, the resultant growth of extracellular bacteria is much slower and more closely resembles the original experiments, while maintaining the maximum intracellular load of $N = 50$ from the literature. An interesting feature is the ratio of intra to extracellular bacteria. In this system, extracellular bacteria are far more prevalent, despite a growth rate of almost half of their intracellular counterparts. Similarly to the previous model in chapter 2, this is due to the ability of bacteria to continue growing intracellularly before being released back in to the extracellular space. In fact the extracellular population can be completely sustained by only intracellular growth.

In order to investigate how much of an effect intracellular growth has, figure 3.7 shows the result of the model with all parameters at their MLEs other than α_E, which is set to zero. Despite only being able to grow intracellularly, the population of extracellular bacteria continues to grow, albeit at a significantly reduced rate:

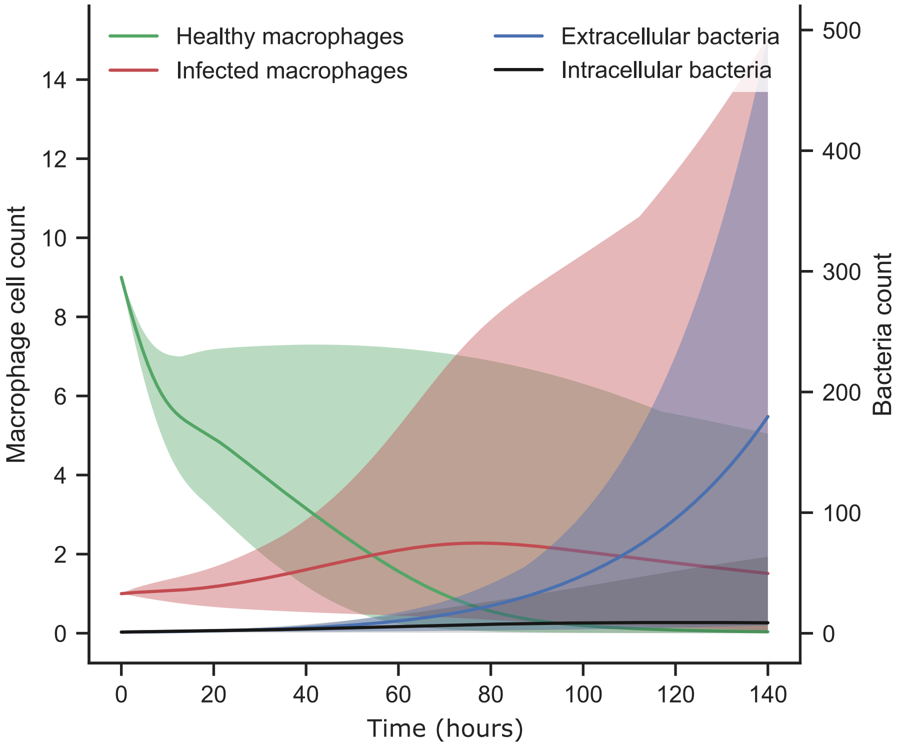

Figure 3.6: Integration of equations 3.1 to 3.6. Solid lines represent parameters set to their MLEs, while the shaded area is created by varying the parameters within their confidence intervals. Parameter sets were created using latin hypercube sampling (LHS) with 2000 sets of parameters chosen.

the final number of extracellular bacteria is an order of magnitude lower than in figure 3.6. This demonstrates that despite the intracellular growth rate being double that of extracellular bacteria, it is the combined ability to grow within and without macrophages that benefits the bacteria.

The mechanism by which intracellular bacteria are released back into the the extracellular space is the death of their host macrophage. There are two routes for this to occur in the model: the natural death of the macrophage which occurs at rate d_I for all macrophages $M^i_{k>0}$; and the rupture of a cell which has reached its carrying capacity which occurs at rate $\alpha_I N$ for macrophages with N intracellular bacteria M^i_N. Naturally, the rate of infected macrophage death (d_I) will affect the total

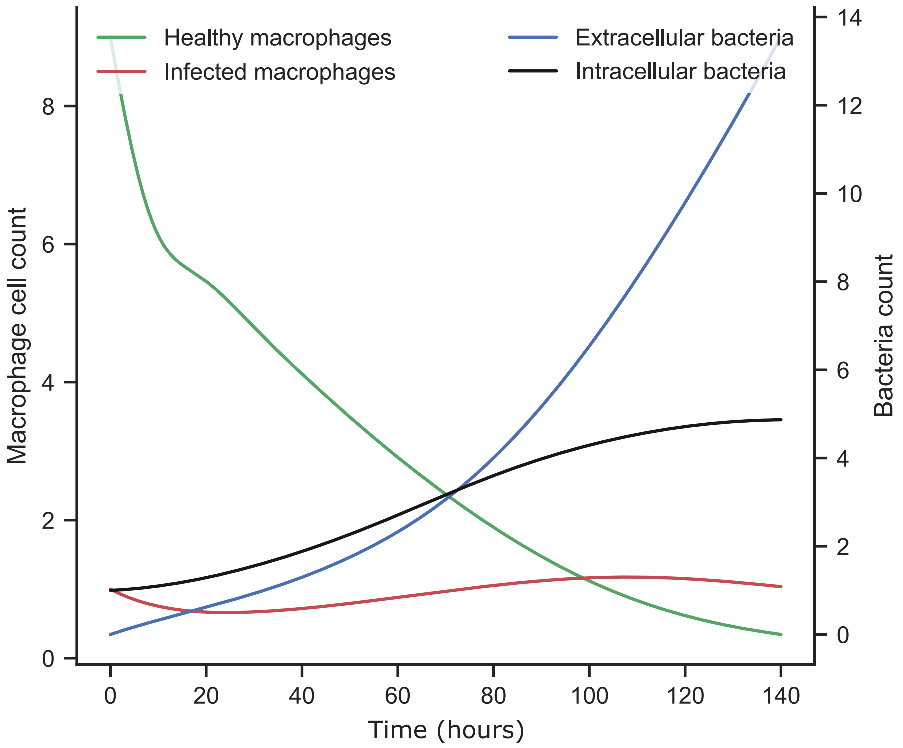

Figure 3.7: Despite the extracellular growth rate being zero, the population of extracellular bacteria can still be maintained through intracellular growth. The resultant population however is an order of magnitude lower than when the bacteria are able to grow extracellularly.

population of bacteria. Since the intracellular growth rate (α_I) is higher than the extracellular growth rate, a lower value of d_I will result in a higher total population of bacteria. This preference for a lower value of d_I is in line with a noted ability of the *Mycobacterium tuberculosis* (Mtb) bacteria described in section 1.7.1. Mtb are able to delay apoptosis of the macrophages when it is preferable for the bacteria to remain intracellular.

3.5.2 Sensitivity

Now that the intracellular bacteria are being explicitly modelled, it is clear from the sensitivity analysis shown in figure 3.8 that the dynamics of both intra and extracellular bacteria are important contributors to the total extracellular bacteria

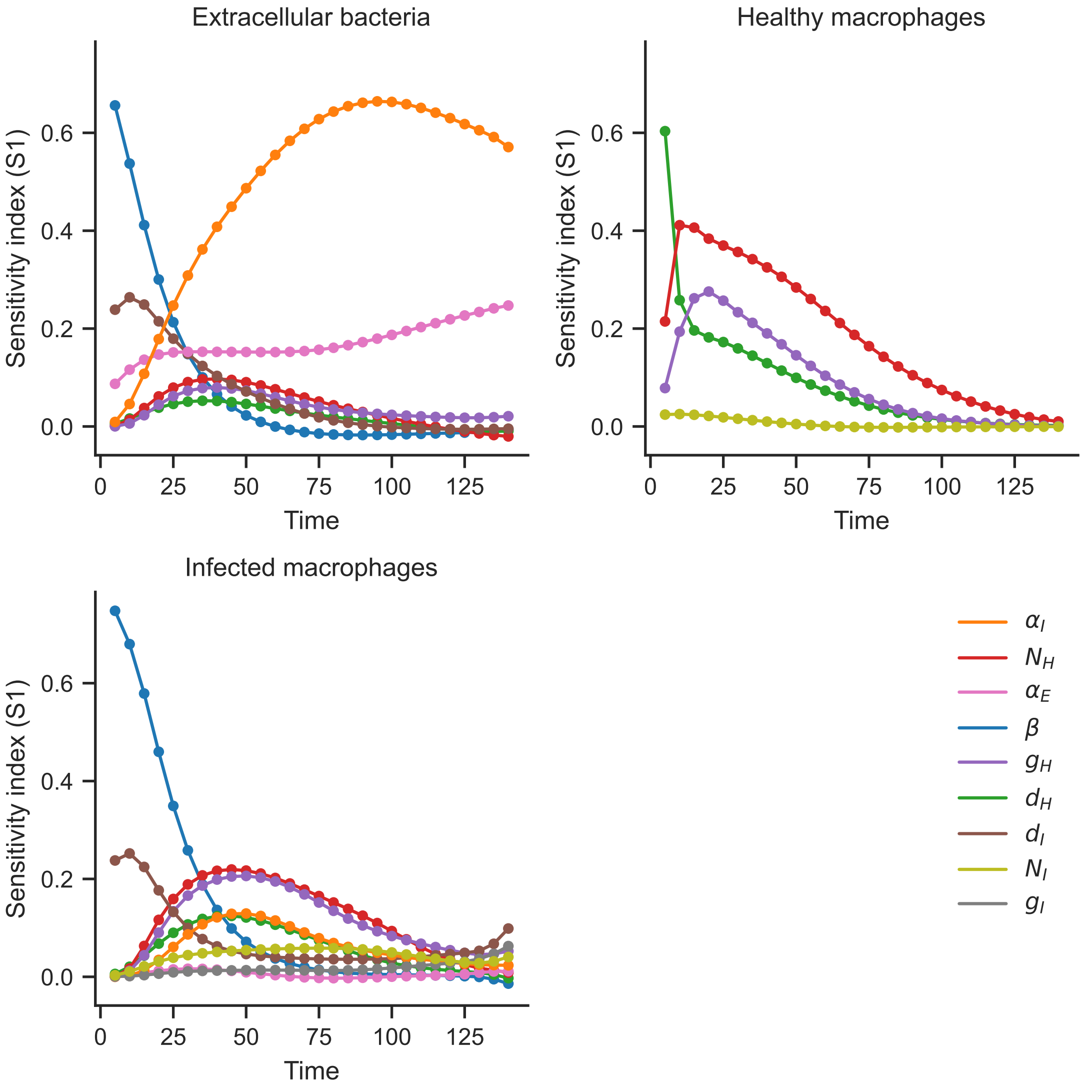

Figure 3.8: Sobol sensitivity analysis measuring how much variance in the model populations can be attributed to variance in the model parameters. Populations are sampled at evenly spaced time points to investigate the sensitivity to parameters throughout the whole time course.

population. As in the previous model, at the start of the simulation, variance in the extracellular bacteria population is dominated by the rate of infection (β). This rapidly drops away however, as the bacteria establish both extra and intracellular populations. As the majority of bacteria are initially intracellular, and the intracellular growth rate (α_I) has larger uncertainty than that of the extracellular growth rate (α_E), for the rest of the 140 simulated hours the dominant contributing parameter to the extracellular bacteria population is the intracellular growth rate followed by the extracellular growth rate. However towards the end of the simulation, once the extracellular population has had time to grow and the number of macrophage hosts is beginning to be depleted, this relationship is trending towards a reversal. One significant difference to the previous model is the reduction in dependence on the death rate of infected macrophages (d_I). While there is some dependence initially, this is quickly dwarfed by the intracellular growth rate. This can be explained by comparing the last two terms of equation 3.1. Initially, the population of infected macrophages is small for all populations of intracellular bacteria, thus the significant difference between these terms will come from comparing d_I and $\alpha_I N$, of which clearly $\alpha_I N$ is much larger.

The dynamics of the healthy macrophages is almost entirely driven by their own birth death process. Parameters such as the the rate of infection β cannot be seen as they did not provide a significant contribution to the variation observed in the macrophages population. The only additional parameter seen belongs to the growth rate of infected macrophages. This is a result of the possibility of infected macrophages dividing to produce one healthy and one infected macrophage.

Initially the major parameters contributing to variation in the infected macrophage population are the rate of infection (β) and the rate of death of infected macrophage (d_I). β drives the rate at which healthy macrophages are becoming infected increasing the infected macrophage population, while d_I does the opposite and brings it back down. The infected macrophage birth death process mirrors that of the healthy macrophages, with the death rate initially having a larger impact, but as the population becomes more dispersed across the Erlang age stages, the growth rate (g_I, N_I) becomes more significant. After the initial stage of the model being dominated by infection, the majority of variance observed in the infection macrophage population is a result of variance in the growth rate of healthy macrophages. At this point in the simulation, the healthy macrophage population is still larger than that of the infected macrophages, thus the availability of healthy macrophages to become infected will be a major driver of the infected macrophage population, compared to its own sustaining growth process.

Adjusting the measure of the model output from the populations over time to the time taken until ${}^{\mathrm{d}E}\!/_{\mathrm{d}t} > 0$ allows the model to identify which mechanisms are involved in delaying the advance of infection. Figure 3.9 shows the resulting sensitivity analysis for this measure. As in the previous chapter, the dominant parameter is the rate of infection (β): quickly internalising the extracellular bacteria will naturally delay any extracellular growth. The next most dominant parameter is d_I: once the bacteria are internalised, a decreased rate of infected macrophage death will result in the bacteria remaining intracellular for longer, thus further delaying the growth of extracellular bacteria. The role of β will be analysed in more detail in section 3.5.5.

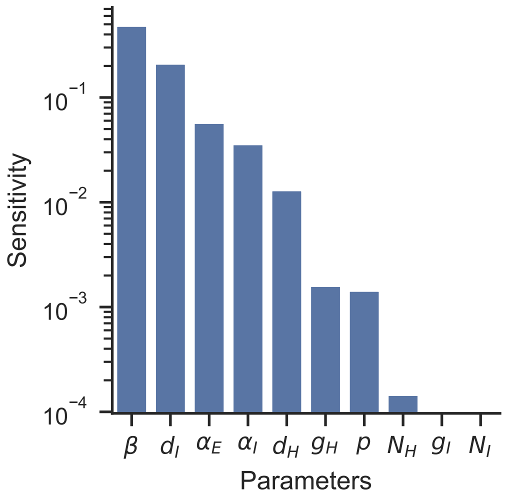

Figure 3.9: Sobol sensitivity analysis bounding the parameters within their 95% confidence intervals. Measured is the fraction of the variance in the value of the time at which $\frac{\mathrm{d}E}{\mathrm{d}t}$ switches sign from negative to positive, which can be attributed to variance in the given parameter.

3.5.3 Identifiability

The additional parameters can once again be analysed for identifiability according to 1.6.7. As in chapter 2, the parameters governing the birth death process are the MLEs of simple distributions, thus are trivially identifiable. The likelihood function for the intracellular growth is shown in figure 3.5, demonstrating a unique maximum and this is identifiable. The other parameters remain fitted as in chapter 2 thus the model is identifiable.

3.5.4 Compartment choice post infection

The model presented in this chapter assumes that when a healthy macrophage (M_0^i) becomes infected, it will transition to the first Erlang component of infected macrophages (M_1^0). Figure 3.10 shows the effect of modifying the model so that when a macrophage becomes infected, it is distributed evenly across all Erlang compartments. Since the birth-death process of healthy and infected macrophages are modelled independently, they do not necessarily have the same number of Erlang stages, thus a simple one-to-one mapping is not possible. The results in figure 3.10 demonstrate the modification to the model results in only a small effect on the overall result, especially on the resultant number of extracellular bacteria, a key measure of the model. There is, however, a small increase in the population of infected macrophages. This is a result of the length of time until division being reduced for newly infected macrophage in the alternative model. Despite this, the base model will continue to be used as it is a simpler model and does not significantly alter the model outcome.

While the assumption that all macrophages move to class M_1^0 after becoming infected only has small impact on the outcome of the model, this modelling hypothesis could be investigated experimentally. Begin with a population of healthy macrophages for which the time since division is known, then infect these macrophages with multiplicity of infection (MOI) 1. By observing the dynamics of the resultant infected macrophages, the effects of the age of the macrophage at the time it was infected could be investigated. If there is no correlation between the age of the healthy macrophage and at time of infection, and the time between infection and subsequent division of the infected macrophage, then it could be concluded the tha assumption of the model is valid.

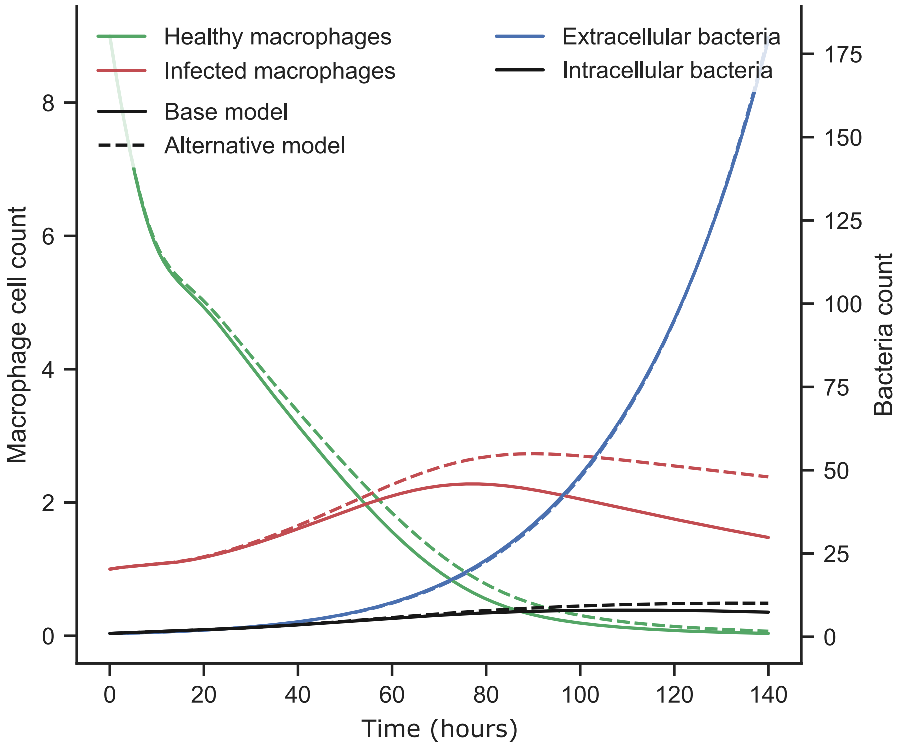

Figure 3.10: The two models simulated are: **a)** the base model: healthy macrophages that become infected always move to the first Erlang compartment of infected macrophages, **b)** the alternative model: healthy macrophages that become infected are equally distributed across Erlang compartments. The change in model has a negligible effect on the extracellular population, but results in a small increase in the population of infected macrophages. In the base model, the time until division for newly infected macrophages is longer than for those in the alternative model, which results in the observed increase in population of infected macrophages.

3.5.5 Predictions

Two computational experiments are used to investigate the following: firstly, the role of the most significant mechanism in delaying exponential growth of extracellular bacteria as shown in section 3.5.2; and secondly, the hypothetical potential of actively removing bacteria from the system.

The role of phagocytosis

Since this model has no way in which to reduce the bacteria population, a method of control may be to delay the extracellular growth, allowing the host's adaptive immune response time to respond. The rate of phagocytosis is the macrophages' only weapon in this system, so the time until the rate of extracellular growth becomes positive is measured against the rate of internalisation of bacteria. In a similar way to the previous version of the model, if the number of healthy macrophages is less than α_E/β, then the rate of change of extracellular bacteria will always be positive. This explains the initially flat part of the curve shown in figure 3.11. Once the scales are tipped in favour of the macrophages there is a rapid increase in the amount of time it takes for the extracellular bacteria to recover, however an interesting feature of this experiment is that there is a point at which increasing the infection rate, β, further has a negative effect. The traditional view of macrophages is that they play a dual role [31] in protecting against Mtb infection: by internalising the bacteria they inhibit their spread, however the bacteria find themselves a safe haven to grow in their host macrophages. The optimal rate of phagocytosis here demonstrates the balance of inhibition and protection.

At their MLE values, the intracellular bacteria have a higher growth rate than the extracellular bacteria, so by continuing to increase β, the bacteria are forced to grow predominantly intracellularly. In tandem, the increase to the rate of infection increases the rate at which the healthy macrophage population is depleted. The resulting situation is a large number of infected macrophages, each with a large number of intracellular bacteria. Once these bacteria begin to be released back into

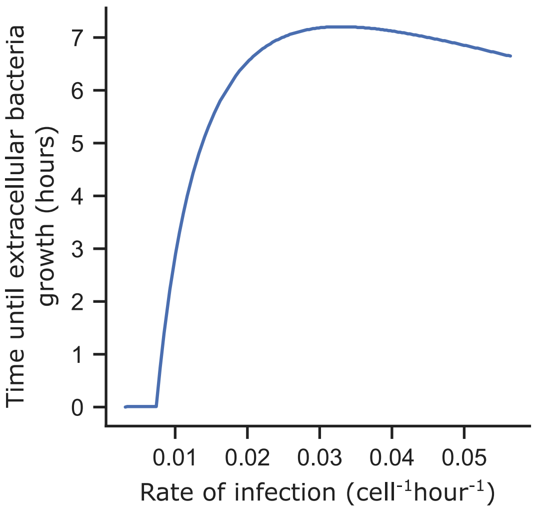

Figure 3.11: Evaluating the time taken for $\frac{\mathrm{d}E}{\mathrm{d}t}$ to change sign from negative to positive for a range of infection rates (β).

the extracellular space, the extracellular bacteria population will have a sudden increase. Further increasing β compounds this effect. The sooner extracellular bacteria become intracellular, the sooner they are released back to being extracellular with a depleted healthy macrophage population, thus the sooner a growing extracellular population is established.

There is a natural delay between the initial infection of a host and the onset of the adaptive immune response. Increasing the time before exponential extracellular growth will result in a smaller extracellular population at the time of the arrival of the adaptive immune response. Identifying an optimal value for the value of β in maximising the delay for extracellular growth speaks to the duality of the role of the macrophage in protecting the host and enabling the protected intracellular growth of Mtb.

Critical kill rate vs intracellular growth rate

The explicit process of successfully killing bacteria and removing them from the system is not directly observable in the biological experiments. This is due to the bacteria being observed at the macro-population level, rather than distinct individuals. Further, this killing of bacteria is usually the realm of macrophages that have been activated through the interaction with T cells and the adaptive immune response. Activated macrophages can kill phagocytosed bacteria by transferring them to the destructive environment of the lysosome [104]. While this is not thought to be an ability possessed by the non-activated macrophages, there is some evidence [19, 24] that they can kill Mtb resulting in early clearance. This process is explored in relation to the intracellular growth rate as a proxy for bacteria virulence.

Figure 3.12 shows a contour plot measuring the change in the number of extracellular population after simulating the system for 140 hours. This change is a result of ranging the parameters controlling intracellular growth (α_I) and the probability that a macrophage kills an extracellular bacteria (p_{kill}). When this occurs, the healthy macrophage is assumed to remain healthy, although in reality there would be a delay immediately after killing a bacteria before a macrophage would be able to perform the same action. The mechanism of killing a bacteria is modelled by including a parameter p_{kill} in equations 3.2, 3.3, and 3.4 as follows:

$$\frac{\mathrm{d}M_0^0}{\mathrm{d}t} = \dots - \underline{(1 - p_{\text{kill}})}\beta E M_0^0 + \dots \tag{3.7}$$

$$\frac{\mathrm{d}M_0^i}{\mathrm{d}t} = \dots - \underline{(1 - p_{\text{kill}})}\beta E M_0^i + \dots \tag{3.8}$$

$$\frac{\mathrm{d}M_1^0}{\mathrm{d}t} = \dots + \underline{(1 - p_{\text{kill}})}\beta E \sum_{i=0}^{N_H} M_0^i + \dots . \tag{3.9}$$

The rate of uptake of extracellular bacteria by healthy macrophages remains the same at $\beta E \sum_{i=0}^{N_H} M_0^i$, however the rate at which healthy macrophages become infected after phagocytosis is reduced by a constant $1 - p_{\text{kill}}$ representing the fraction of interactions that result in the bacteria being killed.

The colour bar in figure 3.12 represents the ratio of the number of extracellular bacteria at the end of the simulation compared to the base model: $p_{\text{kill}} = 0$, $\alpha_I = 0.0407$. The simulation is run for 140 hours, replicating the amount of time the experiments were run for. Note the scale is a composite of two linear scales, centered at 1. The wider central contour level shown in the main figure represents the set of parameters for which the outcome of the model does not change. As the probability for a macrophage to kill a bacteria increases from zero, this contour remains reasonably vertical. It is only once p_{kill} exceeds approximately 0.8 that a dramatic change occurs, forcing the infection to remain lower despite increasing intracellular growth rates.

Naturally, as the rate of intracellular growth decreases towards zero, macrophages are far more able to keep the infection under control, even with limited ability to kill the bacteria. However, as the rate of intracellular growth increases, the infection very quickly becomes unmanageable, even with a high probability for a macrophage to kill a bacteria. This is a result of the population of healthy macrophages being reduced before the bacteria population is fully eliminated, allowing the few remaining intracellular bacteria to grow exponentially before bursting their host macrophages and returning to the extracellular space.

This demonstrates the significant advantage Mtb has against the innate alveolar macrophage. However environmental stochasticity that results in a lower intracellular growth rate, and a higher chance for innate macrophages to kill bacteria, makes it possible for the host to contain the bacteria before the adaptive immune response kicks in. In order to investigate the variation in outcome as a result of stochasticity in the system, a stochastic framework is required. This will be developed in the next chapter.

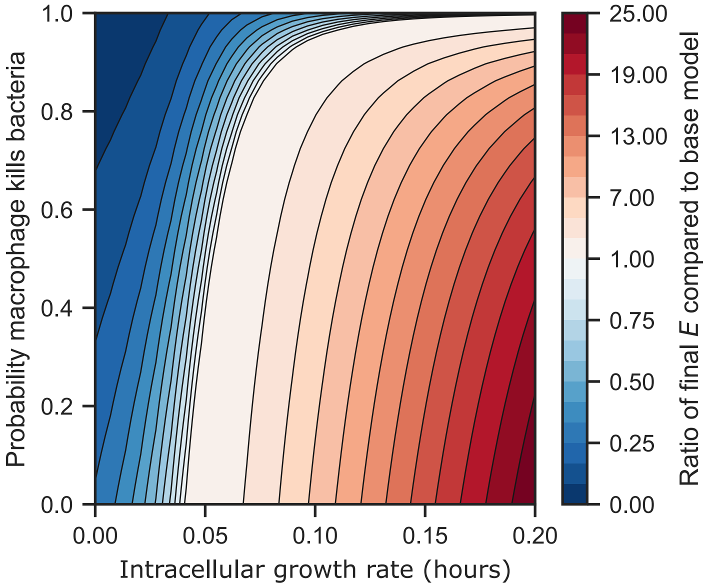

Figure 3.12: By ranging the parameters controlling intracellular growth (α_I) and the probability a macrophage kills an extracellular bacteria, the effect on the resultant population of extracellular bacteria can be measured. Shown is the ratio of bacteria in the alternative model compared to the base model. Note that the colour bar scale is a composite of two linear scales, centered at 1. The value 1 represents that the outcome of the model is the same as the base model. < 1 represents an improvement, and > 1 represents a deterioration in outcome.

3.6 Conclusions

In this chapter the ODE model has been extended to explicitly include the intracellular bacteria and an improved model for the growth and division of macrophages. Through sensitivity analysis the key mechanisms driving extracellular proliferation were identified as bacteria growth rates, and the infection and death of macrophages. By demonstrating the importance of explicitly modelling the intracellular bacteria, feedback to the experimentalist prompted additional data to be collected, explicitly measuring the intracellular growth.

Through simulation of the model equations, increasing the rate at which macrophages internalise bacteria was found to have a protective effect up to a maximum, after which increasing β further reduced the effectiveness. The measure of effectiveness was defined as the time until the rate of change of extracellular population is greater than zero. To begin with, the value of β is too low to counter the extracellular growth. After passing a critical value, the macrophages are able to internalise bacteria quickly enough that the extracellular growth of bacteria is delayed. Increasing the value of β further still continues to add to this delay up to a maximum. After this point, the population of healthy macrophages is depleted too quickly, resulting in extracellular growth once more, thus reducing the delay.

In order to investigate the possibility of bacteria control by the macrophages, a theoretical mechanism was added to the system whereby macrophages have some probability that instead of becoming infected, they are able to kill the bacteria. Through computational analysis, it was shown that a small increase in the virulence of the bacteria, here assumed to correlate with intracellular growth rate, requires a significant increase in the probability for a macrophage to kill in order for the response of the system to remain unchanged. This demonstrates the advantage the bacteria have over the macrophages. The significantly higher growth rate of the bacteria overwhelms the macrophages in the absence of recruitment.

Despite being a more realistic birth death process, some failings of the model are acknowledged. By considering the disease free system, there is no non-zero steady state for healthy macrophages.

$$\frac{\mathrm{d}M_0^1}{\mathrm{d}t} = -(g_H + d_H)M_0^1 + g_H M_0^{N_H} \tag{3.10}$$

$$\frac{\mathrm{d}M_0^i}{\mathrm{d}t} = g_H M_0^{i-1} - (g_H + d_H)M_0^i \tag{3.11}$$

$$i \in \{2, \dots, N_H\}$$

They will either grow exponentially or decay to zero according to the eigenvalue $\lambda = \sqrt[N_H]{2}g_h - g_h - d_h$. The work presented here is focused on the very early dynamics of the interactions between Mtb and alveolar macrophages. As discussed in 1.7.1, the long term macrophage population in the lungs is maintained primarily through recruitment from the blood stream. A simple way to achieve a non-zero steady state for the macrophage population in the model presented here would be to add a constant to the rate of change of the macrophage population. This would represent recruitment, however since the data being modelled is an in vitro experiment, recruitment is not observed and thus is left out of the model. Hence the absence of a non-zero steady state for healthy macrophages does not pose a problem.

Further, the deterministic framework of the model is often considered inappropriate for modelling small populations, such as the macrophage population which rarely exceeds ten in the modelled scenario. In populations of this size, the effects of stochasticity should not be ignored as they are significant in comparison to the unit nature of the agents being modelled. Additionally, by modelling in integer units, the possibility for population extinction arises, whereas in ODE models, such as the one presented in this chapter, a population can always bounce back no matter how small it gets, a phenomenon known as the Atto-Fox problem [44]. To address these shortcomings, in the next chapter the model mechanisms will be restructured as a stochastic agent-based model.

Chapter 4

Model 3: Stochastic Agent Based Model

In this chapter the models which were developed in previous chapters are adapted to a stochastic framework. Since the populations of both bacteria and macrophages are small in the system being modelled, stochastic variation in the event times will have a significant effect on the outcome of the model. Additionally, this will allow the populations to be modelled as integers rather than continuous values, which presents the opportunity for populations to become extinct during the simulation.

4.1 From continuous to discrete populations

In this chapter, many of the mechanisms presented in previous chapters will continue to be implemented, however the model being presented is now a discrete population model being updated stochastically. To describe this process the case study of the classic infectious disease SIR model will be used. The model describes an infection spreading through a population that is initially susceptible S, becomes infected I, then eventually recovers and is immune R.

$$\tfrac{\mathrm{d}S}{\mathrm{d}t} = -\beta\frac{SI}{N} \tag{4.1}$$

$$\tfrac{\mathrm{d}I}{\mathrm{d}t} = \beta\frac{SI}{N} - \gamma I \tag{4.2}$$

$$\tfrac{\mathrm{d}R}{\mathrm{d}t} = \gamma I \tag{4.3}$$

where $N = S + I + R$. A commonly investigated output of this model is the basic reproduction number R_0 which describes the average number of secondary cases in an entirely susceptible population. It can be found by analysing the sign of *odIt* at the start of the epidemic, when the entire population is susceptible: $S \approx N$ thus $\frac{\mathrm{d}I}{\mathrm{d}t} \approx (\beta - \gamma)\, I$, hence if $R_0 := {}^{\beta}\!/_{\gamma} > 1$ then disease will be able to propagate through the population.

The continuos and deterministic model above can be described by a continuous time Markov chain (CTMC) as follows. The state space for the model is $(S, I, R) \in \{0, \ldots, N\}^3$. The transition matrix is simple to construct, however due to the high dimension of the state space, it is not useful to write down. Instead it can be defined by a series of rules:

- For all states with $S > 0$, and $I > 0$ there is a transition from (S, I, R) to $(S-1, I+1, R)$ with rate ${}^{\beta SI}\!/_{N}$,
- For all states with $I > 0$, there is a transition from (S, I, R) to $(S, I-1, R+1$ with rate γI,
- All other states are absorbing states and thus no more changes occur.

Once again the basic reproduction number can be defined R_0, and it can be demonstrated to have the same value as in the deterministic model $R_0 = {}^{\beta}\!/_{\gamma}$. Consider the population is entirely susceptible, that is $S \approx N$ and $I > 0$. There are therefore two possible events that can occur: an infection event with rate ${}^{\beta SI}\!/_{N} \approx \beta I$, and a recovery event with rate γI. R_0 is defined as the average number of secondary cases produced from the average infected individual $(I = 1)$ in a totally susceptible population $(S \approx N)$. Recovery occurs at rate γ so the average infected individual will recover after ${}^{1}\!/_{\gamma}$. Infections occur at rate approximately β. It is assumed that a single infection doesn't significantly reduce the population of susceptibles, thus in time ${}^{1}\!/_{\gamma}$, the infected individual will infect on average ${}^{\beta}\!/_{\gamma} = R_0$ susceptibles.

Finally it can be demonstrated that the continuous deterministic model is the mean field approximation of the discrete stochastic model, that is the mean solution to the discrete model approaches the continuous model as the number of samples approaches infinity. Consider a two state system S with states $S = A$ and $S = B$.

The transition from state A to B occurs at rate δ. The time for a transition event to occur is distributed by the exponential distribution $\mathrm{Exp}(\delta)$, thus the probability that the system is in state A at time t has the same distribution. This can be written as $A(t) = \mathbb{P}[S = A|t] = \delta e^{-\delta t}$. Therefore $\frac{\mathrm{d}A}{\mathrm{d}t} = -\delta A$. Thus the mean field of this system is the differential equation $\frac{\mathrm{d}A}{\mathrm{d}t} = -\delta A$. This logic can be applied to each state of the SIR model, resulting a differential equation for each state. Since the deterministic model considers populations, these differential equations can be summed to result in the same set of equations. The final part of the proof is omitted as it is simple but extensive algebra.

4.2 Data

The data used to parametrise the model presented in this chapter is an extension of the previous data set. In addition to simply being a larger data set, both in terms of the number of experiments performed, and the total duration of the experiment (increased from 140 hours to 200 hours), particular attention was paid to the intracellular growth rate and the infection times of macrophages, both measures that had very few data points in the previous iteration. Table 4.1 shows some simple descriptive statistics of the data relating to the bacteria growth rates. There is a significantly increased quantity of data relating to the intracellular growth rates, while the extracellular growth data remains limited. This will be corrected in chapter 5. The range of values shows the larger rate of growth observed for the extracellular bacteria as compared to the intracellular bacteria.

Table 4.2 shows the additional data relating to the macrophages. Particular attention has been paid to the time to infection, and there is now a large number of data points. There are examples of wait times in the data that extend to the entire duration of the experiment. To be classed as an infection event, the macrophage must be observed to be healthy at least once before being observed to be infected, meaning the minimum possible infection time is two hours (due to the sample rate being once every two hours). Similarly if a healthy macrophage survives from the

Bacteria	Intracellular	Extracellular
Experiment count	131.0000	9.0000
Average starting fluorescence	3.3600	8.2027
Average fluorescence	10.6452	44.5201
Average ending fluorescence	28.9964	126.8290
Average number of data points per experiment	41.5649	36.1111

Table 4.1: Summary of the data relating to intra- and extracellular bacteria growth. There is significantly more data relating to the growth of intracellular bacteria, however the central tendency of the extracellular bacteria suggests a higher growth rate over the intracellular.

beginning of the experiment only to become infected in the last time period, it will have the maximum infection time of 190 hours. This is what is observed in the data, showing the high variance in the macrophage data. Similarly high variances are observed for the time to death and the time to division for infected macrophages, however these variances are slightly reduced for healthy macrophages. How the additional data affects each parameter directly will be discussed at the beginning of the relevant sections throughout this chapter.

Infection status	Healthy				
	Count	Min	Mean	Max	Std
Time to division	108.0000	2.0000	64.3889	180.0000	38.7626
Time to death	52.0000	4.0000	65.8077	184.0000	49.1740
Infection status	**Infected**				
	Count	Min	Mean	Max	Std
Time to division	25.0000	10.0000	71.2800	182.0000	49.5307
Time to infection	220.0000	2.0000	80.0091	190.0000	54.1914
Time to death	206.0000	0.0000	95.1456	194.0000	52.7545

Table 4.2: Summary of the data relating to the macrophages. Macrophages exhibit very high variance for all three measures shown as well as ranges that cover almost the full timespan of the experiments.

4.3 Extracellular growth

As a result of identifying the need for extracellular growth in the absence of macrophage interaction, experiments tracking this are included in this data set. This results in a good picture of how the bacteria grow under normal conditions. For each experiment, the data records the time and the area of fluorescence. As time progresses this area grows exponentially and an exponential curve is fitted to this to obtain the exponential growth rate of extracellular bacteria. Growth curves are fitted using non-linear least squares regression. Some experiments result in smaller errors and therefore produce a more confident estimate of the growth rate. Assuming the errors are distributed normally, these can be used to generate a Gaussian kernel densitiy estimate (KDE) for the distribution of extracellular growth rate α_E. These KDEs can be summed to provide the total distribution of growth rates, from which the maximum likelihood estimate (MLE) can be extracted as shown in figure 4.1.

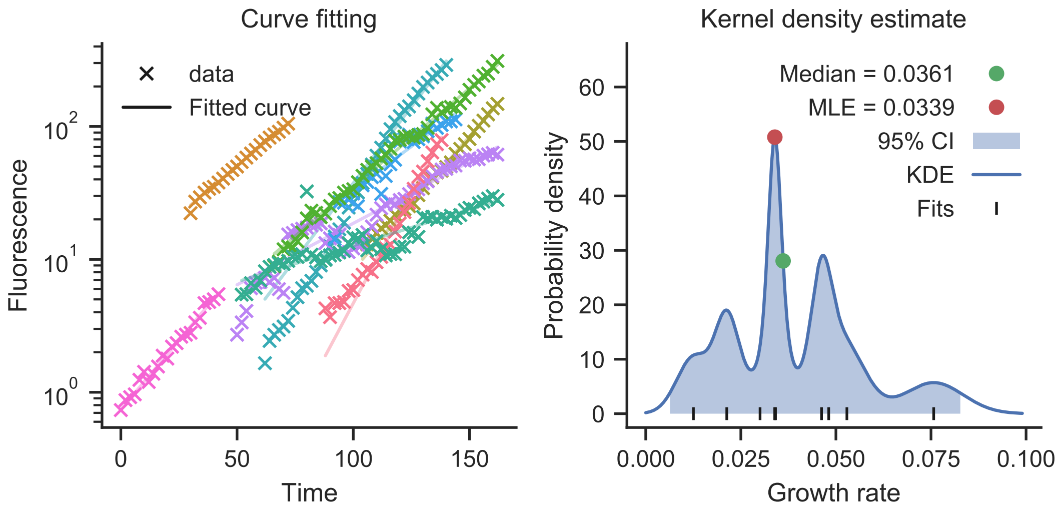

Figure 4.1: Data comprises of individual experiments which track the level of fluorescence from extracellular bacteria over time. Exponential growth curves are fitted (**left**) to the distinct time series data. Each fitted growth rate has an associated error. Taking the fitted value as the mean, and the error as the variance on the mean, a Gaussian KDE can be generated (**right**). The maximum of the KDE is taken as the MLE for the growth rate.

4.4 Intracellular growth

Similarly to fitting the extracellular growth rate, the intracellular growth rates are fitted and their errors are estimated. Intracellular death of the bacteria is not considered for two reasons. Firstly, the innate alveolar macrophages being considered in this study lack the ability to kill the intracellular *Mycobacterium tuberculosis* (Mtb). Secondly, the available data is a one-dimensional time series of the total area of fluorescence, to which an initial load and an exponential growth parameter are fitted. The fitted exponential growth parameter is the resultant difference between bacteria growth and bacteria death. Since bacteria death is not explicitly measured, nor is the relationship between bacteria growth and bacteria death known, only the difference can be established and will be known as the intracellular growth rate α_I.

By taking the fitted value as the mean, and the error of the fitted values as a measure

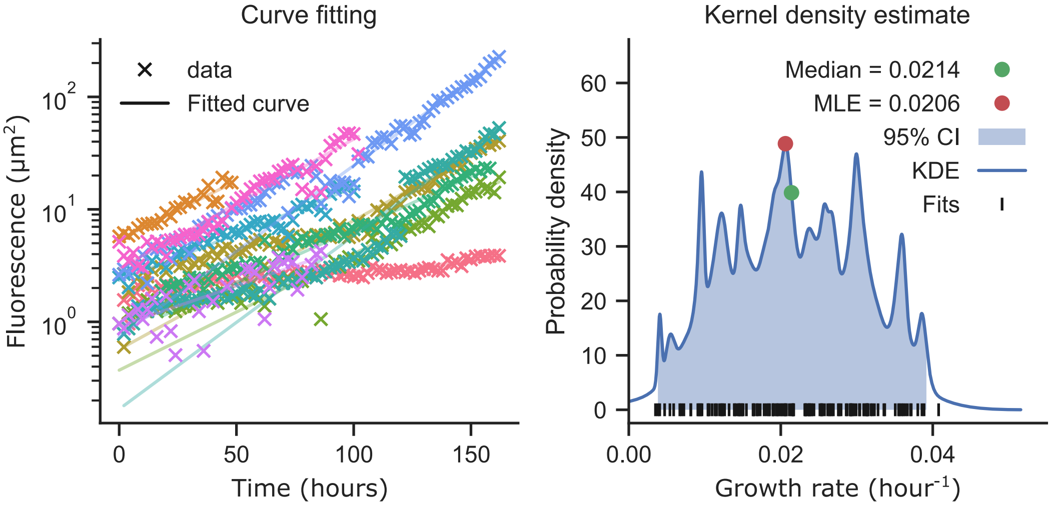

Figure 4.2: (**left**) Exponential growth curves fitted to time series data of fluorescence from intracellular bacteria. (**right**) Guassian KDE generated from fitted intracellular growth rates and their associated variances. The maximum of the KDE is taken as the MLE for the intracellular growth rate.

of the variance, a distribution for each fitted growth rate can be estimated. These distributions are summed over all experiments to find a kernel estimate of the total intracellular growth rate distribution. Then the maximum is found for the maximum likelihood estimate as shown in figure 4.2. The high variance in the distribution of fitted intracellular growth rate α_I speaks of the heterogeneity of the intracellular growth rate within a homogeneous population of Mtb.

4.5 Infection rate

Over the range of repetitions of the experiment, 221 macrophages go through an infection event. For these events, the time of birth of the macrophage, and the time of infection are recorded. The likelihood function is defined similarly to the previous iterations of the model, however here the probability of observing an infection event is conditioned on the event happening within the window of the experiment (approximately 200 hours). The probability of observing a cell that is born at time

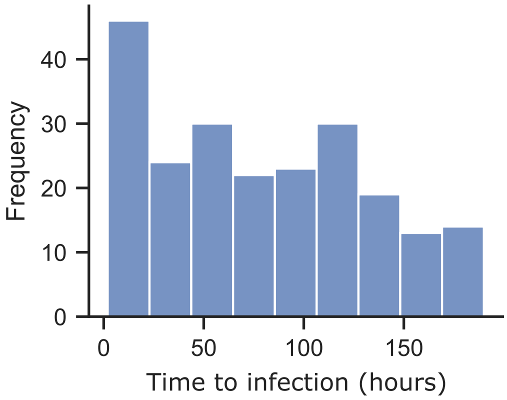

Figure 4.3: Raw data of time from birth to infection for macrophages that become infected during the window of experimentation.

t_0 and becomes infected at time t_1 given that the cell becomes infected during the window of the experiment is given by

$$\mathbb{P}\left(\text{infected by } t_1 \mid \text{infected by } T \cap \text{not infected by } t_1 - 8\right) \tag{4.4}$$

$$= \mathrm{P}\left(N = 0 \text{ in } (t_0, t_1 - 8) \ \cap \ N \neq 0 \text{ in } (t_1 - 8, t_1) \mid N \neq 0 \text{ in } (t_0, T)\right) \tag{4.5}$$

$$= \frac{\exp\left(-\int_{t_0}^{t_1-8} \beta(s)\,\mathrm{d}s\right)\left[1 - \exp\left(-\int_{t_1-8}^{t_1} \beta(s)\,\mathrm{d}s\right)\right]}{\left[1 - \exp\left(-\int_{t_0}^{T} \beta(s)\,\mathrm{d}s\right)\right]}. \tag{4.6}$$

where $\beta(s) = \beta E_0 \exp\left(\alpha_E s\right)$, N is the number of infection events and T is the total length of the experiment. “Not infected by $t_1 - 8$” is used as the data is now captured in eight hour intervals and the model remains using hours for its unit of time. Figure 4.3 shows the distribution of time to infection for healthy macrophages. Since the range of times to infection is significantly larger than in the previous chapters, the estimate for the rate of infection β will be pushed lower. This is a result of the homogeneous mixing between all bacteria and the healthy macrophages. The maximum likelihood estimation of β is $\beta = 2.63 \times 10^{-4}$ as shown in figure 4.4. Note that the increased range in time to infection is a result of the increase in the total duration of the experiments. An observation that prompted conditioning the likelihood function on observing the event within the window of the experiment.

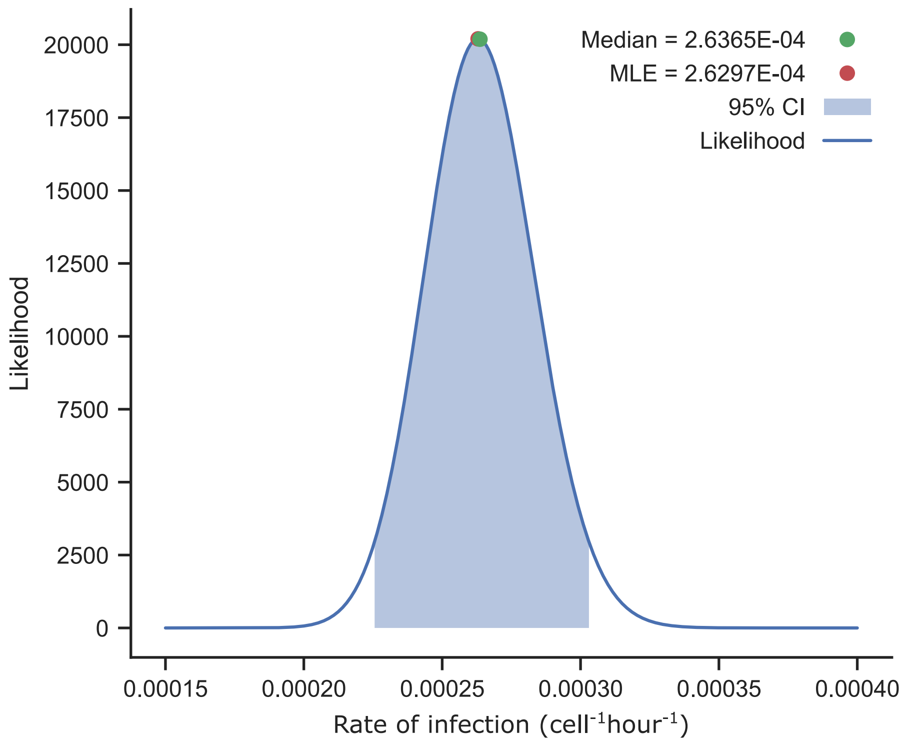

Figure 4.4: Likelihood function for the rate of infection β calculated using equation 4.6.

4.6 Macrophage birth death process

In the previous chapter, an Erlang distribution was fitted to the time of division data and an exponential death rate was fitted to the time of death for both healthy and infected macrophages. The resultant joint distribution modelling the birth death process of a single macrophage does not accurately capture the observed distributions in the data[1] as shown in figure 4.5. This is due to a lack of consideration for the conditional probability of a macrophage dying given that it does not divide. Hence the model being fitted is not a true representation of the model being simulated. To resolve this, the same heuristic model is fitted, however this time the birth and death rates are fitted at the same time using a CTMC.

The CTMC is simply defined by a fixed number (N_g) of states representing the

---

[1] While figure 4.5 is presenting the model from chapter 3, it is being compared against the extended data set used in this chapter.

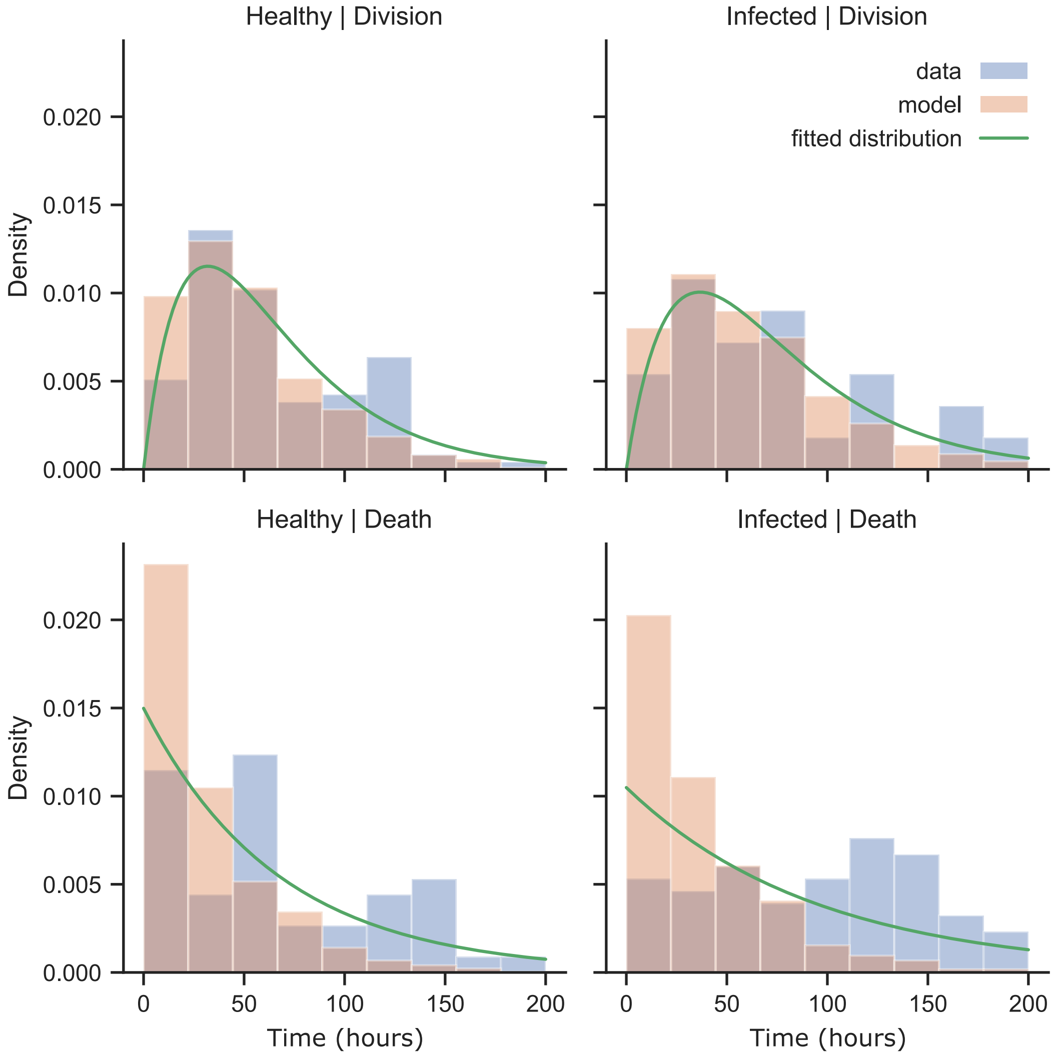

Figure 4.5: An Erlang distribution is fitted to division times and an exponential distribution to death times. Running a stochastic simulation to compare the distribution of birth and death times of the model to the data shows that the model fails to accurately capture death times observed. This is due to the fitting process not considering the dependency between the two mechanisms.

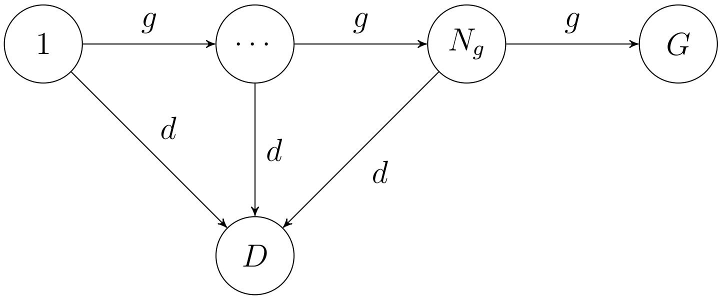

Figure 4.6: A network representation of the CTMC governing the birth death process of a single generation of macrophages. Growth is assumed to follow an Erlang process while the death rate is constant across all stages of life. Macrophages grow by moving through the N_g stages at a fixed rate g and finally transitioning to state G. A cell that reaches state G will divide removing itself from the system to be replaced by two new cells in state 1. Macrophages die by moving from any stage of life to state D. A cell that enters state D is removed from the system and represents a dead cell.

Erlang growth stages, then two additional states, the first representing a macrophage that has divided (G), the second representing a macrophage that has died (D). Macrophages grow by moving through the N_g stages at a fixed rate g and finally transitioning to state G. Macrophages die by moving from any stage of life to state D. This processes is demonstrated in figure 4.6.

To find the optimal values for g, d and N_g, the wait times taken for cells to either divide or die are extracted from the data. The probability of entering state G or D in that time can then be calculated based on the CTMC. Given a healthy macrophage was observed alive at time $T_D - 1$ but dead at time T_D, it can be concluded that it died some time between $T_D - 1$ and T_D. The likelihood of observing this with parameters N_g, g and d is

$$\begin{aligned}&\mathbb{P}\left(X(T_D) = D | X(T_D - 1) \in \{1, \dots, N\}\right) \\ &= \sum_{i=1}^{N_g} \mathbb{P}\left(X(T_D) = D | X(T_D - 1) = i\right)\end{aligned}$$

where

$$\begin{aligned}&\mathbb{P}\left(X(T_D) = D | X(T_D - 1) = i\right)\\&= \left[Kv_i\right]_D \times \left[K^{T_D-1}v_1\right]_i\end{aligned}$$

Here $X(t)$ is the state a macrophage is in at time t, v_i is the vector of zeros with a 1 in the ith place, and $[v]_j$ denotes the jth element of the vector v, and K is the transition rate matrix for the CTMC and is of the form

$$\begin{pmatrix} -g-d & g & 0 & 0 & d \\ 0 & -g-d & g & 0 & d \\ 0 & 0 & -g-d & g & d \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The dimensions of the matrix K are dependent on the number of Erlang stages N_g. $[Kv_i]_D$ represents the likelihood of a macrophage transitioning from state i to state D within one unit of time (1 hour). $\left[K^{T_D-1}v_1\right]_i$ represents the likelihood that a macrophage began in stage 1 at time zero and after $T_D - 1$ hours the macrophage is still alive and in stage i. Summing this term over all values of i would represent the likelihood that the macrophage has neither divided nor died after $T_D - 1$ hours. The equivalent equation can be written down for the macrophages that divide and thus a likelihood function for the data given a set of parameters is defined and can be maximised.

This construction still assumes an Erlang style growth and a fixed exponential death time. The results of the fitting process are shown in figure 4.7. The model now better captures the data and is able to demonstrate the longer death times, it is clear that an exponential distribution is not appropriate, since the model is still underestimating the death times.

To improve further, the model of time to death is extended to also be an Erlang style distribution. Equation 4.7 shows the structure of the transition rate matrix used for this extension while figure 4.8 shows the schematic of the model. As previously,

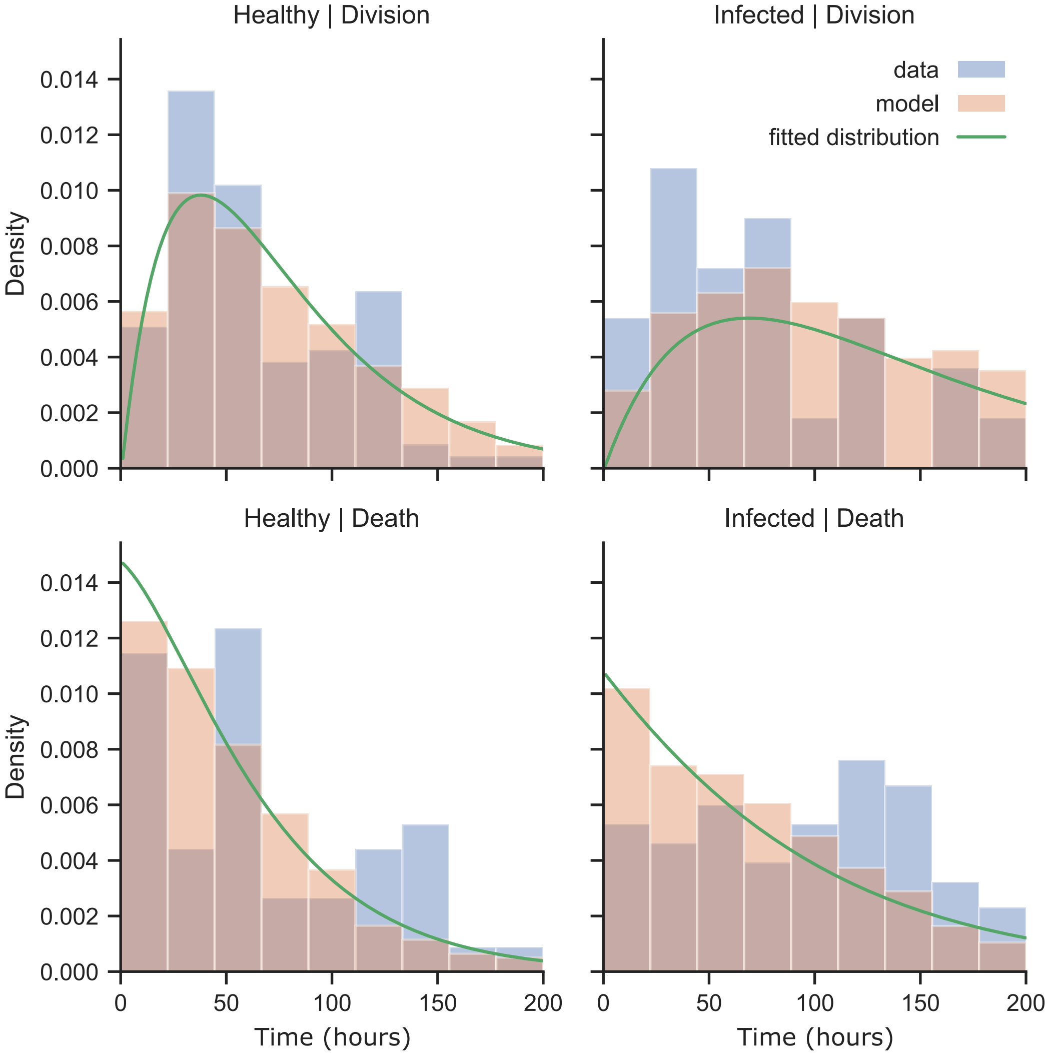

Figure 4.7: Using a CTMC to capture the joint distribution of Erlang growth and exponential death still fails to fully capture the distribution of death times.

macrophages move through the different stages of life at fixed rates g and d. When a macrophage reaches stage G it is considered to divide, therefore the macrophages is removed from the system and replaced by two new macrophages, both in state $(1, 1)$. If the macrophage was infected at the time of division, the intracellular bacteria are split between the two daughter cells according to the model described in section 4.7. When the macrophage reaches stage D it is considered to have died and is simple removed from a system. If the macrophage was infected at the time of death, then the intracellular bacteria are released back into the extracellular space. I.e. the number of intracellular bacteria is added to the total population of extracellular bacteria.

$$K = \left(\begin{array}{c|cccccc} & 1,1 & 1,2 & 2,1 & 2,2 & G & D \\ \hline 1,1 & -g-d & g & d & 0 & 0 & 0 \\ 1,2 & 0 & -g-d & 0 & d & g & 0 \\ 2,1 & 0 & 0 & -g-d & g & 0 & d \\ 2,2 & 0 & 0 & 0 & -g-d & g & d \\ G & 0 & 0 & 0 & 0 & 0 & 0 \\ D & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \tag{4.7}$$

Figure 4.9 shows the final model of the birth death process and how it compares to the data. This is a good fit and the dual Erlang model will be used to model the birth-death cycle of macrophages. Figure 4.10 shows the 95% confidence region for the birth death process.

The parameters d and N_d for infected macrophages are fitted independently of the intracellular load. It was hypothesised that the intracellular load has an effect on the rate of death of the infected macrophages, however the data were not sufficient to demonstrate a relationship between the intracellular load and the rate of death of the infected macrophage. Two models of a load dependent death rate were investigated, but were shown to be unidentifiable. This is shown in appendix C.1.

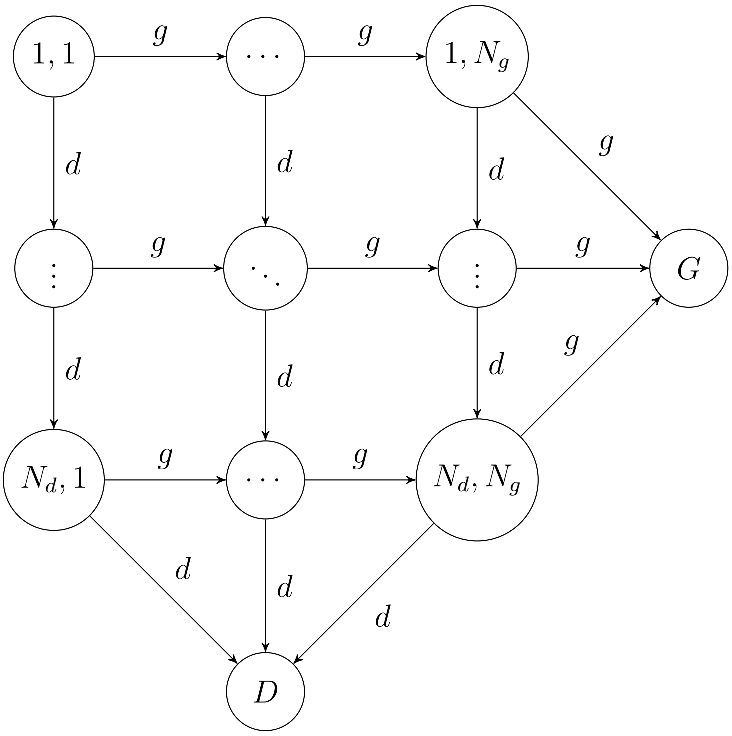
$1,1$
$\cdots$
$1,N_g$
$\vdots$
$\ddots$
$\vdots$
G
$N_d,1$
$\cdots$
N_d,N_g
D
g
d

Figure 4.8: Schematic of the birth death process of macrophages. Macrophages division and death follow two Erlang processes in parallel with parameters (g, N_g) and (d, N_d) for the birth and death processes respectively.

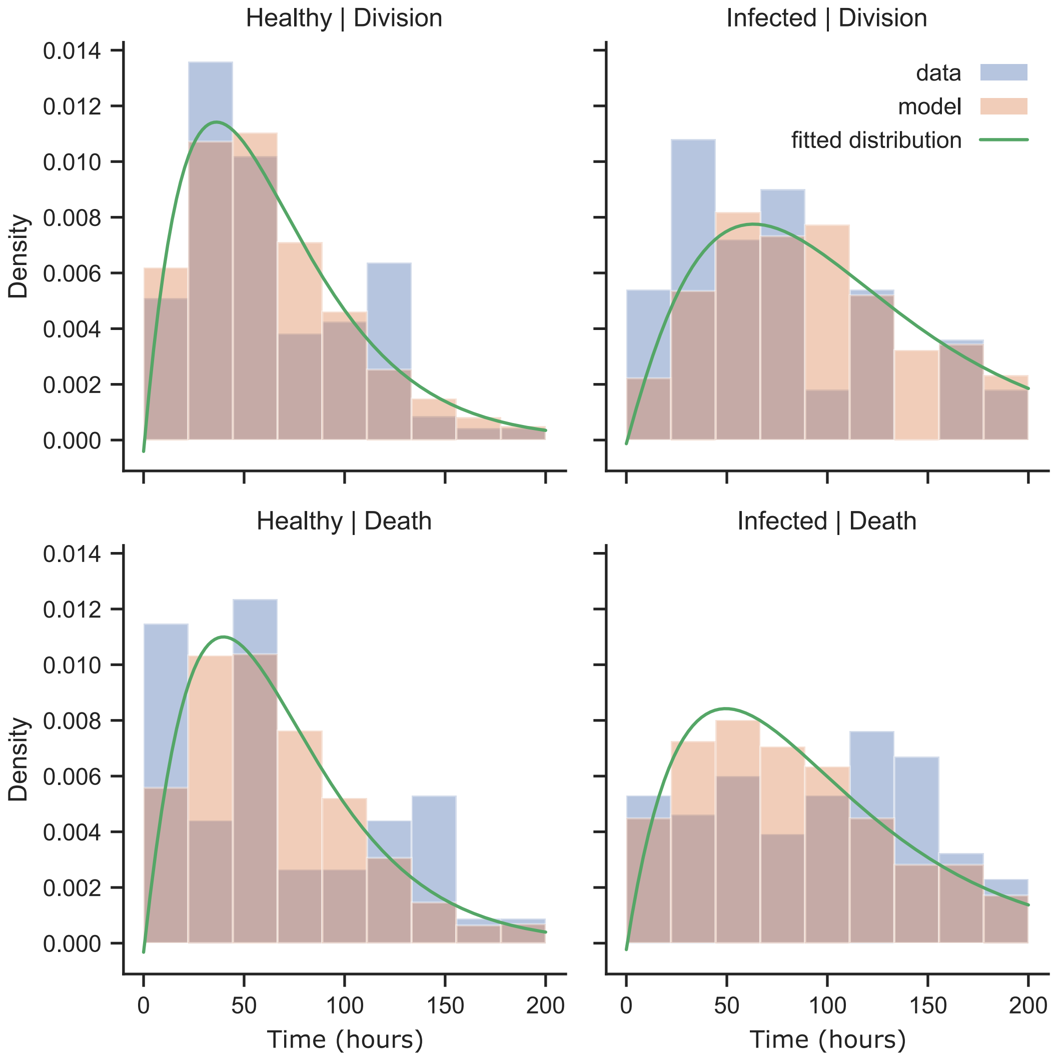

Figure 4.9: Using a CTMC using Erlang distributions for both growth and death results in a good fit to the data.

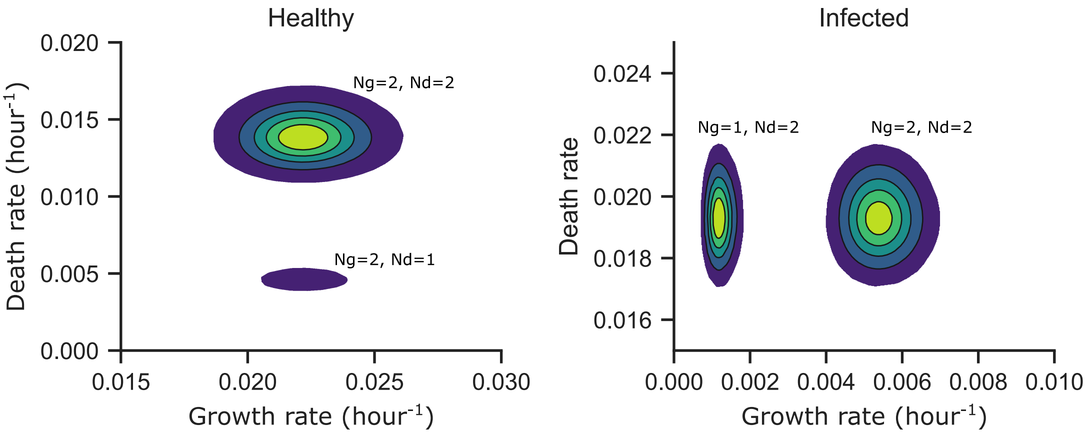

Figure 4.10: 95% confidence regions for the parameters governing the birth death processes of healthy and infected macrophages. The 95% confidence regions are calculated by evaluating the likelihood function over a wide range of parameters and identifying the set of parameters which cover 95% of the total likelihood.

4.7 Infected macrophage division

As in the previous models, when an infected macrophage divides, there is a chance that it either produces two infected macrophages, or one infected macrophage and one healthy macrophage. Figure 4.11 shows the resultant likelihood of fitting a simple bernoulli process. The MLE is slightly reduced compared to the previous model as a result of a new data set being used, although the difference is not significant and both confidence intervals overlap; in the last chapter the 95% confidence interval was $(0.1361, 0.5095)$ while the 95% confidence interval in this chapter is $(0.0661, 0.3844)$.

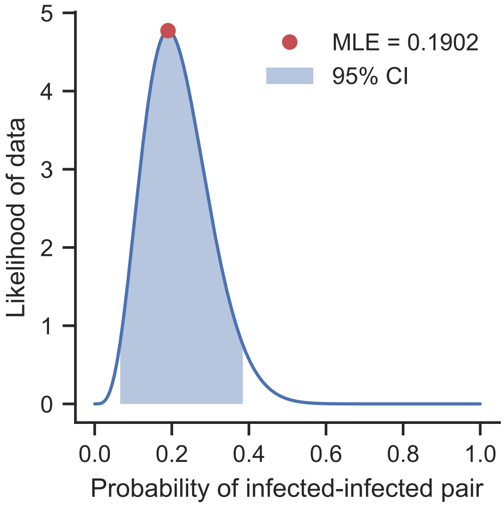

Figure 4.11: Likelihood of observing a infected macrophage dividing into two infected daughter cells.

4.8 Model parameters

The parameters fitted in the previous section and to be used for the rest of this chapter are shown in table 4.3. Comparing these values with those in the previous chapter, the growth rates of bacteria both intracellular and extracellular are approximately equal, however there is a significant change in the reduction in the birth death cycles for both healthy and infected macrophages. In addition, the previous parameters resulted in similar cycles for the two cell types, however in this chapter, the growth of infected macrophages is significantly reduced compared to their healthy counterparts. This is more in line with the literature, wherein the bacteria are said to modify the cell cycle of their host macrophages [127, 128, 130], in order to benefit more from the protective nature of being intracellular.

A further significant difference between the fitted parameters of the two models, is the reduction in the rate at which bacteria are taken up by macrophages. This is a result of the much larger range in the distribution of times between birth and infection for macrophages. Since the model assumes homogeneous mixing between the healthy macrophage and an exponentially growing population of bacteria, increasing the range of times to infection will exponentially increase the force of infection being felt by the macrophage, thus exponentially reducing the value of β.

Parameter	**Description**	**Value**	**95% CI**	**Units**
α_E	Extracellular growth rate	0.0339	$(0.0085, 0.0832)$	hour^{-1}
α_I	Intracellular growth rate	0.0206	$(0.0038, 0.0392)$	hour^{-1}
β	Infection rate	2.6290×10^{-4}	$(2.2558, 3.0304) \times 10^{-4}$	cell^{-1}hour^{-1}
g_H	Erlang growth rate for healthy macrophages	0.0230	$(0.0188, 0.0260)$	hour^{-1}
Ng_H	Erlang growth stages for healthy macrophages	2	$(2, 2)$	scalar
d_H	Erlang death rate for healthy macrophages	0.0146	$(0.0109, 0.0172)$	hour^{-1}
Nd_H	Erlang death stages for healthy macrophages	2	$(1, 2)$	scalar
g_I	Erlang growth rate for infected macrophages	0.0054	$(0.0040, 0.0070)$	hour^{-1}
Ng_I	Erlang growth stages for infected macrophages	2	$(1, 2)$	scalar
d_I	Erlang death rate for infected macrophages	0.0193	$(0.0171, 0.0217)$	hour^{-1}
Nd_I	Erlang death stages for infected macrophages	2	$(2, 2)$	scalar
p	Probability of infected-infected division event	0.1902	$(0.0661, 0.3844)$	scalar
N	Maximum bacterial load	50	cells	

Table 4.3: Model parameters, descriptions, maximum likelihood values and 95% confidence intervals. Note that the confidence intervals for the birth death processes are parameter spaces, and are better captured in a contour plot. These are shown in figure 4.10.

4.9 Model

The model presented in this chapter is a non-spatial agent-based model (ABM) with two distinct populations: the extracellular bacteria and the macrophages. As described in section 1.6.3, to simulate the model, events will be assumed to follow a Poisson point process whose rates depend on the current state of the system, and the system will be evolved through time using the Gillespie algorithm [49]. Thus the definition of the model will be a set of events which occur at a given rate and result in a given action. The event names, rates of occurrence, and resultant actions are depicted in figure 4.12 and are summarised in table 4.4. Further, appendix B shows the python code used to define the model.

The first population modelled is the extracellular bacteria. This population does not require the agent based framework as it is a single value and can be thought of as an environmental variable. The total population of extracellular bacteria E grows exponentially with fixed exponential growth parameter α_E, thus the total rate at which extracellular bacteria are added to the system as a result of extracellular bacteria growth is $\alpha_E E$.

The second population modelled is the macrophage population. Each macrophage is considered as a distinct agent within the system and relies on two internal variables: the age of the macrophage, and the intracellular load of the macrophage. The age of the macrophage is itself defined by two variables given by the two parallel Erlang models fitted in section 4.6 while the intracellular load is simply a count of the number of intracellular bacteria currently residing within the given macrophage. Some events affecting the macrophage will have rates that depend on the current intracellular load; when the intracellular load is zero the macrophage is considered healthy and parameters will be denoted with a H, when the intracellular load is greater than zero the macrophage is considered infected and parameters will be denoted with a I.

The macrophage birth death process is dependent on the infection status (H, I) of the macrophage and is governed by four parameters: $g_{H,I}$, $N_{g(H,I)}$, $d_{H,I}$, $N_{d(H,I)}$. With rate $g_{H,I}$ the macrophage will progress towards division until its division age reaches $N_{g(H,I)}$, at which point the macrophage will divide. The division process involves removing the parent macrophage from the system and adding two new daughter macrophages to the system. If the parent macrophage was infected at the time of division, with probability p both daughters will be born infected, otherwise one daughter will be born with all of the parents intracellular load while the other daughter cell is born healthy. In the case of an infected macrophage dividing to produce two new infected daughter cells, the intracellular load of the parent cell is divided evenly between the daughters. If the intracellular load is an odd number of cells, one daughter will be initialised with one extra bacteria. With rate $d_{H,I}$ the macrophages progress towards death, when a macrophages death age reaches $N_{d(H,I)}$ the macrophage will die. When a macrophage dies it is simply removed from the system. If that macrophage was infected at the time of death, its intracellular load is added to the extracellular population before the macrophage is removed.

Infected macrophages have a maximum load for the intracellular bacteria. The intracellular bacteria grow exponentially in a similar fashion to the extracellular bacteria. Thus for each infected macrophage M_i intracellular growth events occur at rate $\alpha_I I_i$ where I_i is the intracellular load of macrophage M_i. When the intracellular load reaches the maximum load N, the macrophage ruptures releasing the intracellular bacteria back to the extracellular space. Thus the macrophage is removed from the system and its intracellular load is added to the extracellular population.

Finally, healthy macrophages interact directly with the extracellular bacteria through mass action in order to internalise the bacteria. This event occurs with rate parameter β and simulates phagocytosis of the bacteria. Thus for each healthy macrophage M_i, with rate βE one extracellular bacteria is removed from the extracellular space and added to the intracellular load of macrophage M_i, causing it to now be classed as infected. As in previous chapters, when a healthy macrophage becomes infected, its age is reset.

Event	**Rate**	**Action**
Extracellular growth	$\alpha_E E$	$E \to E + 1$
Macrophage growth	$g_{H,I}$, $\forall M_{H,I}(a_g < N_{g(H,I)})$	$M(a_g) \to M(a_g + 1)$
	$d_{H,I}$, $\forall M_{H,I}(a_d < N_{d(H,I)})$	$M(a_d) \to M(a_d + 1)$
Healthy division	g_H, $\forall M_H(a_g = N_{gH})$	$M_H \to \emptyset$
		$\emptyset \to 2M_H$
Infected division	pg_I, $\forall M_I(a_g = N_{gI})$	$M_I(I) \to \emptyset$
		$\emptyset \to 2M_I\ ({}^{I}\!/_{2})$
Infected-healthy division	$(1 - p)g_I$, $\forall M_I$	$M_I(I) \to \emptyset$
		$\emptyset \to M_H + M_I(I)$
Healthy death	d_H, $\forall M_H(a_d = N_{dH})$	$M_H \to \emptyset$
Infected death	d_I, $\forall M_I(a_d = N_{dI})$	$M_I(I) \to \emptyset$
		$E \to E + I$
Intracellular growth	$\alpha_I I$, $\forall M_I(I)$	$M_I(I) \to M_I(I + 1)$
Macrophage rupture	$\alpha_I N$, $\forall M_I(I = N)$	$M_I(I = N) \to \emptyset$
		$E \to E + N$
Infection	βE, $\forall M_H$	$E \to E - 1$
		$M_H \to M_I(a_g = a_d = 1, I = 1)$

Table 4.4: List of possible events within the model. Each event has a rate at which it occurs, and a resultant action if it occurs. $E \to E + 1$ denotes the extracellular bacteria population increasing by 1. $M(a_g)$ represents a macrophage with Erlang division parameter a_g. $M_I(I)$ represents an infected macrophage with I intracellular bacteria.

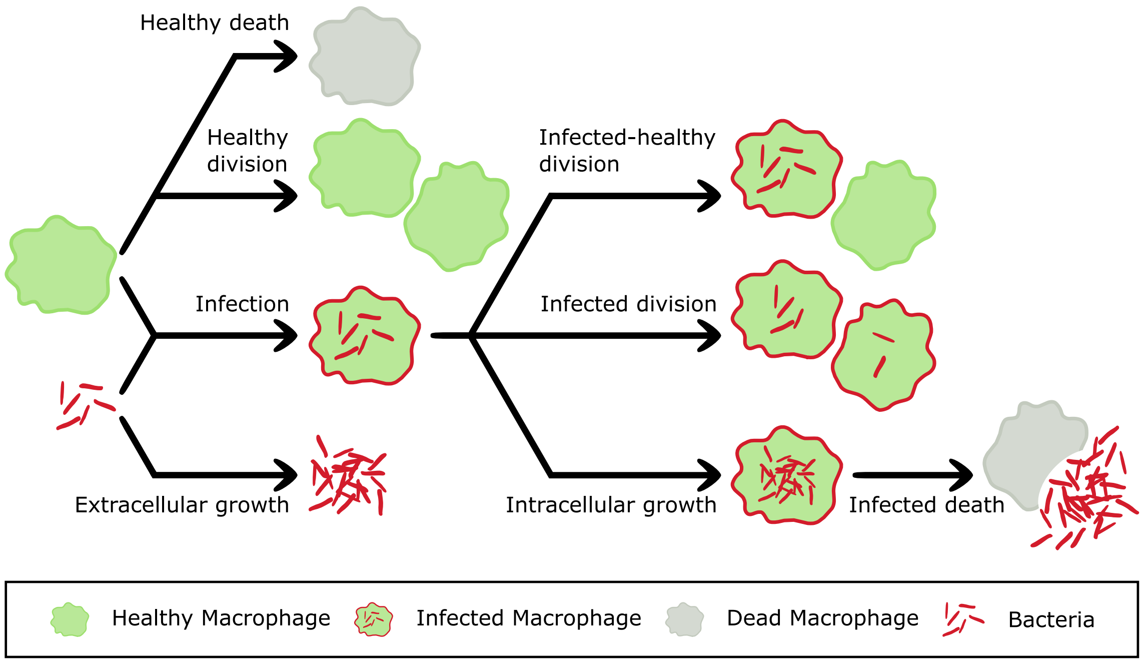
Healthy death
Healthy division
Infection
Extracellular growth
Infected-healthy division
Infected division
Intracellular growth
Infected death
Healthy Macrophage
Infected Macrophage
Dead Macrophage
Bacteria

Figure 4.12: The core mechanisms modelled in this chapter are depicted in this schematic.

4.10 Model results

In this section the general output of the model will be analysed before performing a sensitivity analysis. The results of the sensitivity analysis will then inform the construction of the computational experiments described in section 4.11.

4.10.1 Model output

Figure 4.13 shows the results of running multiple simulations with parameters set to the MLE values. The solid lines represent the average of the simulations while the shaded area shows the standard deviation of the samples. A total of 1000 samples were generated through repeated simulation, stopping at 1000 as after this point variation in the average and the standard deviation was not noticeable.

The model is initialised with ten macrophages and one bacterium. The bacterium is initialised extracellularly for 50% of the simulations and intracellularly for the other 50%. Comparing this to the model output from the previous chapter in figure 3.6, there are some notable differences. In the stochastic model the bacteria population,

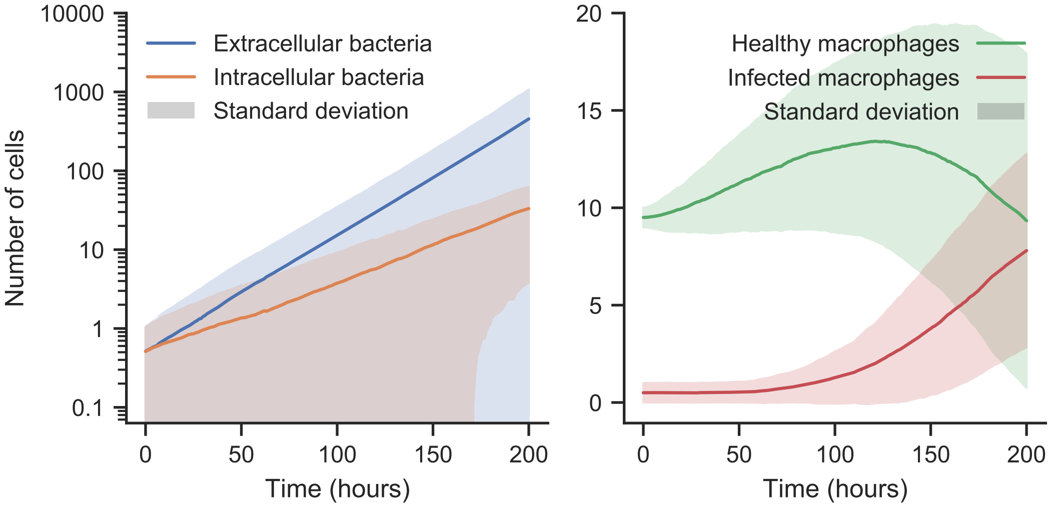

Figure 4.13: 1000 repetitions of the stochastic model with initial conditions of ten macrophages and one bacteria which is initialised intracellularly for 50% of the runs and extracellularly for the other 50%.

both extracellular and intracellular, are each an order of magnitude larger than the deterministic model. Since the intracellular growth rate for the stochastic model is approximately half that of the deterministic model, the effect of the small increase in the extracellular growth rate is dominating the bacteria growth in both states.

In chapter 3 the population of healthy macrophages rapidly dies out before the end of the 140 hour simulation. Comparatively, after the same amount of time in the stochastic framework, the healthy macrophages have just reached their peak. With the population of extracellular bacteria now growing exponentially the rate at which macrophages are becoming infected is beginning to overwhelm the rate of macrophage division.

Figure 4.14 shows the results of inserting the parameters fitted in this chapter into the model developed in chapter 3. The major change in the parameter values is the increase in both the extracellular bacteria growth rate and the intracellular bacteria growth rate. The rapidly growing extracellular population applies a high force of infection on the healthy macrophages, resulting in their population quickly depleting. The healthy macrophages die out after approximately 75 hours, compared to not

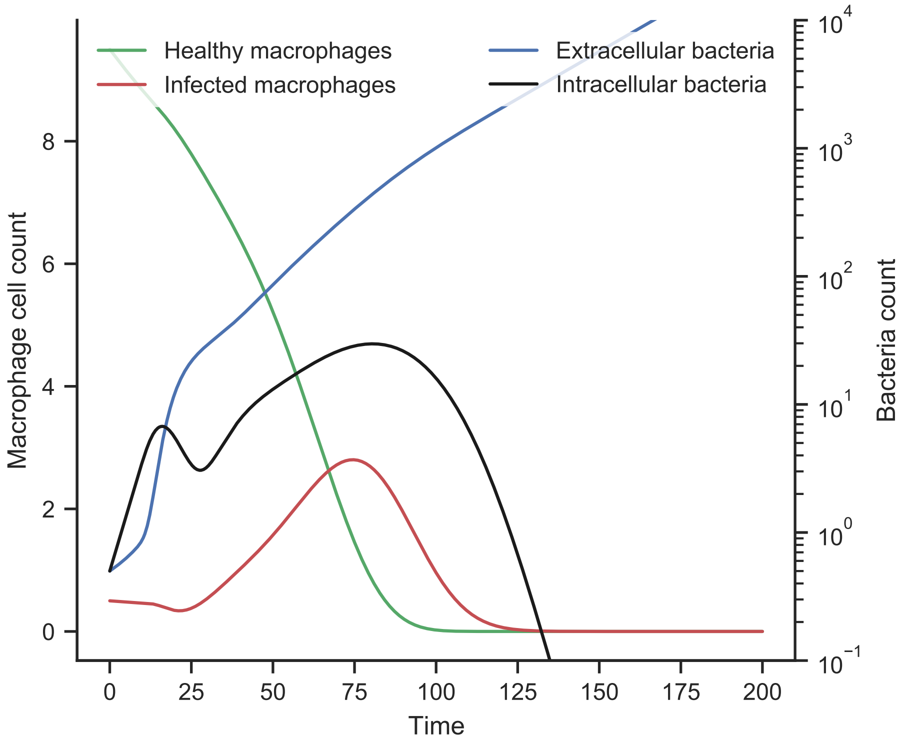

Figure 4.14: Result of running the model presented in chapter 3 with the parameters identified within this chapter.

dying out in the stochastic model, and surviving to 100 hours in the previous chapter. Shortly after the healthy macrophages go extinct, so do the infected macrophages, as a result of internal pressure from their intracellular loads. After approximately 150 hours the only population left is the extracellular bacteria which are growing exponentially.

Putting the parameters from this chapter into the deterministic model from the last chapter demonstrates the need for the model to be stochastic. Since the populations are small, stochastic fluctuations in the number of individual cells will have a larger relative impact on the system than if the population was larger. These fluctuations cannot be captured by the deterministic model.

4.10.2 Sensitivity analysis

In this section a sensitivity analysis will be performed on the model using the Sobol sensitivity analysis method, as described in section 1.6.5. Given the model presented here is a stochastic model, it is not computationally feasible to run the analysis on the same number of parameters as in previous chapters. Additionally, each parameter set needs to be run multiple times in order to smooth out the effects of stochasticity. This results in larger uncertainty on the estimated sensitivity indices. Hence, a balance needs to be found between exploration of parameter space, and repetitions of simulations [185].

Figure 4.15 shows the results of the Sobol sensitivity analysis when measuring the total population of extracellular bacteria, healthy macrophages, and infected macrophages over the 200 simulated hours. Each population is sampled at regular time intervals of 5 hours to observe how the model sensitivity evolves over time. As described in section 1.6.5, the sensitivity indices measure the contribution of each parameter to the variation in the model output.

As in the previous models the variance observed in the extracellular bacteria population is dominated by variance in the extracellular growth (α_E) and the intracellular growth (α_I). In fact no other parameters demonstrate a significant effect on the extracellular population.

Variance in the healthy macrophage population predominantly originates from variance in the birth (g_H) and death (d_H) rates, however as the extracellular bacteria population grows, the extracellular growth rate (α_E) takes overs as the dominant parameter. As the extracellular bacteria population grows larger, fluctuations α_E rate will have an increasing impact on the total number of extracellular bacteria, hence will also have an increasing impact on the force of infection felt by the healthy macrophages. This relationship can be directly observed in the infected macrophages, as towards the end of the simulation, the extracellular growth rate is the largest contributor here too. The probability of healthy-infected division events also has a

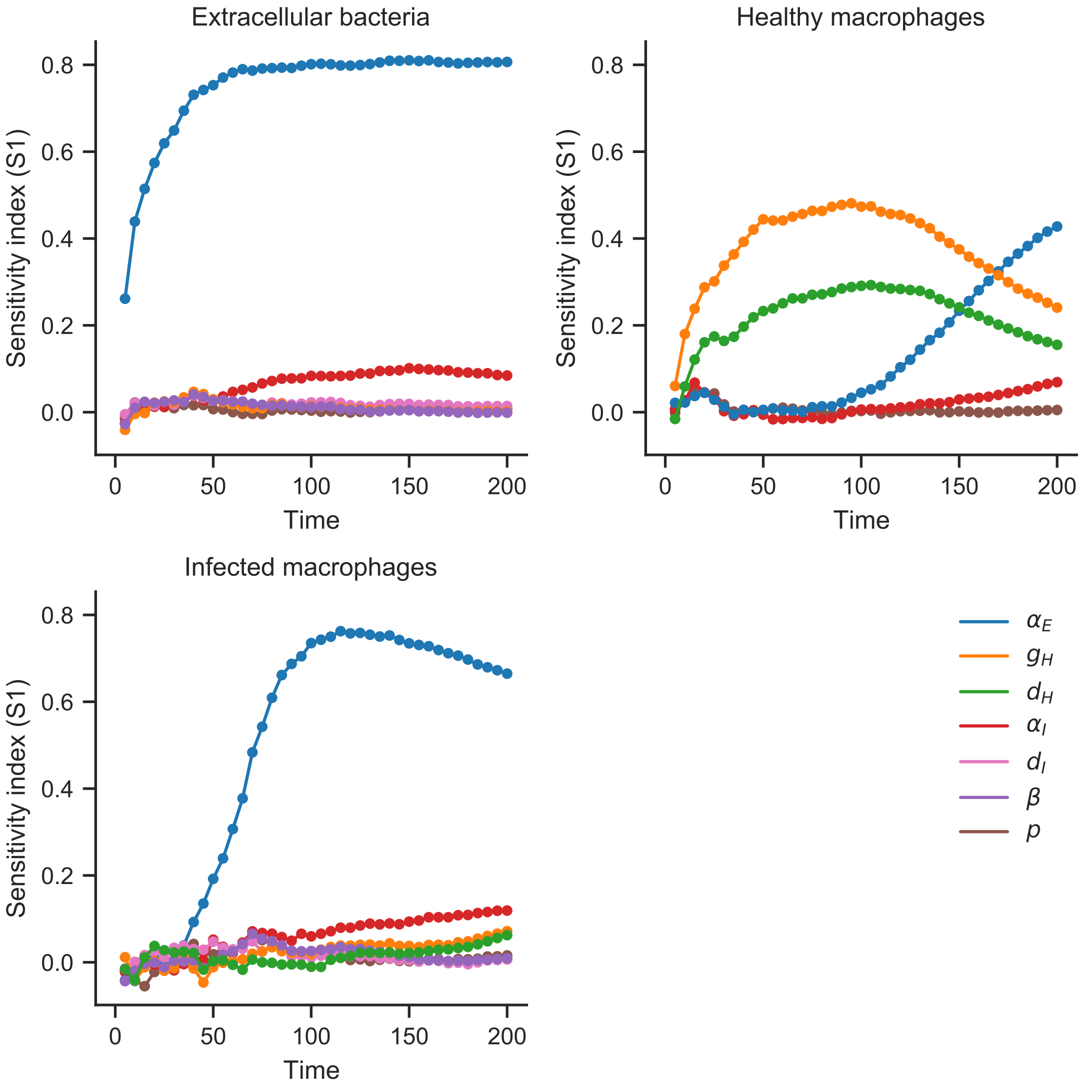

Figure 4.15: Sobol sensitivity analysis measuring the fraction of the variance in the modelled populations over time that can be attributed to variance in the given parameters.

significant effect on the healthy macrophage population, stemming from division in the infected macrophage population.

As seen in figure 4.13, after around 100 hours of simulation is when the infected macrophage population begins to take off. This is just after the sensitivity to the extracellular growth rate has a steep increase to become the dominant parameter for infected macrophages. As alluded to already, this dependence on the extracellular growth rate for the infected macrophages is a result of the force of infection felt by the larger population of healthy macrophages. A pressure which in turn converts these healthy macrophages into infected macrophages. While the extracellular growth rate contributes to the majority of the observed variance, the intracellular growth rate (α_I) and the birth-death process for both populations of macrophages also contribute.

In previous chapters, a second sensitivity analysis was performed measuring the time until $\frac{\mathrm{d}E}{\mathrm{d}t} > 0$. This is not possible in this scenario, as the model is now stochastic. Instead, figure 4.16 shows the sensitivity analysis measuring the time until the extracellular population exceeds ten. While in previous models, this construction has demonstrated sensitivity to the infection rate and the macrophage birth death process, in the stochastic construction presented here, the variation in model output is dominated by variation the the extracellular growth rate α_E. This is due to the low MLE value of β resulting from the homogeneous mixing assumption between extracellular bacteria and healthy macrophages.

The results of the sensitivity analysis show that the dominant mechanism affecting the level of free bacteria during the early interactions is the extracellular growth rate. Since innate alveolar macrophages have no control over this mechanism, it is not natural to apply controls here. Instead, since a common feature of activated macrophages is to inhibit the intracellular growth, and since the data demonstrate a wide range in intracellular growth rates as shown in figure 4.2, this mechanism will be focused on in the next section using computational experiments.

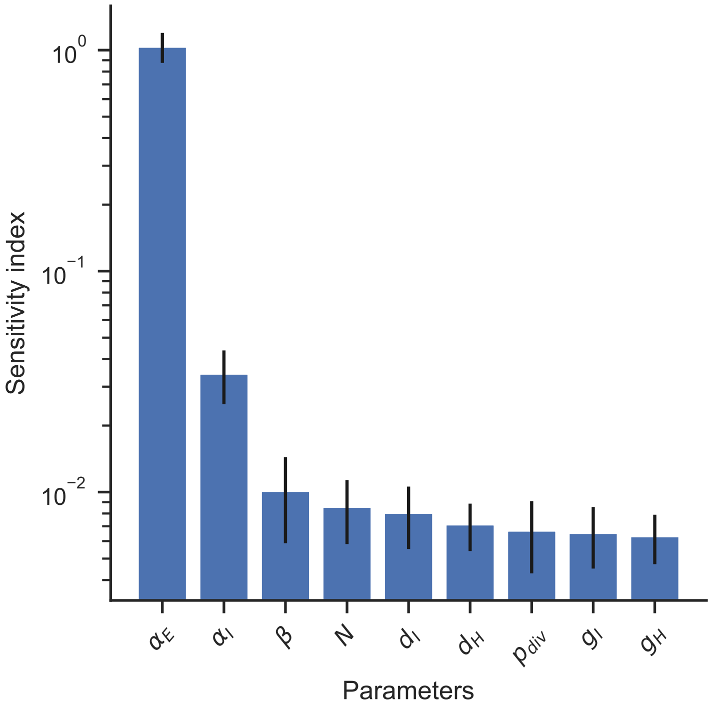

Figure 4.16: Sobol sensitivity analysis measuring the time taken for the extracellular bacteria population to exceed ten. Since stochastic variation will add noise, the time until the gradient of extracellular growth is positive cannot be used. An extracellular population of 10 is large enough that the chance of being significantly reduced is negligible, while being small enough that the simulation is quick, thus a large number of parameters can be sampled.

4.11 Computational experiments

In this section, two computational experiments will be performed, analysing the effect of intracellular growth inhibition as a method of control on the extracellular bacteria population.

4.11.1 Effect of intracellular growth rate

By allowing the rate of intracellular growth to vary from its MLE value of $\alpha_I = 0.0206$ to within the interval $(0.0, 0.1)$, the relationship between this parameter and the

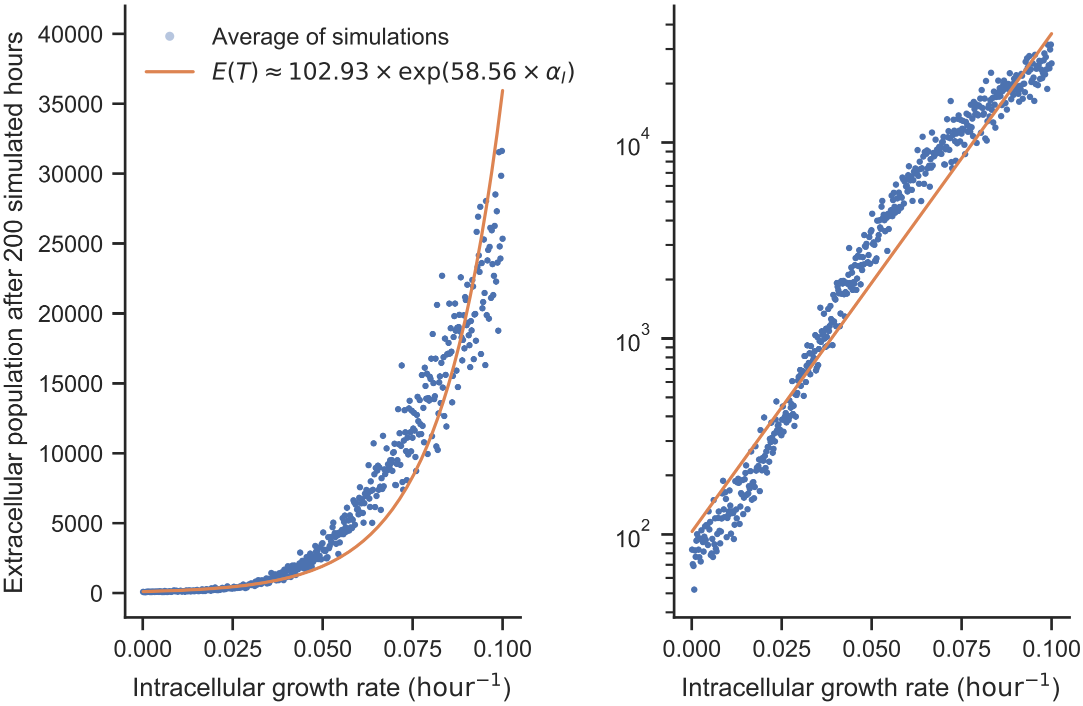

Figure 4.17: The relationship between the resultant extracellular bacteria population and the intracellular growth rate is approximately exponential. For each value of α_I, the simulation was run 100 times. The displayed points are the average of those simulations.

resultant population of extracellular bacteria can be observed. The population of extracellular bacteria after 200 simulated hours is being used as a proxy for bacterial virulance, as a larger population prior to adaptive immunity being engaged will naturally be more difficult to control.

Figure 4.17 shows the approximate exponential relationship between the resultant extracellular population and the intracellular growth rate. While not very surprising, the nature of such an exponential relationship demonstrates how a small increase in the intracellular growth rate can lead to a large increase in total bacteria.

If grown in the absence of macrophages, the expected number of extracellular bacteria after 200 hours is approximately 900. Figure 4.17 shows that for sufficiently low values

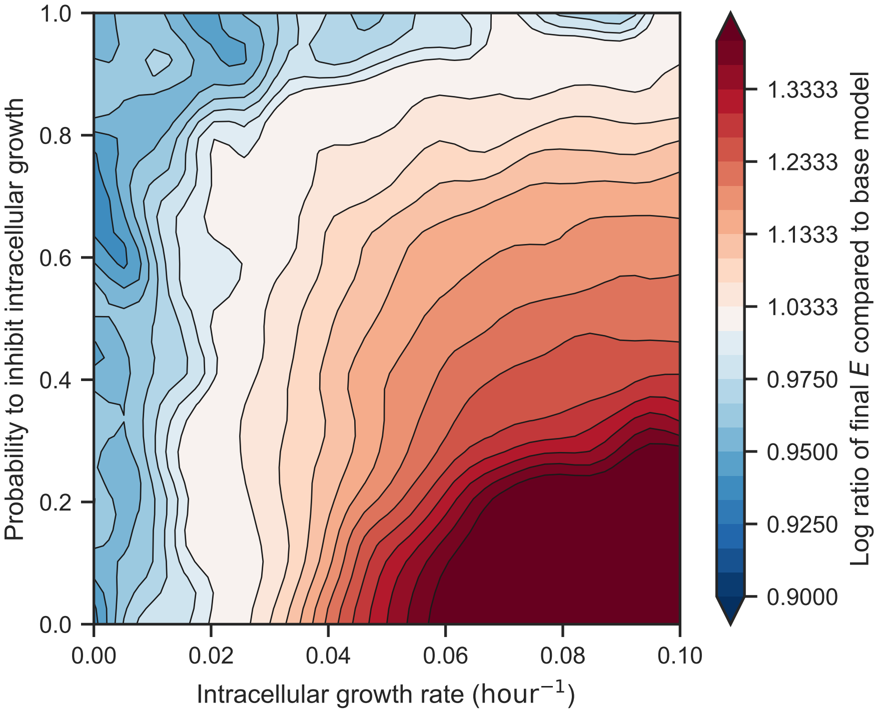

Figure 4.18: A small increase in the intracellular growth rate requires a large increase in the probability to inhibit intracellular growth in order for the outcome not to change. Note the non-linear colour bar centered at 1.

of α_I, the population of extracellular bacteria is well below this limit, demonstrating the potential ability for macrophages to limit the extracellular population.

4.11.2 Effect of probability of growth inhibition

While increasing the intracellular growth rate results in a large increase in the bacteria, due to the exponential nature of the relationship, reducing the intracellular growth rate only results in a small reduction. Including a probabilistic ability for macrophages to completely inhibit intracellular growth is a further mechanism to control the infection. Figure 4.18 shows a heat map of the relationship between intracellular growth rate and the probability to inhibit growth. The measure used is

$$\frac{\log\left(E\left(200;\alpha_I,p_{\text{inhib}}\right)\right)}{\log\left(E\left(200;\alpha_I=0.0206,p_{\text{inhib}}=0\right)\right)}. \tag{4.8}$$

This measure was chosen as it approximately linearises the values of the exponentially growing bacteria population, and allows the reader to quickly identify the boundary between the two parameter sets: those that result in a larger and faster growing extracellular population compared to those that result in a smaller and slower growing extracellular population.

For more virulent strains, demonstrated here with an increased intracellular growth rate, the macrophages must respond with a large increase in the probability to inhibit intracellular growth in order to maintain the same level of extracellular bacteria at the end of the simulation.

The ability for macrophages to inhibit intracellular growth does not provide a benefit, as can be seen by the vertical stratas for α_I close to and below the MLE. The ability for macrophages to inhibit intracellular growth provides a small benefit in the system, as can be seen by the near vertical strata for α_I close to and below the MLE. The magnitude of the benefit improves as α_I increases, as this ability allows macrophages to limit the rate at which the infection gets out of control.

4.12 Conclusions

In this chapter, a stochastic agent-based model has been developed and fitted to an extended set of experimental data. By moving to a stochastic framework, small population effects are able to be modelled more accurately and allow for probabilistic extinction events of the macrophages. In addition the macrophage birth-death process has been improved, such that the distributions of wait times for division and death events from the model more closely resemble the equivalent distributions in the data. Through sensitivity analysis it was shown that both extracellular bacteria and the infected macrophages are highly dependent on the extracellular growth rate, almost to the exclusion of all other mechanisms. This is a result of the homogeneous mixing between healthy macrophages and extracellular bacteria. Due to the homogeneous mixing of macrophages with the total extracellular population, the fitting process pushed the rate of infection very low, limiting the amount of interaction between the two cell types.

Through computational experimentation, the intracellular growth rate of bacteria was targeted as a candidate for bacteria control. Two experiments were carried out. The first investigated the relationship between the intracellular growth rate and the extracellular population. This was shown to be approximately exponential demonstrating the importance of intracellular growth on the proliferation of extracellular bacteria. The second experiment introduced a theoretical mechanism whereby the macrophages inhibit the intracellular growth of the bacteria. This is similar to the previous chapter, but was deemed a more realistic mechanism. The experiment simulated the response in the extracellular bacteria population when changing the probability to inhibit intracellular growth and when changing the average intracellular growth rate. This experiment demonstrated that reducing the average intracellular growth rate by a small amount for all macrophages has a larger impact on the resultant extracellular than completely reducing the growth rate to zero for some macrophages.

The model presented in this chapter has some conceptual similarities with two Francisella tularensis models [184, 186], whereby infected phagocytes are initialised with a single bacteria and the bacteria are able to grow intracellularly. Similar to the comparison between chapter 2 and Wood, Egan, and Hall [176] there are some significant differences in the model formulations. Firstly, healthy macrophages and macrophage division play key roles in the overall dynamics of the system presented here, most importantly natural death of the macrophage. A core focus of the Francisella tularensis is the expected wait times to cell rupture, however natural death is not included. As shown in appendix C.1, with the data available it has not been possible to differentiate the cause of death of the cell between rupture as a result of load and natural death.

Another model of interest from the literature is model of macrophage division and death known as the cyton model [187, 188], whereby division and death events are modelled by two distributions and the rate of division and the rate of death is simply the integration of the respective distributions. This is a similar construction to the model presented here where this model restricts the two distributions to Erlang distributions rather than the arbitrary positive support distributions presented in the cyton model. The cyton model is also extended to allow these distributions to vary from generation to generation. While running the cyton model through a wide model selection process to pick the optimal distributions for division and death wait times may have resulted in an improved quantitative fit to the data, fixing the distributions to be Erlang as has been done here still allows for a lot of flexibility in the resultant distribution while maintaining a connection to the underlying biology that is driving the process.

The current model assumes macrophages are continuously mixing with all bacteria, and can only take in single bacteria at a time. This has resulted in the MLE for the rate of infection β being pushed low, limiting the number of interaction between healthy macrophages and extracellular bacteria. In the following chapter the structure of the extracellular bacteria will be modified to be distinct populations known as aggregates. Macrophages will then interact with the aggregates at a density dependent rate rather than the individual bacteria. With this modification in mind, additional data was also collected by the experimentalist to explicitly measure the growth rate within these distinct aggregates, and also to measure the quantity of bacteria internalised. The latter being used to remove the constraint that only individual bacteria are picked up by the macrophage.

Chapter 5

Model 4: Stochastic Agent Based Model With Aggregated Extracellular Bacteria

In the previous chapter a stochastic agent-based model (ABM) was developed and fitted to experimental data. A major component of this model was the homogenous mixing of the macrophages with all available extracellular *Mycobacterium tuberculosis* (Mtb) bacteria, and the mass action methodology for the rate of infection. In the data, it was observed that some macrophages persist in the healthy state for the full duration of the experiment. Since the extracellular bacteria are growing exponentially, the force of infection also grows exponentially. The model of infection is based on mass action, thus persistent healthy macrophages experiencing an exponentially increasing force of infection will drive the rate of infection down.

In order to address the limitations resulting from the macrophages homogeneously mixing with the entire bacteria population, in this chapter the extracellular bacteria will be considered to exist as distinct aggregated populations. Each aggregate will be assumed to be a distinct population of extracellular bacteria, which form as a result of an infected macrophage releasing its internal load back into the extracellular space. Interaction will occur between bacteria aggregates and macrophages. This construction will maintain the modelling simplicity of homogeneous mixing and not considering a spatial element. In order to parametrise this model, a range of

Bacteria	Aggregate
Experiment count	46.0000
Average starting fluorescence	108.4471
Average fluorescence	183.9347
Average ending fluorescence	342.6384
Average number of data points per experiment	11.5870

Table 5.1: Summary of data relating to the aggregated bacteria growth. In total 46 populations of bacteria were independently tracked and the total area of fluorescence was measured.

additional experiments were conducted by the experimentalist, specifically measuring the growth rate of extracellular bacteria within distinct aggregates.

5.1 Data

The data that was recorded to enable this transition in the model is aggregate specific fluorescence measurements within the experiments. This means that if multiple aggregates are visible within the same frame, they are recorded as distinct bacteria populations, rather than summed as a single population. Table 5.1 summarises the new data.

5.2 Model description

The model construction in this chapter remains as for chapter 4, a stochastic ABM with homogeneous mixing between agents. However, in addition to the macrophages being modelled as agents, so are the extracellular bacteria. Each extracellular bacterium exists in one of several distinct populations of extracellular bacteria. These populations will be called aggregates for the purposes of this study. A new aggregate can only be formed when an infected macrophage dies releasing its intracellular load back in to the extracellular space. When a healthy macrophage interacts with the extracellular bacteria, it will be at the macro scale of macrophages interacting with aggregates rather than directly with individual bacteria. This iteration of the model will act to reduce the effect of homogeneous mixing between all bacteria and the macrophages which resulted in a very low rate of infection being fitted in the previous chapter. Further, this will allow for the possibility of a macrophage to internalise more than one bacteria at a time. Either the macrophage will internalise the entire aggregate resulting in the aggregate being removed from the system and all of the extracellular bacteria within the aggregate transitioning to intracellular bacteria within the macrophage. Or the macrophage will only internalise a subset of the aggregate, in a partial uptake event. This results in the aggregate remaining in the system albeit now with fewer bacteria, and some of those bacteria transitioning to intracellular bacteria within the macrophage. The basis for this mechanism is the experimental observation that macrophages are able to either partially consume or entirely consume an aggregate of bacteria.

5.2.1 Mechanisms

There are several additional mechanisms being introduced in this chapter, the majority of which revolve around the core change in the way extracellular bacteria are modelled. Previously the model considered extracellular bacteria as a single population which is homogeneously mixing with the individual macrophage agents. One limitation of this construction is due to infections being modelled using mass action. Since the population of bacteria is growing exponentially, the force of infection felt by macrophages is also growing exponentially.

As already described, the infection process will now be based on mass action mixing between macrophages and aggregated populations of extracellular bacteria. During an infection event, there will be the possibility for the macrophage to fully consume the aggregate, or only partially. If the macrophage only consumes a fraction of the aggregate, this fraction will be modelled dependent on the current population of bacteria within the aggregate. Immediately after the infection event the macrophage will be evaluated for how likely it is to survive, as some macrophages in the experimental data are observed to internalise a large number of bacteria but not survive the process.

In previous models, when an infected macrophage divides, the number of intracellular bacteria is either split evenly between the two daughter cells or remains entirely in a single daughter cell with the other daughter cell being a healthy macrophage. In this chapter, the model will be extended to allow for a range of possibilities on the division of bacteria between the two daughter cells. As shown in chapter 3, there is insufficient data to be able to assume correlation between parent cells and daughter cells. Therefore the model will remain memoryless in terms of whether a cells parent divided to produce two infected cells, or one infected and one healthy cell. The construction for this is discussed more in section 5.3.4

There is high variability observed in the distribution of intracellular growth rates. The possibility for macrophages to inhibit the intracellular growth rate post infection will therefore be investigated as in chapter 4.

5.3 Parameter Fitting

In this section the additional and altered mechanisms will be described mathematically and have their parameters fitted. The newly fitted parameters and the unchanged parameters from the previous chapter are summarised in table 5.2 at the end of this section (page 156).

5.3.1 Aggregated extracellular growth

The experimental data was expanded to focus specifically on the bacteria as individual aggregated populations instead of a single homogeneous population. Rather than measuring the level of fluorescence over the whole frame, the measurements now focus on specific aggregates within each experiment. The fitting process to model

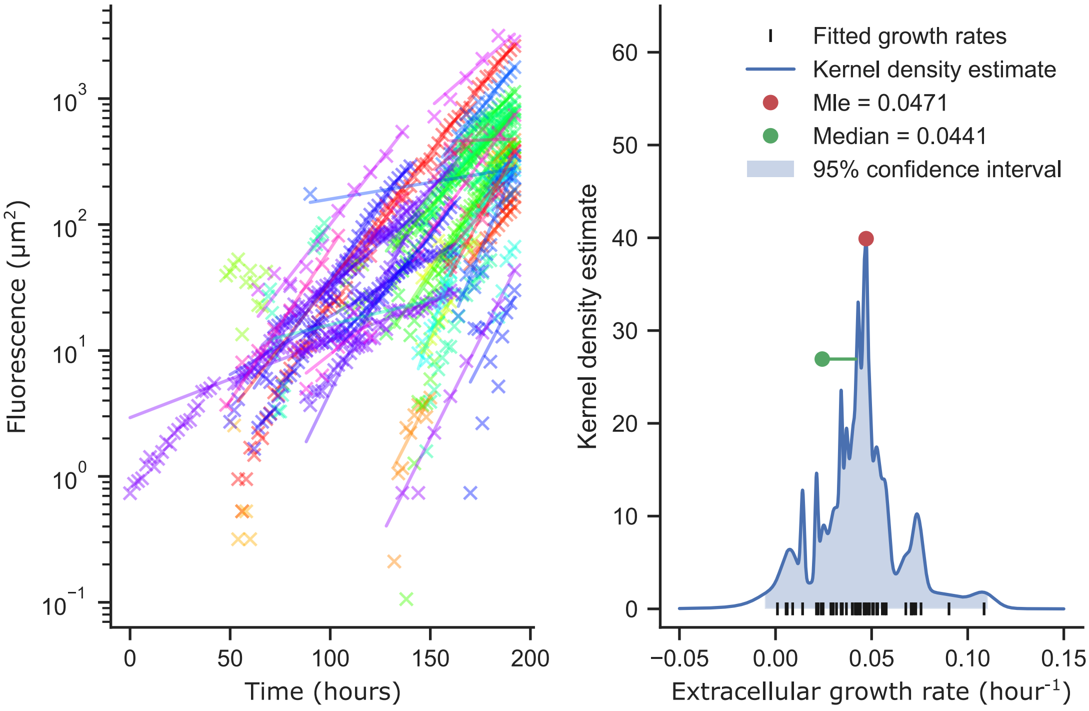

Figure 5.1: Exponential growth curves fitted to the aggregated extracellular growth experiments.

the growth of extracellular bacteria remains unchanged and the outcome is shown in figure 5.1.

5.3.2 Infection

Bacteria are now modelled as aggregates, therefore mass action mixing occurs between the aggregates and the macrophages rather than directly between the bacteria and macrophages. Motivated through discussion with the experimentalist, it was observed that when a macrophage interacts will a population of bacteria, it may internalise the entire population or only part of it - here called partial uptake. While it was observed in the experiments, it was not quantitatively measured in the data. Bootstrapping is used to estimate the likelihood of parameters: the underlying process determining the number of aggregates at any given time is repeatedly sampled generating a probability distribution.

As for previous chapters, the initial number of bacteria aggregates will be assumed to be one. The parameter that is bootstrapped over will be the probability of a macrophage consuming an entire aggregate. This probability will be assumed to be independent of the size of the aggregate. When a death of an infected macrophage occurs in the data, the number of aggregates will always be assumed to increase by one, and when an infection of a healthy macrophage event occurs, the number of aggregates will decrease by one with some fixed parameter to be determined. This is in line with the observations in the data, as bacteria do not disperse post macrophage apoptosis, and the observation within the data that the aggregates are not always internalised fully is the foundation of this mechanism.

The underlying data for fitting the infection process is the same as in previous chapters, and has simply been extended to include additional experiments. The time of birth, infection and death for each infected macrophage is measured. Within a single experiment, these are used to define a probability distribution of the number of extracellular aggregates, since this data is not recorded.

Figure 5.2 shows the raw experimental data to be used for fitting the infection process. Each line represents a single macrophage, while the colours represent different experiments. Since four of the nine experiments exhibit a large amount of activity relating to infection events, the space of possible aggregate time series is prohibitively large. Instead, bootstrapping will be used to estimate the model parameters. A set of time series of the number of aggregates will be generated according to the observed infection and infected macrophage death events and the probability of partial uptake, p_{partial}. A set of examples is shown in figure 5.3.

The likelihood function is defined as in the to previous chapter, by measuring the force of infection felt by a macrophage between the time it was born to the time it became infected. The force of infection is defined by the number of active aggregates at any time, which is itself unknown and parameter dependent. The solution used here is to first bootstrap the number of possible aggregates by assuming the simulation starts with a single aggregate, then observe infection and infected macrophage death events

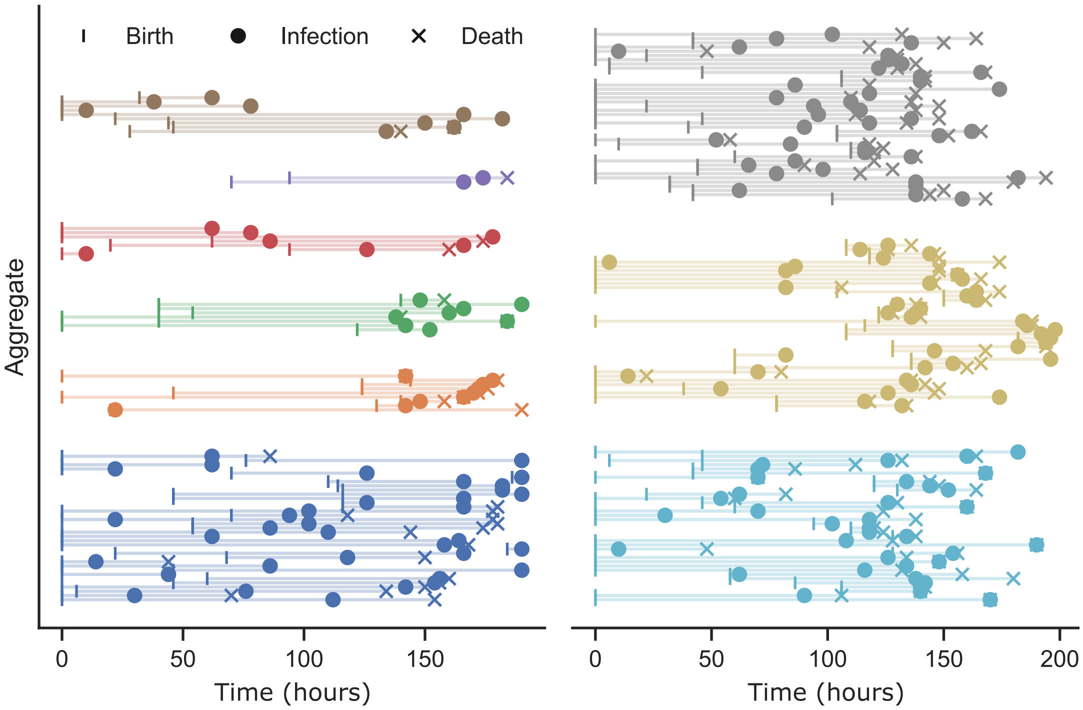

Figure 5.2: The individual experiments in the data set for this chapter show a lot of activity regarding infection times of the macrophages. In this plot each coloured group represents an experiment, and each line represents a cell that becomes infected at some point during the experiment. Of the ten experiments that tracked macrophages at the individual level, there were 9 that observed infection events, four of which demonstrate a lot more activity than the other five. Since macrophages move around while infected, more infections early on the experiment will rapidly increase the number of aggregates available, thus further increasing the number of observed infection events.

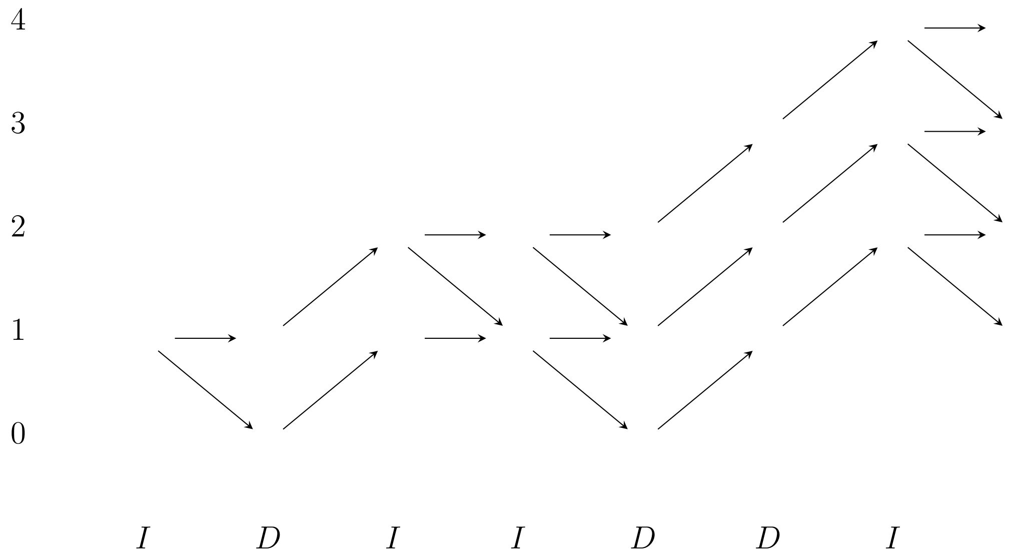

Figure 5.3: The number of aggregates in an experiment is not measured, so the infection rate is bootstrapped over the possible number of aggregates. When a macrophage dies, a new aggregate is formed. When a macrophage becomes infected, there is a Bernoulli process which determines whether the macrophage consumes the entire aggregate, resulting in the number of aggregates going down, or the macrophage only partially consumes the aggregate, resulting in no change to the number of aggregates. This plot shows the an example experiment and how the possible number of aggregates changes according to events within a single experiment. Along the x axis an I represents a healthy macrophage internalises either all or part of an aggregate, and a D represents the death of a macrophage.

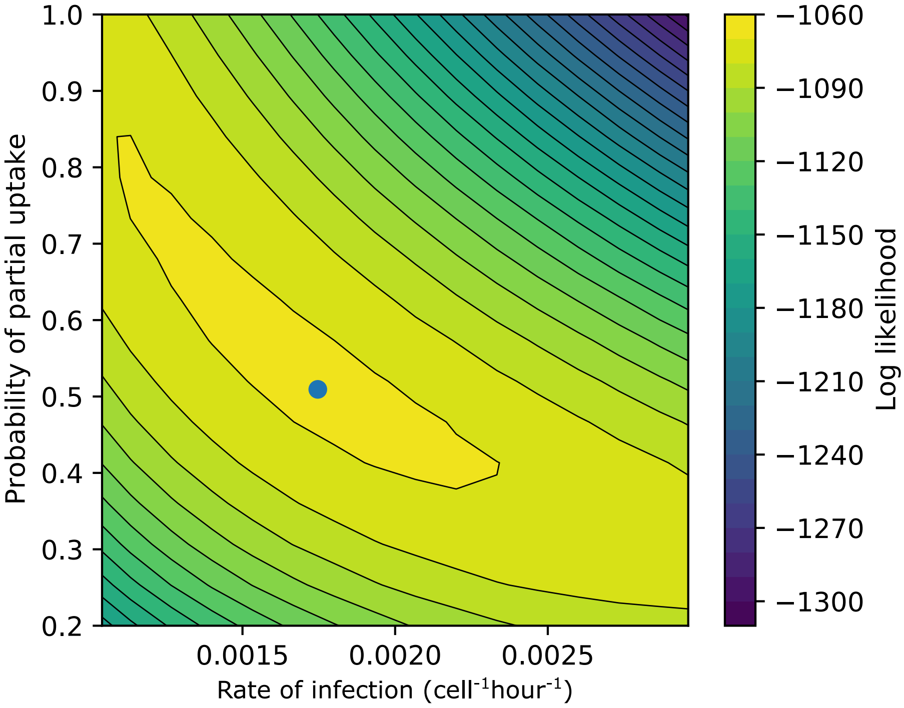

Figure 5.4: Likelihood landscape for the rate of infection (β) and the probability of a macrophage partially internalising an extracellular aggregate (p_{partial}).

to infer the current number of aggregates. These events affect the population in the two ways. Firstly, when a macrophage interacts with an aggregate to become infected, there is some probability that it will consume the whole aggregate, or only part of it, resulting in the number of aggregates either staying the same, or reducing by one. Secondly, when an infected macrophage dies, it will deposit its intracellular load into the extracellular space creating a new aggregate, thus increasing the population of aggregates by one. The likelihood landscape for the probability of partial uptake (p_{partial}) and the rate of infection (β) is shown in figure 5.4. During a partial uptake, the number of bacteria internalised is assumed to be uniformly distributed on the number of bacteria available. This is explored further in appendix C.2.

5.3.3 Post infection death rate

It was observed in the experiments that some macrophages that internalise a large number of bacteria do not survive to the next sampled time period. Investigating the data highlighted that the number of bacteria internalised correlated with the probability of the macrophage not surviving. To address this mechanism, the probability of a given bacterial load (x) causing the macrophage to die will be defined by

$$f(x) = \frac{x}{a + x}. \tag{5.1}$$

Analogously the probability the macrophage survives will be given by $1 - f(x)$. f therefore defines a Bernoulli process for each possible load x. The value of a is determined using least mean squares using the data of the initial bacteria load of a newly infected macrophage, and whether the macrophage survived to the next time sample in the experiment. The data and the fitted curve f are shown in figure 5.5.

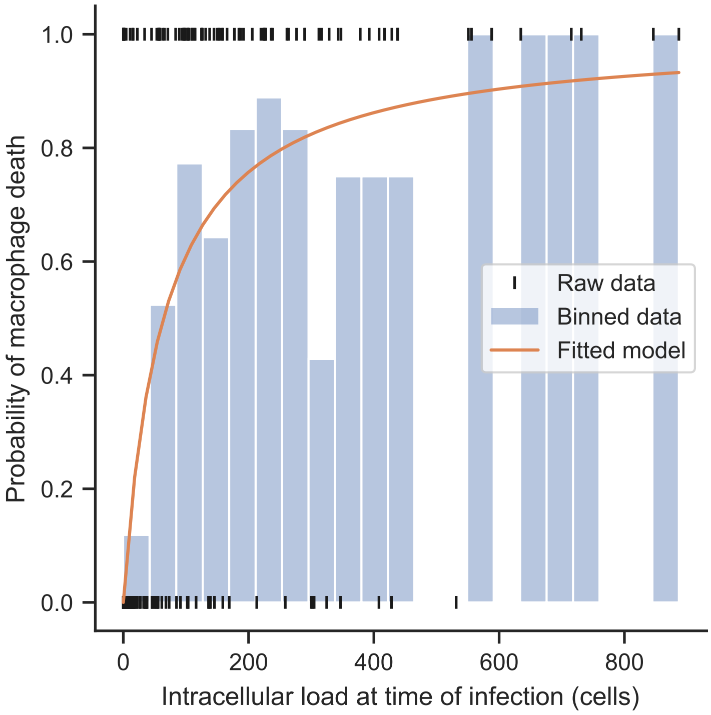

Figure 5.5: Scatter plot of the number of initially internalised bacteria compared to whether that cell died within the next eight hours. The data is binned into 21 bins of equal width order in better to visualise the trend. Using least mean squared difference fitting results in an maximum likelihood estimate (MLE) of $a = 64.03$.

5.3.4 Infected macrophage divisions process

In previous chapters, infected macrophages have been assumed to divide in one of two ways, governed by a simple Bernoulli parameter. Either a macrophage divides into an infected-healthy pair of daughter cells, in which case all the intracellular bacteria will move to the infected cell, or the parent cell will divide into two infected daughter cells, in which case the intracellular bacteria will be divided evenly between the two new cells. In this section an extension to this model will be investigated, allowing for more variation in the divisions of bacteria. The model will be an affine sum of two binomial distributions, as follows:

$$p(i, j) = p_{\text{ratio}} \text{Bin}\left(j, p_{\text{clump}}\right) + (1 - p_{\text{ratio}}) \text{Bin}\left(j, p_{\text{spread}}\right). \tag{5.2}$$

$p(i,j)$ defines the probability that an infected macrophage with i intracellular bacteria divides to produce a daughter cell with j intracellular bacteria. The hypothesis of the model, based on observations from the data, is that intracellular bacteria either reside clumped together within their host cell, or spread throughout. Consequently, when the host cell divides, in the case of clumped bacteria, the intracellular load will transfer en masse to only one of the daughter cells. Conversely, bacteria that are dispersed will be split between the two daughter cells. The structure of the intracellular population - either clumped together or spread throughout - is controlled by the parameter p_{ratio}. The remaining two parameters p_{clump} and p_{spread} define the probability of each intracellular bacteria moving to each daughter macrophage, resulting in the two binomial functions. Since there are two daughter cells produced during division $\sum_{i=0}^{j} p(i,j) = 2$.

Figure 5.6 shows how the likelihood function responds to each of the three parameters. The results mirror the results of previous chapters, with a slight increase in the proportion of infected-healthy pairs, rising from a probability of 0.7 to 0.86. The MLE values for p_{clump} and p_{spread} show there is very little middle ground, either the division results entirely in infected-healthy pair, or the division is approximately in half.

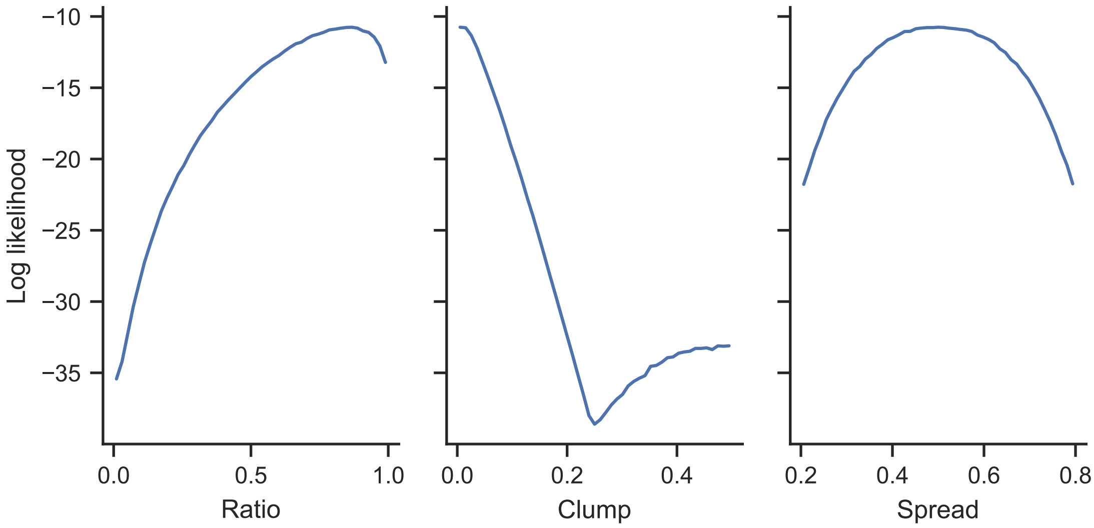

Figure 5.6: The maximum log likelihood as a function of each of the parameters governing the infected macrophage division process. $p_{\text{ratio}} = 0.86$ shows that most division events will be of the clumped form. $p_{\text{clump}} = 0.008$ shows that clumped divisions are dominated by completely infected-healthy divisions. Finally, $p_{\text{spread}} = 0.4941$ shows that spread divisions divide the intracellular load on average equally between the daughter cells.

5.3.5 Maximum number of intracellular bacteria

In the previous chapters the maximum number of intracellular bacteria was fixed at $N = 50$. Figure 5.7 shows the relationship between the time a macrophage spent infected and the level of intracellular load at the time of the cells death. This is also contrasted between cells that were born infected vs those that became infected. Cells that are born infected are more commonly born with a lower load and thus higher loads are rarely observed in cells that are infected for shorter times, however cells that become infected demonstrate some much higher loads with shorter infection durations. In both cases, the level of intracellular load is significantly higher than the previously assumed $N = 50$.

Figure 5.8 shows the model response to increasing the maximum number of intracellular bacteria all the way to the observed 600, where the effect is being measured

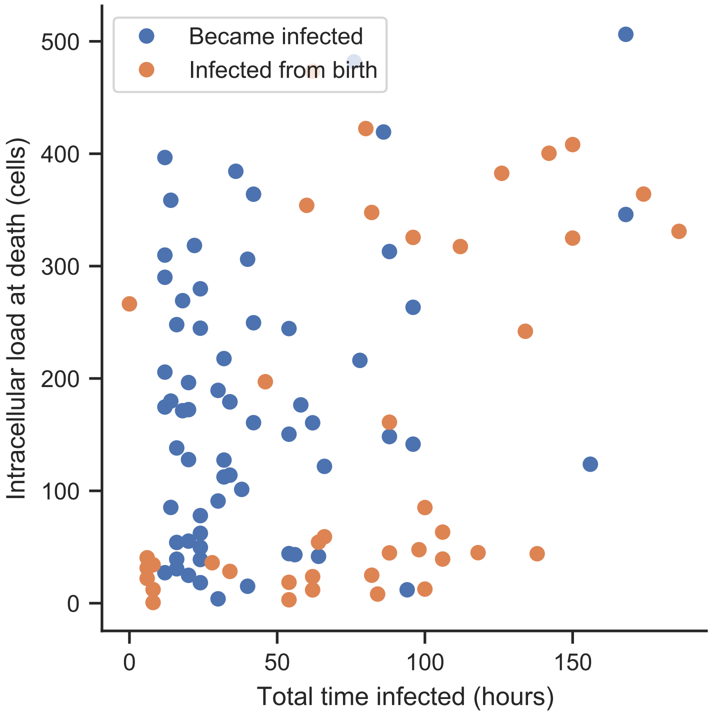

Figure 5.7: Data showing the relationship between the time spent infected and the intracellular load at the time of the cells death. There are a significant number of cells that contain more than the assumed maximum load. Cells infected from birth tend to be born with a small number of bacteria, resulting in higher loads more commonly being observed in cells that were infected for a long time. Cells that become infected have a more dispersed distribution, many dying with a high load only shortly after they became infected.

as

$$\frac{\log\left(E\left(200;N\right)\right)}{\log\left(E\left(200;50\right)\right)}. \tag{5.3}$$

$E\left(200;N\right)$ is the population of extracellular bacteria after simulating for 200 hours with the maximum intracellular bacteria fixed at N. This measure is chosen to demonstrate the proportional change in the response of the bacterial population to varying parameters.

There is a slight decreasing trend of approximately 2%, as would be expected. Increasing the maximum allowed load further (to 10^6) shows this reduction plateaus. The baseline for the comparison is calculated as the average of 5000 runs, while the compared values while ranging the parameter N are averaged over 1000 runs each. This reduction is due to computational limitations. The plateau suggests that macrophages die naturally, before bursting can take effect. This is shown in figure 5.8 (right).

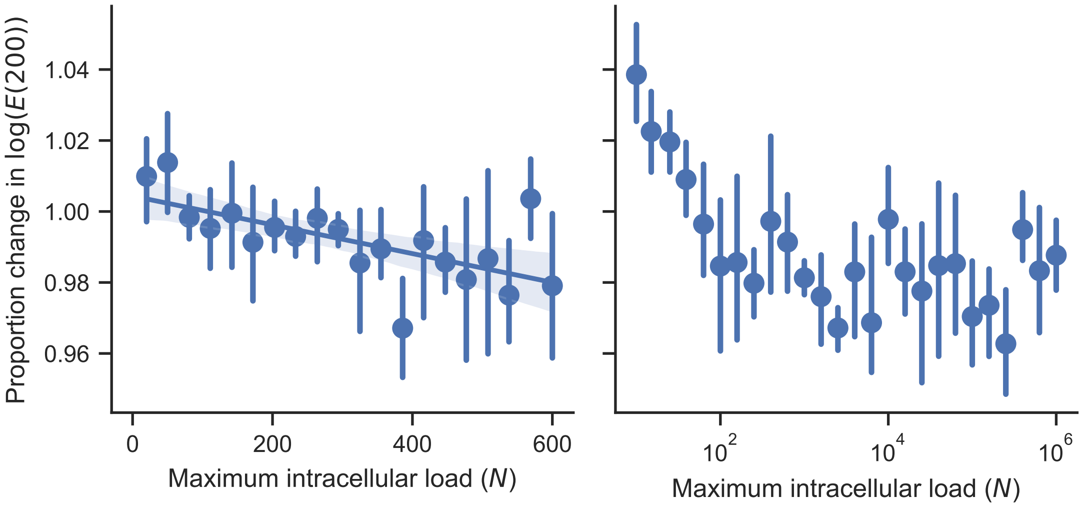

Figure 5.8: Measuring the effect of the maximum intracellular load on the resultant extracellular bacteria population (see equation 5.3). There is a very small decrease of approximately 2% when increasing the maximum intracellular load from the assumed 50 to the observed 600. The baseline is calculated from the average of 5000 runs, while the compared values are averaged from 1000 runs due to computational limitations. The error bars are a bootstrapped 95% confidence interval and the fitted line is a simple linear regression, again shown with bootstrapped 95% confidence interval. (**left**) Linear x scale up to the observed 600. (**right**) Log x scale increasing the limit beyond feasibility to observe change.

Parameter	**Description**	**Unit**	**Value**	**95% CI**
α_E	Extracellular growth rate	hour^{-1}	0.0471	$(-0.0057, 0.1104)$
α_I	Intracellular growth rate	hour^{-1}	0.0206	$(0.0038, 0.0392)$
β	Infection rate	cells^{-1}hour^{-1}	0.0017	$(0.0011, 0.0025)$
p_{partial}	Probability of partial aggregate uptake	scalar	0.5093	$(0.3600, 0.9467)$
a_{uptake}	Shape parameter for pdf of macrophage death post infection	cells	64.03	$(52.90, 75.16)$
g_H	Erlang growth rate for healthy macrophages	hour^{-1}	0.0230	$(0.0188, 0.0260)$
Ng_H	Erlang growth stages for healthy macrophages	scalar	2	$(2, 2)$
d_H	Erlang death rate for healthy macrophages	hour^{-1}	0.0146	$(0.0109, 0.0172)$
Nd_H	Erlang death stages for healthy macrophages	scalar	2	$(1, 2)$
g_I	Erlang growth rate for infected macrophages	hour^{-1}	0.0054	$(0.0040, 0.0070)$
Ng_I	Erlang growth stages for infected macrophages	scalar	2	$(1, 2)$
d_I	Erlang death rate for infected macrophages	hour^{-1}	0.0193	$(0.0171, 0.0217)$
Nd_I	Erlang death stages for infected macrophages	scalar	2	$(2, 2)$
p_{ratio}	Probability of clumped division event	scalar	0.8612	$(0.4881, 0.9906)$
p_{clump}	Binomial parameter for clumped division event	scalar	0.0082	$(0.0001, 0.0539)$
p_{spread}	Binomial parameter for dispersed division even	scalar	0.4941	$(0.3053, 0.6946)$
N	Maximum bacterial load	cells	50	$(20, 1000)$

Table 5.2: (On previous page) Complete and final list of model parameters along with descriptions, their fitted values, and 95% confidence intervals. All fitted values are fitted using maximum likelihood methods, and 95% confidence intervals are estimated using the likelihood surface generated.

5.4 Model definition

As in the previous chapter, the model is a stochastic ABM. Both macrophages M_i and bacteria aggregates A_i are modelled as agents. The only variable modelled for the aggregates is the total number of bacteria within the aggregate E_i. The variables modelled for the macrophages are the two counters modelling the Erlang growth and death g_i, d_i and the current intracellular load of the macrophage I_i.

Extracellular bacteria grow exponentially at a fixed rate α_E, thus for a given aggregate A_i this rate is scaled by the number of bacteria in the aggregate to $\alpha_E E_i$. The only way for the number of bacteria within an aggregate to reduce is through interaction with macrophages through infection. If the number of bacteria within an aggregate reaches zero, it is considered dead and removed from the system.

The macrophage birth death process is modelled exactly as in the previous chapter. The two Erlang ages g_i and d_i are tracked. g_i and d_i are incremented at rate $g_{H,I}$ and $d_{H,I}$ respectively depending on whether the macrophage is healthy H or infected I. If g_i reaches $N_{g(H,I)}$ then the macrophage will divide. Healthy macrophages are simply removed from the system and replaced by two new healthy macrophages, each with values $g_i = 1$ and $d_i = 1$. When an infected macrophage undergoes division, the intracellular load is divided between the two daughter cells according to the distribution defined in equation 5.2. This can result in either two infected daughters, or one infected daughter and one healthy daughter. If d_i reaches $N_{d(H,I)}$ then the macrophage dies, producing a new aggregate with their intracellular load.

Intracellular growth is again modelled as exponential growth. Thus for each macrophage M_i with I_i intracellular bacteria, intracellular growth events occur at rate $\alpha_I I_i$. When the intracellular load reaches the maximum load of N, the macrophage bursts. Thus this macrophage is flagged as dead and a new aggregate of extracellular bacteria is added to the system. The aggregate is initialised with N bacteria.

Infection events occur between healthy macrophages and extracellular bacteria aggregates according to mass action with rate parameter β. When an infection event occurs between aggregate A_i (with E_i extracellular bacteria) and macrophage M_i, the entire aggregate is internalised with probability p_{partial}. Otherwise the number of bacteria internalised by the macrophage is drawn from the uniform distribution as such: $I \sim \text{Uniform}[1, E_i - 1]$. Once the load to be internalised has been determined, the probability that the macrophage will survive the infection event is evaluated according to equation 5.1. If the macrophage survives then the bacteria are internalised, whereas if the macrophage does not survive then the bacteria remain in the aggregate and the macrophage is flagged as dead.

By construction, each event and its associated rate of occurrence is now only dependent on the current state of the system, thus the model is Markovian. With the rates and events now defined, the model can be simulated using the Gillespie algorithm. At each time step the system is evaluated and a list of events is compiled with corresponding rates and functions to apply if that event occurs. This list is then sampled in order to determine which event occurs next and at what time, resulting in a continuous time Markov chain (CTMC).

5.5 Model results

In this section a collection of results generated from the model will be presented. First the qualitative output of the model as a time series will be analysed, followed by a range of distributions and how they compare with the analogous distributions in the data. As in previous chapters, a Sobol sensitivity analysis will be performed and used to inform which mechanisms to explore through computational experiments. Finally the computational experiments will be performed to establish the relationships between methods of control of intracellular growth and the rate of phagocytosis.

5.5.1 Time series

The model is run for 200 simulated hours in line with the length of the experiments. Initial conditions for the model are two macrophages and a single bacteria which alternates between being initialised intracellularly and extracellularly. To mimic the experimental design would be to initialise without extracellular bacteria, however there are results in which macrophage infection events occur prior to any infected macrophage death events, thus extracellular bacteria must have been present.

Figure 5.9 shows the average of 6000 runs of the simulation with parameters fixed at their MLE values. The standard deviation is also shown to demonstrate the variability between runs. The macrophage population is in general able to sustain itself, however it is unable to contain the infection, with the extracellular bacteria growing exponentially, in relatively few distinct aggregates. This is due to the fitted infection rate still being relatively low. For three macrophages and a single aggregate, it is expected to take approximately 200 hours to observe a single infection event.

The core function of the macrophages is not to ultimately eradicate the invading bacteria, but rather to delay their growth prior to the adaptive immune system being activated. Section 5.5.4 investigates the way in which the varying mechanisms contribute to delaying or enhancing the bacteria population growth.

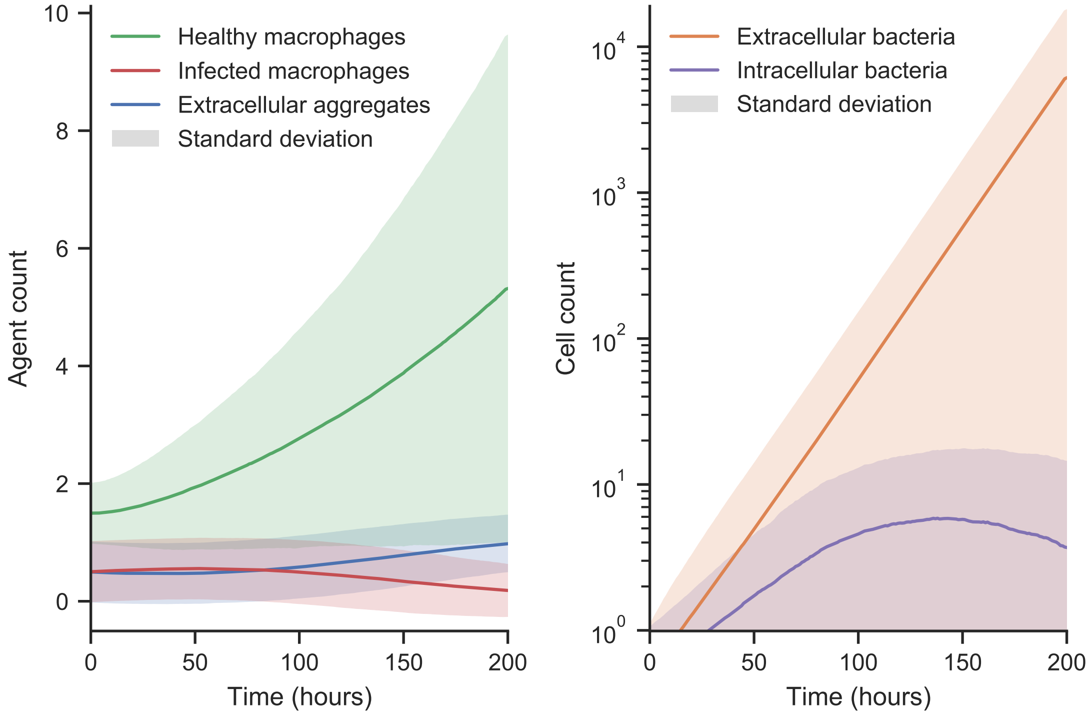

Figure 5.9: 6000 realisations of the model with parameters kept at their MLEs. Solid lines show the mean of the simulations and the shaded regions show the standard deviation. At 200 simulated hours, the healthy macrophages have sustained their population, but are unable to control the infection and extracellular bacteria are growing exponentially.

5.5.2 Validation

In this section a range of distributions will be taken from both the experimental data and the simulated model. By comparing like-for-like distributions, it is possible to evaluate how successfully the model is capturing the qualitative nature of the real-world system. The statistics chosen broadly represent the key mechanisms being modelled: the birth death process of macrophages; the infection process; and the intracellular bacteria lifecycle.

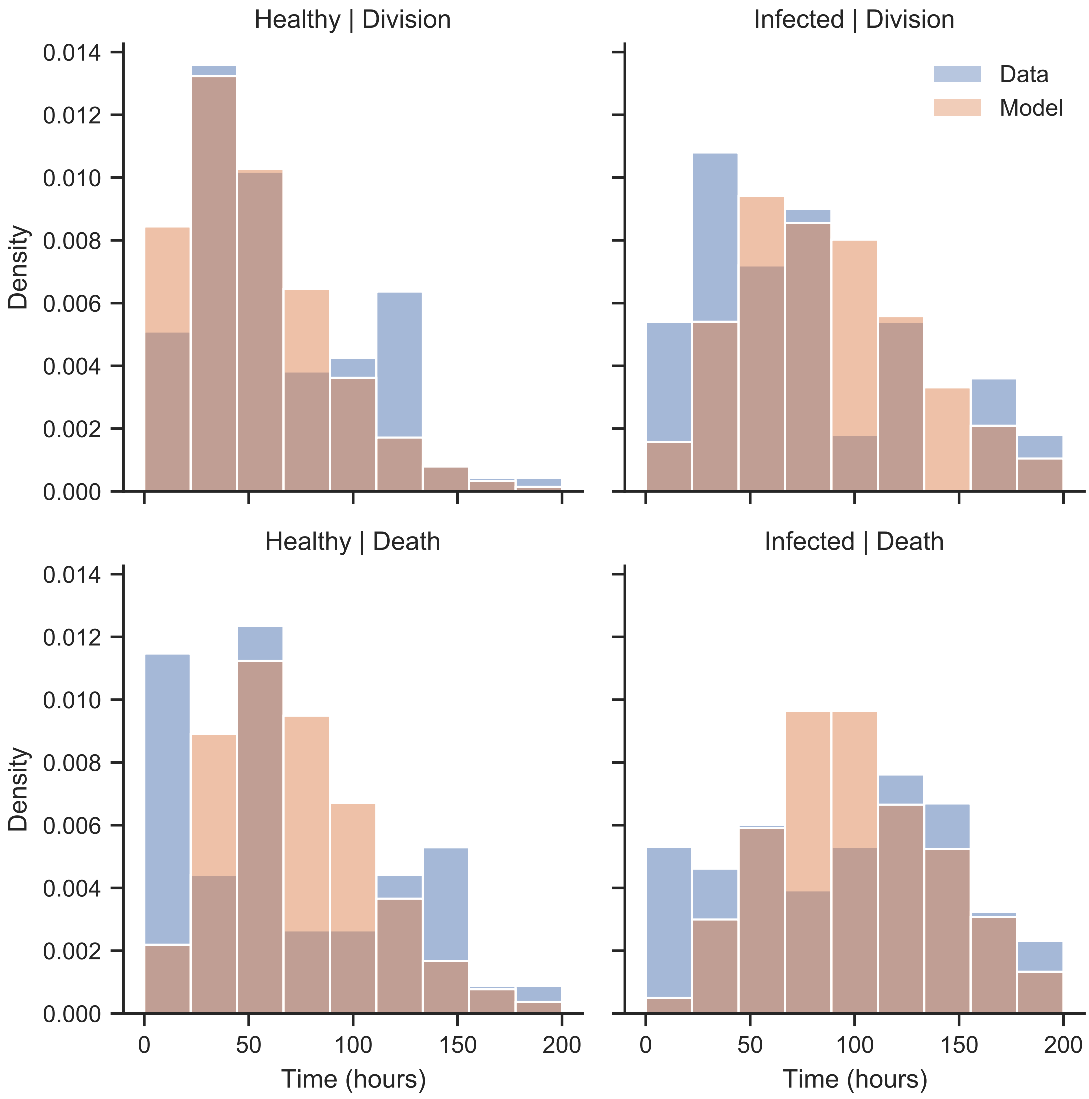

Figure 5.10: Comparison of distribution of wait times for macrophage birth death processes between the model and the data. The division process fits well, while the death process has a flatter distribution in the data than the model, especially for infected macrophages.

Macrophage birth death process

The model for the macrophage birth death process is a CTMC comprised of two Erlang processes running in parallel, as described in section 4.6. Figure 4.9 showed that this model fits the data well when considered in isolation from the other mechanisms involved in the full system. Figure 5.10 is the analogous plot showing the distribution of wait times for birth and death events in the full model, and how they compare to the data.

The model provides a good qualitative fit to the data, in particular for the division processes. The infected death wait times have a flatter distribution in the data than in the model, however the median of the distributions remain close to each other at approximately 100 hours.

Infection wait times

The distribution of times from a macrophage being born to a macrophage becoming infected are shown in figure 5.11 for both the model and the data. The model is underestimating the distribution of ages at which macrophages become infected. This is a symptom of the model being homogeneously mixed. From the start of the simulation, all macrophages are interacting with all possible aggregates, however in the experiments macrophages move around to seek out bacteria. Thus there is a delay, and hence older macrophages are observed to become infected in the data, but not in the model. Reducing the infection rate does not correct for this however, as it merely changes the observed number of infection events, rather than the distribution of ages at which macrophages become infected.

Intracellular loads

The data measuring intracellular load at the time of death of the macrophage is not used to fit the maximum intracellular load used in the model. This parameter has been defined based on values in the literature. This is because there is a large

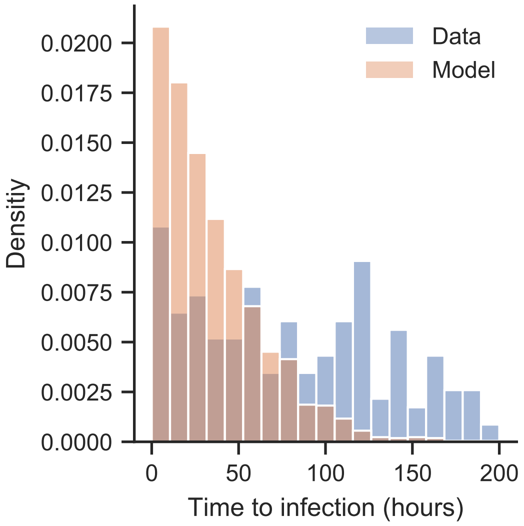

Figure 5.11: Distribution of wait times from a macrophage being born to that macrophage becoming infected. The model skews to the left when compared to the flatter distribution observed in the data. This is likely a result of the homogeneous mixing assumption of a non-spatial modelling. Since all agents are always able to interact, interactions are modelled as a Poisson process thus having exponential wait times.

discrepency between the value observed in the data (up to 600) compared to the maximum value found in the biological literature (approximately 120 [118]) and the maximum value used in the modelling literature (approximately 100 [157]). Section 5.3.5 demonstrated that changing the maximum load used in the model has only a small effect on the outcome. Thus, here the maximum intracellular load is extended to 600 to allow the distribution of intracellular loads generated by the model to be compared to those observed in the data. Figure 5.12 shows the distribution of intracellular loads at the time of infection and the time of death of infected macrophages, comparing the model to the data. The load at infection fits the data well, only slightly underestimating the initial load. Similar to the distribution of infection times, this is likely due to the homogeneous mixing and hence premature interaction between macrophages and aggregates.

The distribution of the load at the time of death of the infected macrophages does

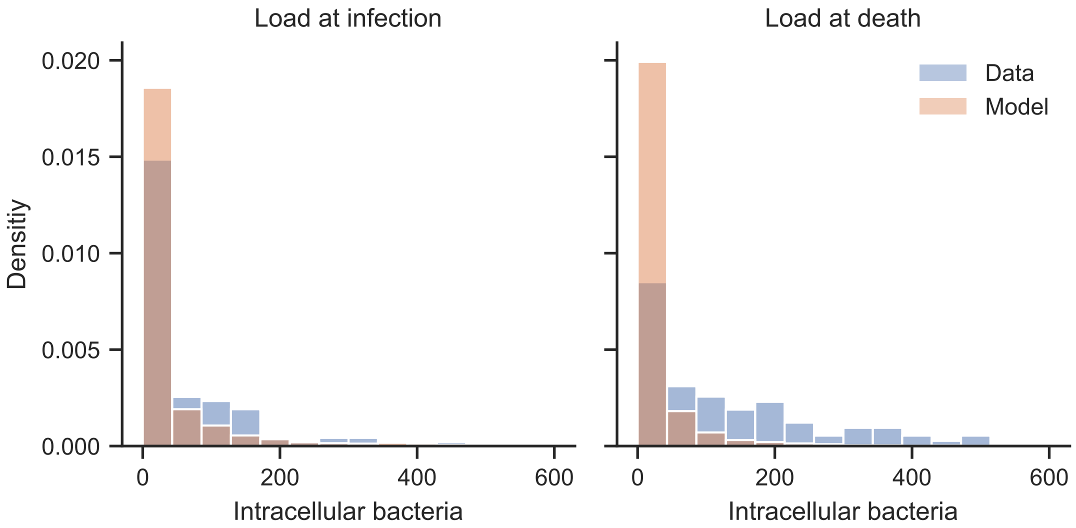

Figure 5.12: The intracellular load at the time of infection and the time of death is measured and compared. (**left**) The distribution of loads at the time of infection are close, with the model slightly underestimating, but still able to capture the tail. (**right**) The distribution of the intracellular load at the time of death is significantly heavier in the tail than the data.

not fit well, the loads are significantly underestimated compared to the data. This implies that infected macrophages in the model are dying prematurely. The biological literature refers to a process whereby bacteria are able to delay apoptosis of the infected host cell [127, 128, 130]. Allowing for separate death rates between healthy and infected macrophages was intended to capture this, however the MLE for the two groups was very similar and it was not possible to fit a load dependent death to the data (see appendix C.1).

Goodness of fit tests

Table 5.3 shows the results of the goodness of fit tests outlined in section 1.6.6 applied to the distributions presented in the previous 3 sections. Both the Kolmogorov-Smirnov (KS) and the Anderson-Darling (AD) tests are applied to each distribution

	Kolmogorov-Smirnov		Anderson-Darling	
	t-stat	$c(5\%)$	t-stat	$c(5\%)$
Healthy division	0.1884	0.1320	5.8831	1.9610
Healthy death	0.2241	0.1905	8.0976	1.9610
Infected division	0.2788	0.2770	2.2709	1.9610
Infected death	0.1598	0.1049	14.5264	1.9610
Infection wait times	0.4060	0.0937	142.2180	1.9610
Load at infection	0.3383	0.1310	78.8970	1.9610
Load at death	0.6245	0.1071	267.8059	1.9610

Table 5.3: Test statistics and the corresponding critical values at the 5% significance level for the KS test and the AD test for the range of distributions presented in this section.

with the null hypothesis that the results from the model and the observed data are drawn from the same distribution. The only result for which the test statistic is close to the critical value is the wait times on division of infected macrophages. Despite the KS test statistic being only slightly higher than the critical value, the p-value for this test was $0.0382 < 0.05$ thus the null hypothesis can still be rejected.

While the goodness of fit tests demonstrate that the extracted distributions from the model are not quantitatively the same as the observed data, the qualitative behaviour of the system remains in line with the experiments. The model presented is a mechanistic model and has been developed based on current understanding of the biological dynamics. While a statistical model may have performed better in the goodness of fit tests, the qualitative similarities of the mechanistic model to the data provide a greater understanding of the underlying processes.

5.5.3 Sensitivity analysis

As for previous chapters, two versions of the Sobol sensitivity analysis are applied. Firstly, variation is measured in the total populations of extracellular bacteria,

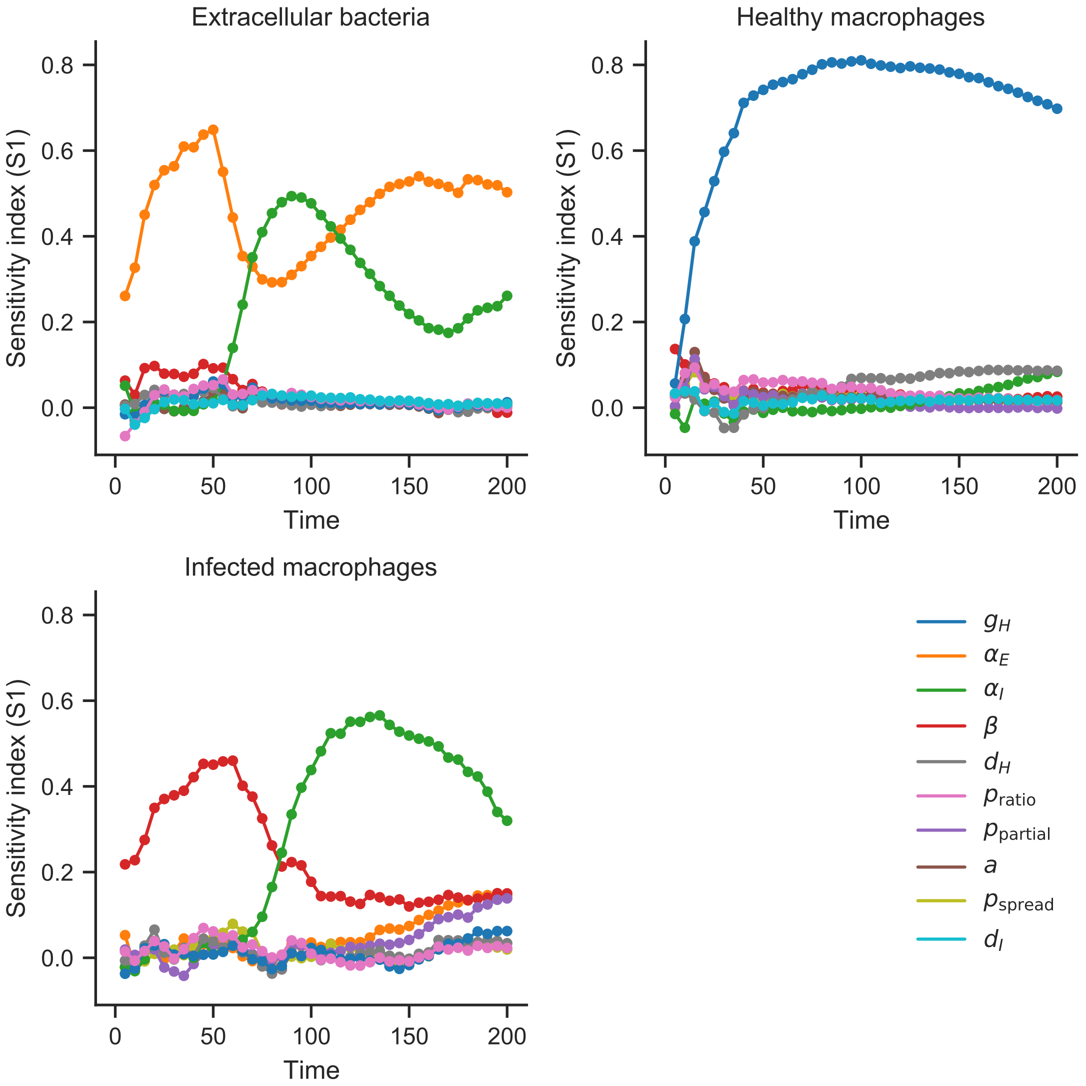

Figure 5.13: Sobol sensitivity analysis measuring the fraction of the variance in the modelled populations over time that can be attributed to variance in the given parameters.

healthy macrophages, and infected macrophages, sampled every 5 hours up to a maximum of for 200 hours. This is shown in figure 5.13. Secondly, variation in the time taken for the extracellular population to reach 1000 cells is measured, shown in figure 5.14. The initial conditions for the sensitivity analysis are ten macrophages and a single bacteria which alternates between beginning intracellularly and extracellular. The parameters are allowed to vary within their 95% confidence intervals shown in table 5.2. For the 12 parameters, 1000 parameter sets are sampled using the Saltelli sampler, as described in section 1.6.5. Since the model is stochastic, each set of parameters is evaluated over a large number of model simulations, using convergence of the mean to determine when to stop simulating.

The majority of variance in the population of extracellular bacteria is a result of variance in the extracellular growth rates (α_E) and the intracellular growth rate (α_I). Initially, α_E dominates, however at around 100 hours, the α_I begins to increase. This is in line with the growth of intracellular bacteria as seen in figure 5.9. However, this relationship does not last long, and the extracellular growth takes over again as the extracellular bacteria are free to grow exponentially and not constrained by remaining intracellular.

Healthy macrophages are again driven predominantly by their own birth death process, however two parameters of the infected macrophage division process are also present, showing that the production of healthy macrophages from healthy-infected division events is not insignificant in the system.

The initial stage of the model sees the infected macrophage population being controlled mostly by the rate of infection (β). However as the intracellular population is allowed time to grow, the driver changes to the intracellular growth rate. This demonstrates the effect of the intracellular pressure rupturing infected macrophages on the macrophage population.

As in previous chapters, the time until extracellular bacteria form the majority of the system is also investigated. The sensitivity analysis is shown in figure 5.14. Intracellular growth rate has a clear relationship to the extracellular population, as

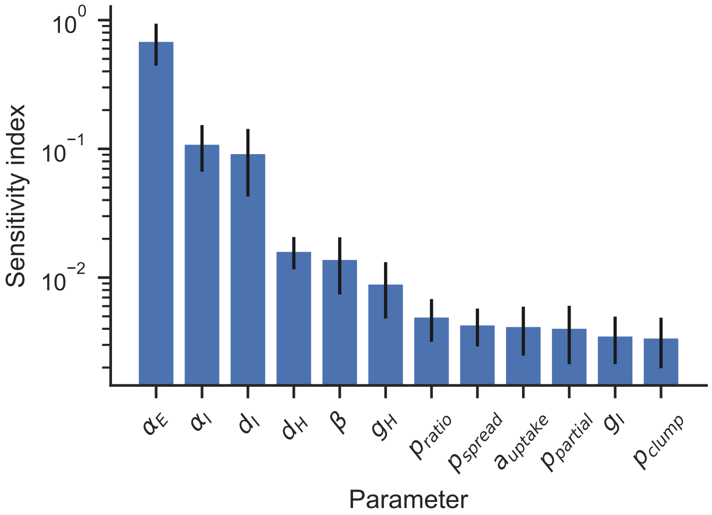

Figure 5.14: Sobol sensitivity analysis tracking the variation in the number of simulated hours required for the extracellular population to exceed 1000 cells. The vast majority of variation is attributed to the extracellular growth rate, with the intracellular growth rate and the infected macrophage death rate having the next highest contributions.

the faster bacteria can grow intracellularly, the more the extracellular population will increase by when the host macrophage dies. In this iteration of the model, the extracellular bacteria growth rate is approximately double that of the intracellular rate, therefore increasing the rate at which infected macrophages die will result in bacteria spending a higher proportion of time extracellularly, and thus able to grow faster. Increasing the rate at which healthy macrophages die will reduce the number of macrophages available to internalise bacteria. Similarly, reducing the infection rate reduces the number of infection events, both resulting in bacteria spending more time extracellularly and growing faster.

5.5.4 Experiments

In this section two computational experiments will be performed, analysing the response of the model to variation in parameters governing key mechanisms. First the intracellular growth rate will be varied against the infection rate demonstrating a non-linear relationship, presenting an advantage to the bacteria over the macrophages. Second, an additional mechanism will be considered, the possibility for macrophages to inhibit intracellular growth. The probability of growth inhibition will again be considered against the varied infection rate showing that a large increase in this probability is required to achieve the same result as a smaller increase in the infection rate.

Intracellular growth rate against rate of infection

Two key mechanisms affecting the macrophages' ability to control the bacteria population are the intracellular growth rate (α_I) and the rate of phagocytosis (β). As shown in the sensitivity analysis, the intracellular growth rate (α_I) and the rate of infection (β) are significant parameters in the development of the system. Figure 5.15 shows the proportional change in $\log(E)$ compared to the unmodified model:

$$\frac{\log\left(E(200;\alpha_I,\beta\right)}{\log\left(E(200;\alpha_I^\star,\beta^\star\right)}. \tag{5.4}$$

$E(200;\alpha_I,\beta)$ is the population of extracellular bacteria after 200 hours given parameters α_I and β. $\alpha_I^\star$ and $\beta^\star$ are the previously calculated MLE values. For each pair of (α_I,β), the model is run multiple times and the measure of interest, in this case the population of extracellular bacteria after 200 hours, is averaged across realisations. The parameters were allowed to vary within limits proportional the their MLE values: approximately $\pm 50\%$.

It is not a surprising result that as the rate of infection increases and the intracellular growth rate decreases, the resultant population of bacteria is reduced. This is due to bacteria spending more time intracellularly with a slower growth rate. There is however a non linear relationship between the two parameters. As the rate of intracellular growth increases, a larger proportional increase in the infection rate is required to offset the resultant increase in total bacteria. This demonstrates the inherent advantage the bacteria have over the macrophages.

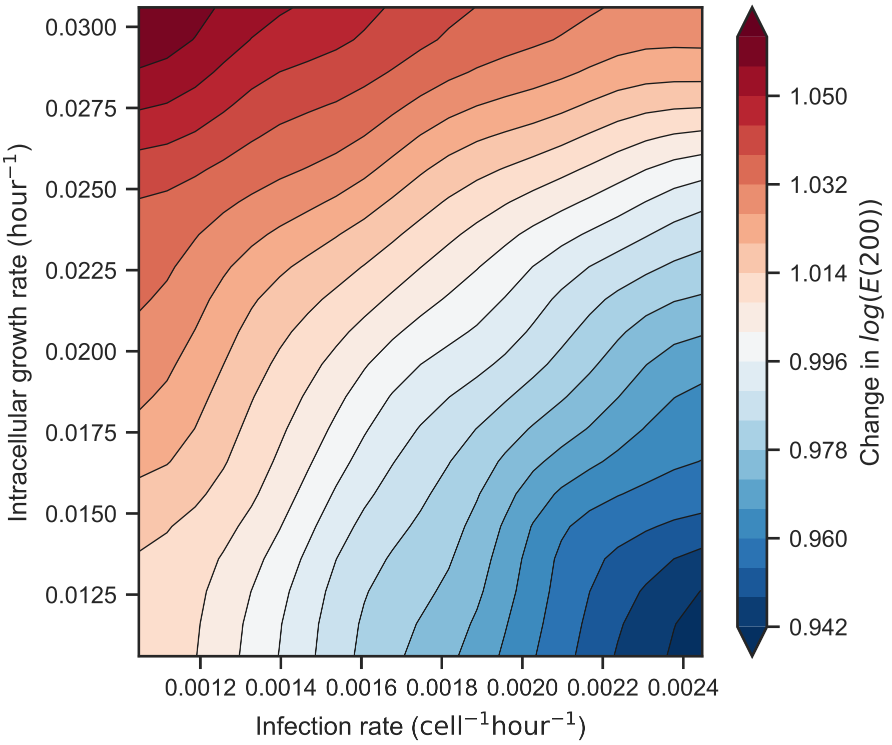

Figure 5.15: The intracellular growth rate and the rate of infection are allowed to vary within approximately ±50% of their MLE values. The resultant extracellular population after 200 simulated hours is compared to the base model as a ratio of logs (see equation 5.4). An increase to the intracellular growth rate requires a larger proportional increase in the rate of infection to offset the increased rate of bacteria growth.

Probabiliy of inhibiting intracellular growth against rate of infection

The biological literature references the ability for activated macrophages to inhibit the growth of intracellular bacteria [104]. While it was not possible to fit this mechanism to the data despite observing varied intracellular growth rates, this mechanism can be investigated with the model. An additional parameter is included: p_{inhib}. When a healthy macrophage interacts with an extracellular aggregate, either all or some fraction of the bacteria are internalised. The macrophage then has some probability

(p_{inhib}) that it will inhibit the growth of these bacteria. This status will remain until the macrophage dies and releases the bacteria back into the extracellular space.

Figure 5.16 shows the response from the model when allowing this parameter to range fully between 0 and 1, against the rate of infection. The rate of infection, β, is once again allowed to vary within ±50% of its MLE. Note the y-axis is flipped to mirror figure 5.15. The effect of a minority of macrophages being able to completely stop intracellular growth compared to all macrophages having a reduced intracellular growth is drastically reduced. The vertical nature of the heat-map shows that the rate of infection holds significantly more power in affecting the outcome of the model compared to the probability of growth inhibition, especially given that β is only allowed to range with a limited domain, whereas p_{inhib} is permitted to vary across its entire domain.

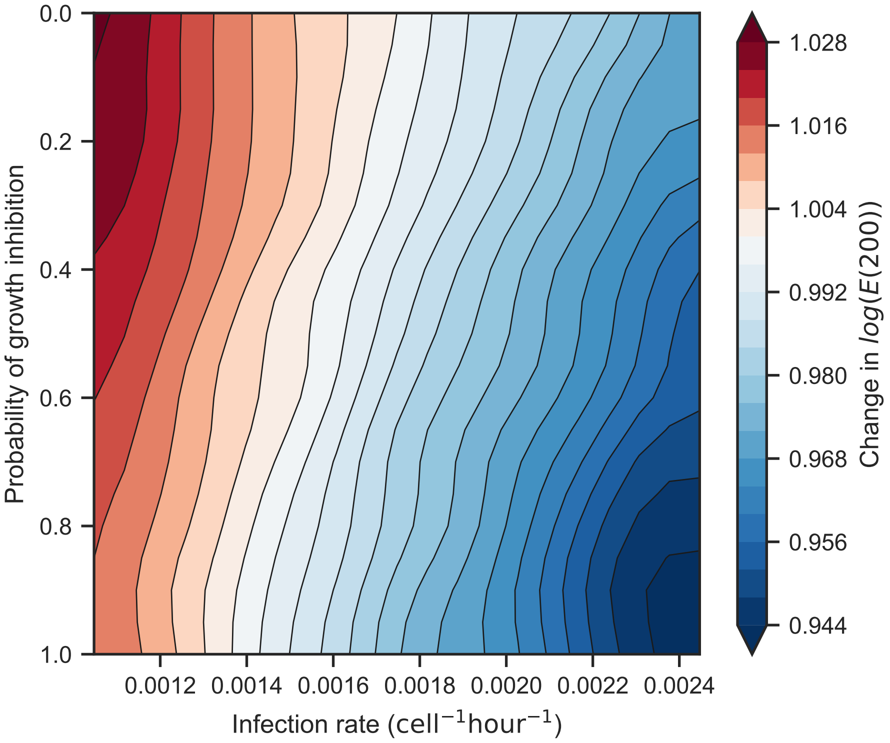

Figure 5.16: A new mechanism is introduced to allow macrophages to inhibit any intracellular growth according to a fixed probability. This probability and the rate of infection are allowed to vary: the probability of growth inhibition varies completely between 0 and 1, while the rate of infection varies within ±50% of its MLE. Note the y-axis is flipped to mirror figure 5.15.

5.6 Conclusions

In this chapter the model has been undergone a fundamental change in the way extracellular bacteria are modelled. It has transitioned from a single population bacteria all of which are interacting with macrophages, to distinct populations of bacteria (aggregates) which interact with macrophages at the macro-scale. Additional mechanisms were included, such as the probabilistic division of intracellular loads during infected macrophage division events and the probabilistic uptake of aggregated bacteria. The parameters for these new mechanisms were fitted using maximum likelihood methods, demonstrating smooth likelihood surfaces to be optimised over.

The model presented in this chapter is the final model to be developed in this thesis. Through an iterative process of model design and evaluation against the underlying dataset, a novel model has been constructed and parametrised. By comparing biologically relevant distributions in the model and the data, the model was validated both qualitatively and quantitatively. This process demonstrated the trade off between model accuracy and model complexity. Since a key goal of this thesis was to develop a simple model based on biologically relevant mechanisms, model simplicity and explainability justified the drop in quantitative accuracy. For example, while the birth death process of the macrophages replicate the data well, the wait times on time to infection demonstrated a left skew. This is due to the non-spatial aspect of the model. Macrophages do not move completely randomly and therefore a fixed rate of interaction is not able to capture the dynamics of macrophages actively moving towards aggregates.

Performing a sensitivity analysis identified the key mechanisms involved in the output of the model. Naturally the bacteria growth rates contribute significantly to the total population of bacteria, similarly to the previous model. However now that the rate of infection is not tied to the total number of bacteria, rather the total number of aggregates, the rate of infection also plays a significant role in the early stages of the model evaluation. Since the extracellular growth rate is tied to the virulence of the bacteria, it was not focused on for methods of control. Instead, the intracellular growth rate and the rate of phagocytosis were investigated.

The first computational experiment compared the resultant population of extracellular bacteria when running the model with different combinations of parameters for intracellular growth rate and the rate of infection. What this experiment revealed was that a proportionally larger increase in the infection rate is required in order to offset the effects of an increase in the rate of intracellular growth. This highlights the advantage that the Mtb bacteria have over the innate macrophages. Naturally the optimal result was achieved through high infection rates and low intracellular growth rates, and the inverse produced the worst results.

The second computational experiment performed measured again the resultant population of extracellular bacteria when allowing the rate of infection to vary alongside the probability of growth inhibition. The near vertical strata in the results of this experiment demonstrate the much higher affinity for the rate of infection to have an effect than the probability of growth inhibition. Thus an overall reduction in the intracellular growth rate is a significantly better method of bacteria control than the complete inhibition of intracellular growth in a subset of macrophages.

Further developments of the modelling work presented in this thesis should include the addition of a spatial element. By modelling macrophage movement as well as a potential proxy of the chemical signalling by macrophages, the distribution of wait times of time to infection will more accurately represent those observed in the data. In the following chapter a full discussion of the work presented here will be presented as well as a summary of the thesis outcomes.

Chapter 6

Discussion

The core goal of this thesis was to develop a mathematical model of the in vitro interaction between innate alveolar macrophages and *Mycobacterium tuberculosis* (Mtb) bacteria. The model development process was an organic and iterative progression whereby additional data and additional mechanisms were incorporated in each chapter. The iterative nature of the model development and the regular collaboration with the experimentalist has meant that feedback from the model development process was used to inform on additional data of interest and resulted in additional data being collected. A particular focus was to maintain biological explainability of the modelled mechanisms and the mathematical structure used. The subsequent goals of the thesis were to use the models to identify the dominant mechanisms affecting the outcome of the system, and to investigate methods of controlling the bacteria population through computational experiment. Dominant mechanisms were identified through a combination of mathematical and sensitivity analysis. These were then investigated and theoretical scenarios were explored through simulation and computational experiments.

The initial model presented in chapter 2 was a continuous and deterministic ordinary differential equation (ODE) model. Intracellular bacteria were included implicitly in the infected macrophage (M_I) class. It was shown that the assumption of an average number of intracellular bacteria was not sufficient and skewed the population of extracellular bacteria through the rate of death of infected macrophages term. In the second model, presented in chapter 3, the intracellular bacteria were modelled explicitly through additional compartments in the model. Additionally, the birth death process of the macrophages was extended to include the Erlang distribution for the growth rates. This was shown to be a better fit to the data, as well as with the current biological understanding of the cell life cycle. The next stage of the model development process, presented in chapter 4, transitioned the model from the continuous and deterministic ODE framework to a discrete and stochastic agent-based model (ABM) framework, as ABMs are better able to model more complex mechanisms, such as the infected macrophage division process. The macrophage birth death process was further extended to improve the qualitative fit to the available data by modelling death rates with an Erlang distribution. In the final iteration of the model, presented in chapter 5, the extracellular bacteria underwent a paradigm shift, each population of extracellular bacteria now being considered as distinct agents, named here as aggregates. As an ABM, it was possible to include additional mechanisms such as the partial uptake of an aggregate by a healthy macrophage, and the division of intracellular bacteria according to the affine sum of two binomial distributions.

Sensitivity analysis of the developed models demonstrated that the macrophage dynamics dominate the system early on, however in the later stages it is the growth rates of the bacteria that drive the state of the system. The measure used in the sensitivity analysis was the total population of each cell type throughout the duration of the model simulation. An alternative measure was also used, measuring how much time was required before the rate of change of the population of extracellular bacteria was positive or until the extracellular bacteria population exceeded a defined limit. This second measure showed the rate of infection to be the dominant mechanism, however the relationship was shown to be non-monotonic. That is, there is an optimal value for the rate of infection at which the delay on extracellular bacteria growth is maximal. Increasing the rate of infection beyond this optimal value worked to reduce the delay. This delay is of interest as it is hypothesised that maximising the time until extracellular bacteria are able to grow freely may have a beneficial impact on disease progression for the host. The result presented here could be tested experimentally. Since the rate of infection amounts to the rate at which moving macrophages locate a population of bacteria, this could be controlled by measuring the initial distance between the macrophage and the extracellular bacteria. The hypothesis would then be that the relationship between this distance and the total extracellular bacteria after a fixed amount of time is non-monotonic.

Mechanisms of bacteria control were investigated through computational experiments. The rate of infection and the rate of intracellular growth were frequently measured as significant parameters by the sensitivity analyses and since both of these mechanisms involve both bacteria and macrophages, these were targeted for investigation. A theoretical mechanism was introduced, whereby macrophages have the opportunity to inhibit intracellular growth, and this was compared to a universal reduction in intracellular growth rate for all macrophages. It was shown that the smaller reduction across all macrophages had a larger impact on reducing the final population of bacteria than the complete inhibition of intracellular growth in a smaller population of macrophages. As for the previous result, a possible experiment to test this result is as follows. Two populations of macrophages must be cultivated, one of resting macrophages, the other of activated macrophages. Then a series of experiments can be performed with varying ratios of cells from each population, signifying the modelled probability of a macrophage inhibiting intracellular growth. The ratio of activated to resting macrophages can then be compared to the resultant bacteria population with the hypothesis of their being a correlation.

6.1 Literature review

A literature review of both the biological and the mathematical literature provided a foundation of concepts, mechanisms, and methodologies relevant to developing mathematical models of Mtb. While the development of the model was primarily driven by observations within the data, the literature provided motivating examples and supplemented this process. The approach used to identify resources was based on the method of systematic review, whereby a specified search term is used to identify a large quantity of literature. Through a series of inclusion criteria, this body of literature was reduced to identify relevant materials. Finally, additional sources were found through review articles identified within the primary search.

As described in section 1.7, the literature search was focused on resources specific to Mtb, for which a large subset was focused on the processes based around granulomas. Since Mtb is a widely studied disease, the large number of materials from a range of journals provided a good foundation of knowledge of the biology and mathematical modelling of Mtb infection and the immune response to date. The study of within-host bacterial infections, either in vitro or in vivo, is a vast and ever growing area for which each bacterial species will have specific mechanisms necessitating specific models. For this reason the literature review remained focused on Mtb, although work on other bacteria such as Francisella tularensis demonstrates some similarities [176, 184, 186].

6.2 Non spatial model

The literature examined focuses in the majority on the formation of granulomas and the interactions relating to the adaptive immune response, often considering the variability in outcome and how that corresponds to the strong heterogeneity in host responses to Mtb exposure. Since the granuloma is inherently a spatial object, this results in most models in the literature having some form of spatial element, whether that is agents being modelled on a lattice, or a spatial dependency in a set of partial differential equations (PDEs) [136, 148, 149, 151, 153, 157, 166]. The focus of this study was the in vitro innate immune response, with a goal of minimising complexity while allowing key mechanisms to be identified. For simplicity it was decided to not include a spatial element, thus assuming homogeneous mixing between macrophages and bacteria and modelling the infection process using mass action. Additionally, the data used to parametrise the model does not contain sufficient detail to fit the movement of macrophages. In multiscale systems, macrophages are often modelled to follow cytokine concentrations released by other macrophages that have encountered bacteria such as interferon-γ (IFNγ) and interleukin-10 (IL-10).

This process would have to be estimated using parameters from the literature. The biggest impact of not including a spatial element was on the infection process. The assumption of homogeneous mixing resulted in elevating the force of infection, thus overestimating the rate at which macrophages become infected. This is discussed more in the following section. The benefit of parametrising a non-spatial model is the vast reduction in the number of parameters that are required and, consequently, the number of assumptions that must be made about how particles in the system are moving. The resultant model is demonstrated to capture the qualitative dynamics well and the reduced complexity of the model results allows for a larger number of computational experiments to be run, increasing the confidence in the resulting outcome of the computational experiment. Despite not being the primary focus of the experiments, the raw image data that was captured during the experiments could, in the future, be used to analyse the movements of macrophages and fit a spatial model, whilst being cautious of the assumptions that doing so would require. By including a spatial element, the distribution of wait times in the infection process can be flattened and will have a better qualitative fit to the distribution of wait times shown in figure 5.2.

6.3 Infection process

There are three core points of discussion for the infection process. Firstly, the extracellular population during infection events is inferred using the previously fitted extracellular growth rate. Secondly, as discussed in the previous section, the model assumes homogeneous mixing between macrophages and bacteria. And finally, in chapter 5, the probability for a macrophage to partially consume a bacterial aggregate is inferred using a bootstrapping process.

The infection process is fitted to the measured time of birth and time of infection of macrophages. The assumption of the model is that the force of infection felt by the macrophage during this time is proportional to the level of extracellular bacteria. Under a mass action assumption the rate of infection events is proportional to the population of healthy macrophages and the population of bacteria, $\beta M_H E$. Therefore the force of infection felt by the healthy macrophages is βE. In chapters 2, 3 and 4 this is the total extracellular population, while in chapter 5 this is the number of distinct aggregates. The bacteria population during infection events is not tracked, thus these populations are estimated using the fitted growth rate of the extracellular bacteria and the assumption that the system is initialised with a single extracellular bacteria. In the experimental design, it is stated that the macrophages are infected with a multiplicity of infection of 1:1, and are washed prior to initialising the experiment to remove any extracellular bacteria. This implies that the model should assume no extracellular bacteria at the start of the experiment. However, there are observations within the data demonstrating macrophages that become infected prior to the release of any intracellular bacteria. The release of intracellular bacteria is assumed to occur only after the death of an infected macrophage; bacteria are assumed to not be able to escape a still living macrophage. Since the model assumes the rate of infection is proportional to the number of extracellular bacteria, these infection events are evaluated as having a likelihood of zero for all possible parameters. This is resolved by assuming an initial population of extracellular bacteria. A consequence is that the population of extracellular bacteria is being overestimated, which would result in a higher force of infection being modelled than is observed, resulting in a lower than expected rate of infection.

Whilst validating the model, it was shown that the infection process underestimates the time at which macrophages become infected; that is, macrophages become infected earlier in the model than observed in the data. The initial conclusion is that the infection rate is being fitted too high, which contrasts against the argument that assuming an initial extracellular population would result in a lower than expected rate of infection. However, adjusting the rate of infection either up or down does not affect the resultant distribution of infection times. The source of the problem is the assumption of homogeneous mixing. The core assumption of mass action and homogeneous mixing is that particles move randomly, and thus interact at a constant and fixed rate. However, macrophages are known to follow chemical signals released by other macrophages that have encountered bacteria such as IFNγ and IL-10, and thus the interactions are not an entirely random process. This causes the underestimate of infection times, as macrophages permanently feel a force of infection, which grows exponentially over the course of the simulation. Thus, the observation of a macrophage that exists for a long time prior to becoming infected drives the infection rate parameter down, as a high rate of infection would evaluate to very low likelihood. Future work should investigate alternative methods for modelling the infection process, such as a delay mechanic or using a Michaelis-Menten model. The impact of fitting a spatial model is discussed in the previous section.

Finally, the process of macrophages only partially consuming bacteria aggregates is a mechanism that is observed in the image data of the experiment, however the process was not quantitatively measured, as it was not a focus of the experiment. Thus the process was parametrised by bootstrapping the possible aggregate populations and evaluating a likelihood function. Evaluating the likelihood function required large number of repetitions due to the variation in the model output, and the resultant likelihood plane was not very steep, meaning uncertainty in the model parameters is high. Additionally, the sensitivity analysis demonstrated low sensitivity to the parameter governing the probability of partial uptake. This shows that the impact of this mechanism is low and will not change the overall conclusions of the model.

In summary, firstly, the extracellular bacteria was assumed to exist from the start of the simulation, and inferred using the previously fitted extracellular growth rate. The impact of this is an overestimate of the extracellular population and thus an underestimate of the rate of infection. However, in the model, macrophages become infected earlier than in the data. Future work could explicitly track the extracellular population in order to limit the error resulting from the assumed population levels. Secondly, the model is non-spatial and assumes mass action mixing in order to minimise complexity in the model and reduce the number of parameters that require fitting. This is the cause of the underestimation of the infection times. Future work should investigate the possibility of fitting a spatial model, and how it affects the infection process. Finally, the probability of partial uptake is inferred using a bootstrapping process. The likelihood function for this parameter, and the results of the sensitivity analysis demonstrate that it has a small impact on the outcome of the model. In order to investigate the role of uptake, data should be collected which quantitatively measure this.

6.4 Macrophage division

Macrophage division is not a process that is included in models in the mathematical literature. Instead, macrophage populations are assumed to be maintained by recruitment of new cells. This is a result of the majority of mathematical literature constructing in vivo models directly. While biological literature does suggest macrophage populations are maintained through proliferation in addition to recruitment [89–92], it is also shown that during infection the recruitment is the dominant form [101]. Macrophage division was included in this study as it was observed commonly within the data set, and the purpose of this study was to recreate the dynamics observed within this data set. Should future work consider translating this model from in vitro to in vivo, it should be noted that the macrophage division will not translate well, as it will be dominated by recruitment.

Related to macrophage division is the process governing intracellular bacteria when their host macrophage divides. Observations from the biological experiments show infected macrophages dividing to produce either two infected daughter cells, or one infected daughter and one healthy daughter. As a result of multiple observations within the data, this mechanism was included in the model, despite not appearing in either the biological or mathematical literature. As for macrophage division, it should be noted that this mechanism will not translate to an in vivo model.

Despite not translating well to an in vivo model, it was important to include these mechanisms in the modelling framework, as a core goal of the work presented here was to develop a mathematical model that is able to capture and replicate the observed dynamics, before extending those and exploring hypothetical scenarios. Through sensitivity analyses it was often shown that the birth death process of macrophages was a significant driver of the simulated system, further supporting its inclusion in the model.

6.5 Data

The models presented in this study are intrinsically related the data from which they are parametrised. Discussed here is the duration of the experiments, the format of the data that was shared, and the fact that the model presented in this study was fitted to a single data source.

The length of the experiments is capped at 140 and 200 hours. This is a result of the bacteria population growing to a point at which it is difficult to continue analysing the images. Since the adaptive immune response requires approximately two weeks [85, 140], it is not unreasonable to continue modelling the interactions of the adaptive immune response for the full duration of the experiment. The innate immune response does include macrophage activation, which is not included in the model fitting process due to no data being available for this process. Proxy mechanisms representing activated macrophages are, however, investigated through computational experiments. A result of these limitations is that, were the experiments to last longer than 200 hours, there would not be much additional information to gain without first increasing the focus of the experiments to include additional data relating to the activation of the macrophages.

While there was exposure of the raw image data, as shown in section 1.5, the full data set was shared as numerical tabulated data of event times and population sizes. Additionally all data regarding macrophages is at the individual level, so while the model considers a population of macrophages, there is no macrophage population data to compare against. This is a result of the small frame that was required to measure the bacteria. Due to this format of the data, all mechanisms had to be fitted independently. As discussed above, a beneficial next step for this work would be to develop the model as a spatial model, using the raw image data to parametrise the movement of macrophages.

Lastly, each model presented in this thesis is fitted to a set of experiments originating from a single laboratory, with a single experimental design. This is unusual compared to the literature, in which the majority of models are parametrised using a range of experiments from differing sources [136, 148, 149, 151, 153, 157, 166]. This is necessary in larger scale models as it would not be feasible to measure everything in a single experiment. However, a drawback of this is that parameters from different sources may not be compatible within the same model. The benefit of fitting entirely to a single experiment is that parameters being fitted to different mechanisms will be reliably comparable to each other. Future work should validate the models presented in this study against a similar but alternative source of data to corroborate the findings.

6.6 Model fitting

Maximum likelihood methods are used throughout this thesis to parametrise the mechanisms being modelled. As mentioned in the previous section, the structure of the data means that mechanisms are fitted in isolation from each other. While there is a strong benefit in this process, in that at most three parameters are being optimised at any one time, there is the drawback that the interplay between mechanisms is not well captured. For example, the intracellular growth rate is fitted independently of the infected macrophage death rate. While fitting a load dependent death rate was investigated, it was not possible to identify the relationship between these two mechanisms from the data available.

During the validation process in chapter 5, a series of distributions were extracted from the data and the model in order to make a comparison between the two. An interesting future direction of research would be to develop a optimisation technique to parametrise all or most of the parameters in parallel, by comparing statistics generated by the model to statistics in the data. This would require the optimisation of several objective functions while searching a high dimensional parameter space. The result, however, would be a model that is inherently designed to capture the interplay between mechanisms which, as has been shown in this thesis, is an important part of the system.

Chapter 7

Conclusions

The primary goal of this body of work was to develop a mathematical model of the in vitro interactions between innate alveolar macrophage and invading *Mycobacterium tuberculosis* (Mtb) as observed in a novel experimental data set. The secondary goal was to use these models to investigate the dominant mechanisms which affect the outcome of the system and to investigate methods of controlling the bacteria population. The models were parametrised using a novel data set collected in a single study described in section 1.5.

A range of novel mathematical models have been fitted to a novel data set. The models capture the in vitro interactions between macrophages and Mtb over a period of approximately 200 hours. This covers the period of time immediately after inoculation, prior to the adaptive immune response kicking in. It has been shown that fitting distinct mechanisms independently, using relatively simple techniques such as maximum likelihood estimation, results in a robust model with which computational experiments can be performed.

Each chapter builds on the knowledge gained in the previous chapter, and each model accounts for the limitations of the previous model. Preliminarily, a simple ordinary differential equation (ODE) model was fitted to sample data as a proof of concept. This model showed that early capture of extracellular bacteria was key to reducing the resultant load, however it also demonstrated that explicitly modelling intracellular macrophages was required. This was due to populations of bacteria being overestimated as a result of the assumed value of intracellular load. It was also demonstrated that exponential wait times are not an appropriate model for modelling macrophage division. These drawbacks were addressed in the following chapter.

Fitting the macrophage growth rate with an Erlang distribution improved the qualitative fit. Modelling this distribution in ODEs resulted in a variable number of equations, to account for the parametric number of growth stages for the cells. In addition, the intracellular bacteria were modelled compartmentally, further adding to the variability in the dimension of the system. The introduction of these mechanisms resulted in a more qualitatively accurate model in which the maximum intracellular load, the only parameter taken from the literature, no longer adversely affected the output. Using this model it was shown that while increasing the infection rate has a protective effect, this is only true up to a point, after which it becomes detrimental again. This demonstrates the dual role that macrophages play in Mtb infections. While they are initially protective for the host, eventually they will succumb, and become protective for the bacteria. When modelling populations in the single digits, finite size fluctuations play an important role. Migrating the model to a stochastic framework enabled investigation of these effects.

Along with migrating to a stochastic framework, the data set being worked with was improved through the addition of extra experiments and measurements. This improved the confidence in parameter fitting, and highlighted the observed inconsistencies in the macrophage birth death process. This was remedied by modelling the process as a continuous time Markov chain (CTMC), which parametrised both parts of the birth death process in parallel. Through sensitivity analysis, it was shown that the bacteria growth rate is the dominant factor in the outcome of the system. The homogeneous mixing assumption between healthy macrophages and the entire extracellular bacteria population resulted in a very low maximum likelihood estimate (MLE) for the rate of infection, thus infection events were rare at the beginning of simulations. In reality, bacteria exist in distinct populations. By reformulating the extracellular bacteria to be modelled as agents of aggregated populations, the mass action mixing terms increase and infection events are observed earlier in the simulation.

The final version of the model re-evaluated the model of bacteria from a single population to individual populations (aggregates). This reduced the effect of homogeneous mixing on the rate of infection, however was not able to fully resolve the issue, as shown while validating the model using distributions extracted from both the model and the data. The model showed that a small reduction in the intracellular growth rate across all macrophages provided a more protective effect than completely inhibiting growth in a minority of macrophages.

Through sensitivity analysis and computational experiments it was shown that the major predictor for the outcome of the initial interactions of macrophages and Mtb is the virulence of the bacteria, demonstrated by higher growth rates. While investigating the role of macrophages in limiting the infection, it was shown that it is more beneficial to partially limit the rate of intracellular growth within all macrophages, rather than fully limit the intracellular growth within a few macrophages. The next most significant mechanism was the rate of phagocytosis. While the growth rate of bacteria and the death rate of macrophages had a monotonic relationship with the outcome, this is not the case for phagocytosis. The core effector at the macrophages' disposable is how rapidly they can internalise the bacteria. It was shown that the rate of phagocytosis can be increased up to an optimum, after which increasing further is detrimental.

This thesis has explored, through an iterative process of novel model development, the complexities of the interactions between macrophages and *Mycobacterium tuberculosis*. It has been shown that a mechanistic model construction, parametrised with maximum likelihood estimation, results in a simple yet robust model which can be used to investigate the processes within a complex biological system. Finally the model was used to identify key mechanisms contributing to the outcome of the system, and through computation experiments, a range of control scenarios were tested.

Appendix A

Model 2: Colour coded equations

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \alpha_E E - \beta E \sum_{i=0}^{N_H} M_0^i + d_I \sum_{i=0}^{N_I} \sum_{k=1}^{N} k M_k^i + \alpha_I N^2 \sum_{i=0}^{N_I} M_N^i \tag{A.1}$$

$$\frac{\mathrm{d}M_0^0}{\mathrm{d}t} = 2g_H M_o^{N_H} + g_I M_1^{N_I} + (1-p) g_I \sum_{k=2}^{N} M_k^{N_I} - \beta E M_0^0 - d_H M_0^0 - g_H M_0^0 \tag{A.2}$$

$$\frac{\mathrm{d}M_0^i}{\mathrm{d}t} = g_H M_0^{i-1} - \beta E M_0^i - d_H M_0^i - g_H M_0^i \tag{A.3}$$

$$\begin{aligned}\frac{\mathrm{d}M_1^0}{\mathrm{d}t} &= p g_I \left(2M_2^{N_I} + M_3^{N_I}\right) + g_I M_1^{N_I} + \beta E \sum_{i=0}^{N_H} M_0^i \\ &\quad - d_I M_1^0 - g_I M_1^0 - \alpha_I M_1^0 \end{aligned} \tag{A.4}$$

$$\begin{aligned}\frac{\mathrm{d}M_k^0}{\mathrm{d}t} &= p g_I \left(M_{2k-1}^{N_I} + 2M_{2k}^{N_I} + M_{2k+1}^{N_I}\right) + (1-p) g_I M_k^{N_I} + \alpha_I (k-1) M_{k-1}^0 \\ &\quad - d_I M_k^0 - g_I M_k^0 - \alpha_I k M_k^0 \end{aligned} \tag{A.5}$$

$$\frac{\mathrm{d}M_k^i}{\mathrm{d}t} = g_I M_k^{i-1} + \alpha_I (k-1) M_{k-1}^i - d_I M_k^i - g_I M_k^i - \alpha_I k M_k^i \tag{A.6}$$

Color	Mechanism
🟫 Maroon	Extracellular growth
🟦 Blue	Infection
🟥 Red	Infected macrophage death
🟩 Green	Intracellular growth
🟧 Orange	Healthy macrophage growth
🟪 Pink	Infected macrophage growth
◼ Grey	Healthy macrophage death

Table A.1: Each colour in equations A.1 to A.6 refers to a separate mechanism being modelled in chapter 3.

Appendix B

Model code for chapter 4

```
class constant:
    def __init__(self):
        # model parameters
        # bacteria growth
        self.aE = 0.0339
        self.aI = 0.0206

        # macrophage birth death process
        self.gh = 0.0230
        self.ghN = 2
        self.dh = 0.0146
        self.dhN = 2
        self.gi = 0.0054
        self.giN = 2
        self.di = 0.0193
        self.diN = 2

        # infection
        self.beta = 2.629e-4

        # maximum intracellular load
        self.N = 50

        # probability of infected-infected division
        self.p_div = 0.1902

        # simulation parameters
        self.t0 = 0.0
        self.maxT = 200.0
        self.num_cells = 10
        self.e0 = 1
        self.inf_start = True


class Macrophage:
    def __init__(
        self,
        born=0,
        age_g=1,
        age_d=1,
        load=0,
        divided=False,
        died=False,
        alive=True,
    ):
        self.born = born
        self.inf_start = load > 0
        self.age_g = age_g
        self.age_d = age_d
        self.load = load
        self.divided = divided
        self.died = died
        self.alive = alive


class Experiment:
    def __init__(
        self,
        t=0,
        con=constant(),
    ):
        self.con = con
        self.t = self.con.t0

        # initialise macrophage population
        self.macrophages = [Macrophage() for _ in range(self.con.num_cells)]
        # initialise bacteria population
        self.bacteria = self.con.e0

        # does the bacteria begin extra-
        # or intra-cellularly
        if self.con.inf_start:
            self.infect()

    def healthy_macrophages(self):
        """return list of all healthy macrophages"""
        return [
            mac for mac in self.macrophages if mac.load == 0 and mac.alive
        ]

    def infected_macrophages(self):
        """return list of all infected macrophage"""
        return [mac for mac in self.macrophages if mac.load > 0 and mac.alive]

    def infect(self):
        """infect a random healthy macrophage with 1 extracellular bacteria"""
        if self.bacteria > 0:
            mac = np.random.choice(self.healthy_macrophages())
            mac.load += 1
            mac.age_g = 1
            mac.age_d = 1
            self.bacteria -= 1

    def bacteria_grow(self):
        """grow the extracellular bacteria population by 1"""
        self.bacteria += 1

    def healthy_grow(self):
        """progress a random healthy macrophage towards division. If it
        reaches division flag it as divided and create two new macrophages"""
        mac = np.random.choice(self.healthy_macrophages())

        if mac.age_g == self.con.ghN:
            mac.divided = self.t
            mac.alive = False
            self.macrophages.append(Macrophage(born=self.t))
            self.macrophages.append(Macrophage(born=self.t))
        else:
            mac.age_g += 1

    def healthy_death(self):
        """progress a random healthy macrophage towards death. If it
        reaches death flag it as dead."""
        mac = np.random.choice(self.healthy_macrophages())
        if mac.age_d == self.con.dhN:
            mac.died = self.t
            mac.alive = False
        else:
            mac.age_d += 1

    def infected_grow(self):
        """progress a random infected macrophage towards division. If
        it reaches division, flag it as divided and create two new
        macrophages. Determine the division of intracellular load
        according to the probability p_div"""
        mac = np.random.choice(self.infected_macrophages())
        if mac.age_g == self.con.giN:
            mac.divided = self.t
            mac.alive = False
            if np.random.rand() < self.con.p_div:
                self.macrophages.append(
                    Macrophage(born=self.t, load=mac.load // 2)
                )
                self.macrophages.append(
                    Macrophage(born=self.t, load=mac.load - mac.load // 2)
                )

            else:
                self.macrophages.append(
                    Macrophage(born=self.t, load=mac.load)
                )
                self.macrophages.append(Macrophage(born=self.t, load=0))
        else:
            mac.age_g += 1

    def infected_death(self):
        """progress a random infected macrophage towards death. If
        it reaches death, flag it as dead"""
        mac = np.random.choice(self.infected_macrophages())
        if mac.age_d == self.con.diN:
            mac.died = self.t
            mac.alive = False
            self.bacteria += mac.load
        else:
            mac.age_d += 1

    def mac_loads(self):
        """return a list of intracellular loads of infected macrophages"""
        return [mac.load for mac in self.infected_macrophages()]

    def intra_grow(self):
        """increase the intracellular load of a random infected macrophage
        by 1. The probability of choosing an infected macrophage is weighted
        by its current intracellular load"""
        load_weights = np.array(self.mac_loads())
        load_weights = load_weights / np.sum(load_weights)
        mac = np.random.choice(self.infected_macrophages(), p=load_weights)
        if mac.load == self.con.N:
            mac.died = self.t
            mac.alive = False
            self.bacteria += mac.load
        else:
            mac.load += 1

    def generate_rates(self):
        """For each possible event in the model define the rate at which
        it occurs and the function to carry out if it does happened."""
        num_healthy = len(self.healthy_macrophages())
        num_infected = len(self.infected_macrophages())
        num_bac, bac_g_rate = split_bac_pop(self.bacteria, self.con.aE)

        rates = []

        # healthy growth
        rates.append((num_healthy * self.con.gh, self.healthy_grow))
        # healthy death
        rates.append((num_healthy * self.con.dh, self.healthy_death))
        # infected growth
        rates.append((num_infected * self.con.gi, self.infected_grow))
        # infected death
        rates.append((num_infected * self.con.di, self.infected_death))
        # infection
        rates.append((num_bac * num_healthy * self.con.beta, self.infect))
        # intra growth
        rates.append(
            (np.sum(self.mac_loads()) * self.con.aI, self.intra_grow)
        )
        # extra growth
        rates.append((bac_g_rate, self.bacteria_grow))

        return rates


def run_sim(con=None):
    """Run the simulation using the Gillespie algorithm"""
    if con is None:
        con = constant()

    env = Experiment(con=con)

    while env.t < env.con.maxT:

        rates = env.generate_rates()
        sum_rates = np.sum([rate[0] for rate in rates])

        # end simulation if no possible events remaining
        if sum_rates == 0:
            break

        cum_rates = [0] + list(
            np.cumsum([rate[0] / sum_rates for rate in rates])
        )

        r1 = np.random.rand()
        r2 = np.random.rand()

        for event, (p1, p2) in enumerate(zip(cum_rates, cum_rates[1:])):
            if p1 < r1 <= p2:
                rates[event][1]()

        env.t += -np.log(r2) / sum_rates
```

Appendix C

Additionally considered mechanisms

Throughout the model development process, a range of additional mechanisms were hypothesised and considered against the data that was available. In this appendix a selection of these additional mechanisms are presented and discussed.

C.1 Load dependent death rate

The death rate of macrophages is modelled as a constant rate, with differing rates for healthy and infected macrophages. Here it is investigated whether the death rate of infected macrophages can be fitted such that it is dependent on the intracellular load.

To modify the current death rate to be dependent on load, the Erlang death rate d is modified to be a function of the intracellular load I.

$$d(I) = \begin{cases} d_0 & I = 0 \\ d_a \exp\Big(d_b \log\left(I\left(t\right)\right)\Big) + d_l & I > 0 \end{cases}$$

This equation is based on the Gompertz model of mortality. The Gompertz model gives a death rate dependent on age, but this model requires a death rate dependent on load. Since the intracellular bacteria grow exponentially in time, the log of the intracellular bacteria is used in the place of age in the Gompertz model. With the addition of two more parameters, the optimisation process is very computationally intensive. The computed likelihood surface is very flat, showing that there is an issue with model identifiability.

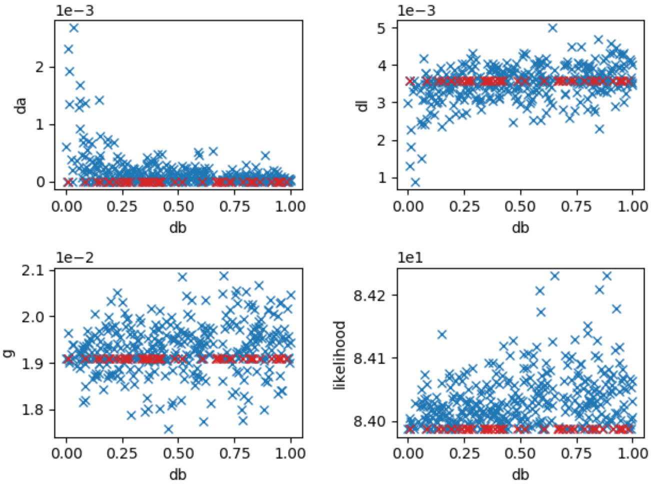

Figure C.1: The result of fitting d_a, d_l and g when d_b is fixed to some value between 0 and 1.

C.1.1 Investigating the relationship of d_a and d_b

In all runs of the optimisation the value of the Erlang parameter N optimises to two, so for the rest of this section N will be assumed to be two in order to reduce the execution time of the optimisation. It is clear that the likelihood surface is non-trivial, so in order to investigate the shape of it, the value of d_b is fixed to some value between 0 and 1, then the optimisation is rerun to find the values of d_a, d_l, and g. This also results in a value of the likelihood for this parameter set and thus four plots can be generated: each of the three fitted parameters against the range of d_b, and d_b against the resultant likelihood. Ideally these four plots would demonstrate clear minima, demonstrating the optimal parameter choices. This is not what is observed (see figure C.1).

For the values of d_l and g there is a clear line through the middle (highlighted in red) showing that these two parameters are likely independent of d_b. The plot of d_a against d_b shows that as d_b gets large, d_a gets small making the death rate approximate a constant. Finally the likelihood plot shows no clear minimum as the bottom boundary of the data is completely flat and no optimal value can be determined.

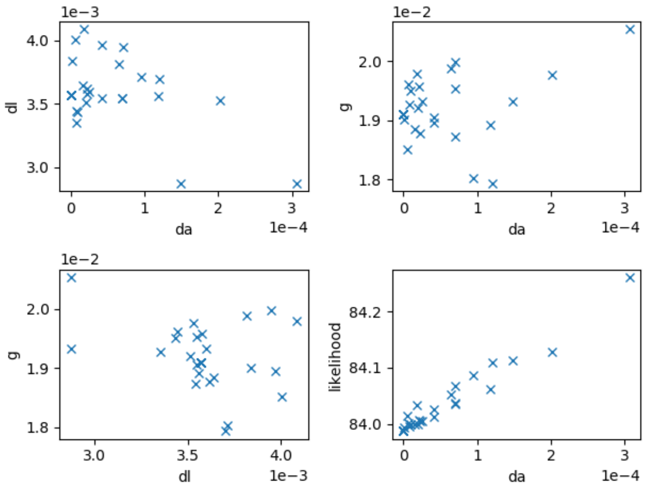

Figure C.2: Result of multiple optimisations forcing $d_b = 1$, i.e the death rate is linearly dependent on $I(t)$.

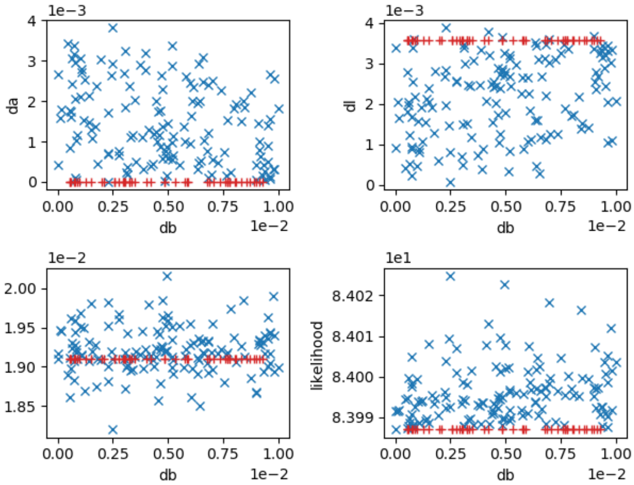

Figure C.3: Fitting Gompertz model with $d_b < 0.01$.

By simplifying the model to be a linear function of the intracellular load and ranging the value of d_a, the likelihood is minimised as $d_a \to 0$, that is the death rate is constant (see figure C.2).

Figure C.1 shows that for $d_b > 0.01$, d_a optimises to zero, making the function $d(I)$ approximate a constant. For this reason, the procedure is run again focusing on values of $d_b < 0.01$. Figure C.3 shows that when $d_b < 0.01$ the optimisation yields a very irregular likelihood surface, and the most likely value of d_a is still regularly 0.

The conclusion for this model of load dependent death rates for macrophages is that the model is not identifiable based on the current data and therefore shall not be used.

C.1.2 Weibull

Since the Gompertz model was not identifiable, a Weibull model of mortality is investigated as well, where $d(I) = d_a \log(I)^{d_b} + d_l$. Two runs of the optimisation with the Weibull mortality yields that d_a and d_l should be both be zero. This makes $d(I) \equiv 0$. Running the optimisation procedure over a range of d_b values gives the optimisation shown in figure C.4. There is a large range in the likelihood of values of d_b, while d_a consistently optimises to zero.

Setting the death rate to be constant $d(I) = d_l$ and running the optimisation results in consistent convergence of the optimal value. Plotting a range of values yields the plot in figure C.5. The likelihood function is a smooth function of d_l, demonstrating identifiability. It is thus concluded that the death rate should remain fixed and independent of the intracellular load.

In conclusion two models of load dependent death were considered, however both were not identifiable using the available data, and analysing the response of the

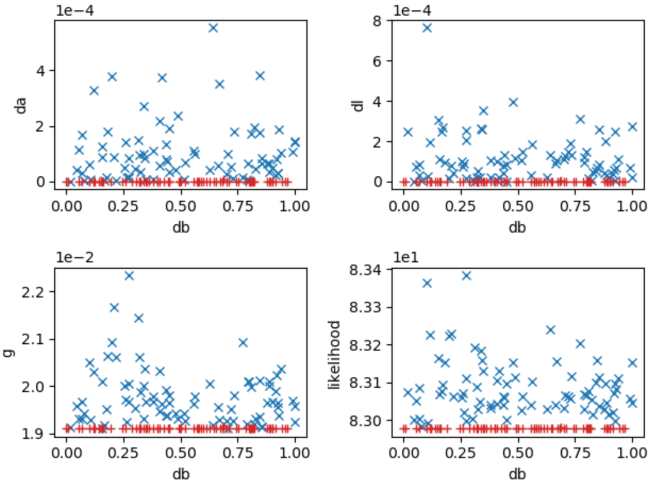

Figure C.4: Fitting Weibull over a range of values for d_b.

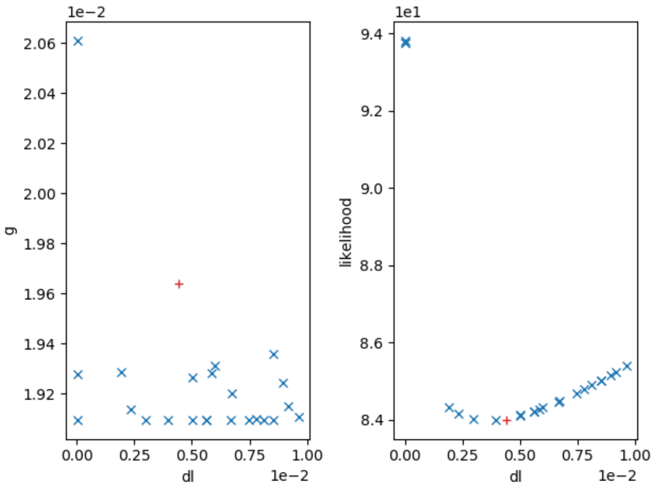

Figure C.5: Output of the optimisation for a range of values of d_l where the death rate $d(I) = d_l$ is a constant.

optimisation procedures demonstrated the tendency towards the constant model. Thus the model will remain as a fixed death rate for infected macrophages.

C.2 Infection uptake distribution

In chapter 5 the number of bacteria internalised during an infection event is allowed to vary away from one. In this section a range of models will be considered to evaluate how many are internalised.

C.2.1 Binomial distribution

There are 221 data points consisting of an estimate of the number of extracellular bacteria, and the number of bacteria internalised. To get a range of values for how many bacteria a macrophage picks up, a binomial distribution is fitted to the data using maximum likelihood. This results in $p = 0.2371$. However plotting the scatter plot of the data, and example values from a binomial distribution shows that this model is clearly unsuccessful, since the binomial distribution does not have enough variance as shown in figure C.6.

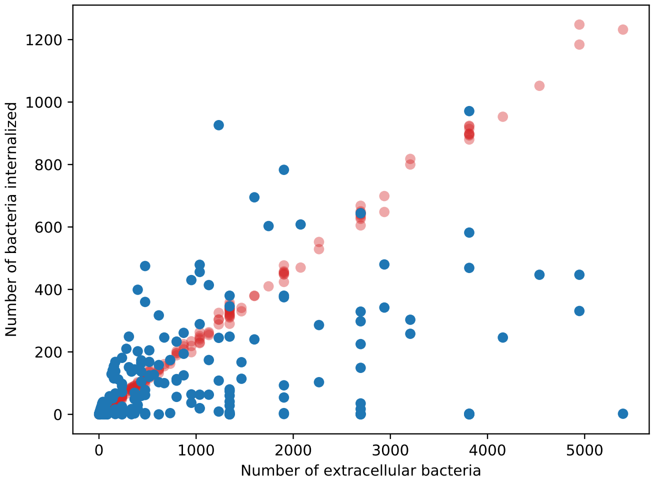

Figure C.6: Scatter plot of the data compared to example samples data drawn from a binomial distribution with $p = 0.2371$. Data is shown in blue, the proposed model is shown in red.

C.2.2 Uniform distribution

Since the binomial distribution is not a good fit, so an alternative distribution is required. Typically a negative binomial is used when more variance is required, however this has the problem that k can be larger than n, and sampling from this distribution will return a lot of zeros. In order to force it to fit the data well, the variance would need to be increased very high and the distribution would need to be truncated at n. This however results in an approximation of the uniform distribution. Thus, in the interest of simplicity of the model, and speed of the simulation, the uniform distribution of integers between 1 and n will be used. The result is shown in figure C.7.

For small number of extracellular bacteria this fits well, however as the number of extracellular bacteria gets larger than about 1000, the model begins to overestimate the number of bacteria phagocytosed.

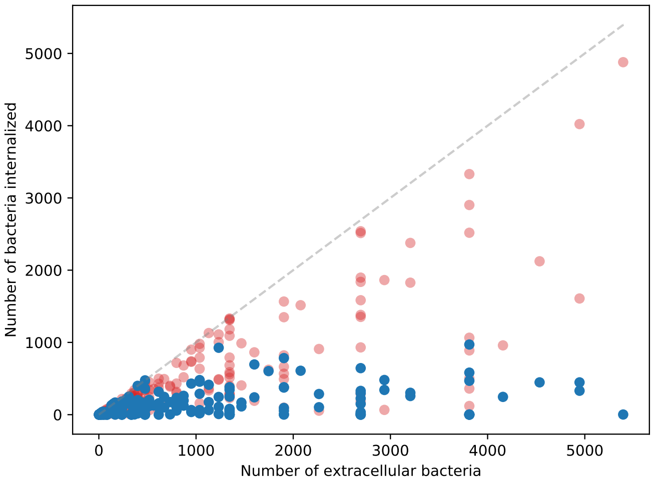

Figure C.7: The uniform distribution overestimates the number of bacteria internalised when the extracellular population gets large.

C.2.3 Truncated uniform distribution

Finally, a truncated uniform distribution is investigated, by truncating the uniform distribution, so that when n is below some value N_{trunc}, the uniform distribution from 1 to n is used, however if n is larger that N_{trunc} then the uniform distribution on 1 to N_{trunc} is used.

Despite appearing to fit the data better, this model does not have a mechanistic justification. Therefore it was decided that a flat uniform distribution will more realistically reflect the real-world system being modelled. Additionally, since the total population of extracellular bacteria only reaches such large numbers towards the end of the model simulations, the decision not to include the truncation parameter in the final model will not affect the output of the model in a significant way.

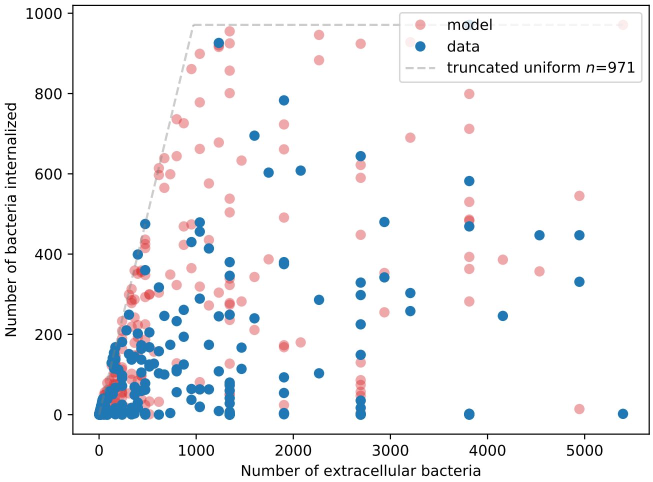

Figure C.8: The truncated uniform distribution provides a better qualitative fit to the data than the uniform distribution.

Abbreviations

ABM: agent-based model

AD: Anderson-Darling

BAL: bronchoalveolar lavage

BCG: Bacillus Calmette–Guérin

BMDM: bone marrow derived macrophage

CDC: Center for Disease Control and Prevention

CFU: colony forming unit

CI: confidence interval

CTMC: continuous time Markov chain

GFP: green fluorescence protein

GFP HQ: green fluorescence protein high quality

HIV: Human Immunodeficiency Virus

IFNγ: interferon-γ

IGRA: interferon-γ release assay

IL-10: interleukin-10

KDE: kernel densitiy estimate

KS: Kolmogorov-Smirnov

LHS: latin hypercube sampling

LTBI: latent TB infection

MAP: maximum a posteriori

MCMC: Monte Carlo Markov Chain

MDR-TB: multi-drug resistant TB

MLE: maximum likelihood estimate

MOI: multiplicity of infection

Mtb: *Mycobacterium tuberculosis*

NHP: non human primate

ODE: ordinary differential equation

PDE: partial differential equation

PN: petri net

PRCC: partial rank correlation coefficient

TB: tuberculosis

TNF: tumor necrosis factor

TNF-α: tumor necrosis factor alpha

TST: tuberculin skin test

WHO: World Health Organisation

XDR-TB: extensively drug resitant TB

Bibliography

[1] W. H. Organization. *Global Tuberculosis Report 2019*. World Health Organization, 2019. ISBN: 978-92-4-156571-4.

[2] W. H. Organisation. *WHO: The Top 10 Causes of Death*. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death. 2018.

[3] G. A. Colditz et al. "Efficacy of BCG Vaccine in the Prevention of Tuberculosis: Meta-analysis of the Published Literature". In: *JAMA* 271.9 (1994), pp. 698–702. DOI: `10.1001/jama.1994.03510330076038`.

[4] J. C. Kester and S. M. Fortune. "Persisters and beyond: Mechanisms of Phenotypic Drug Resistance and Drug Tolerance in Bacteria". In: *Critical Reviews in Biochemistry and Molecular Biology* 49.2 (2014), pp. 91–101. DOI: `10.3109/10409238.2013.869543`.

[5] G. Delogu and D. Goletti. "The Spectrum of Tuberculosis Infection: New Perspectives in the Era of Biologics". In: *The Journal of Rheumatology Supplement* 91.0 (2014), pp. 11–16. DOI: `10.3899/jrheum.140097`.

[6] T. Lillebaek et al. "Stability of DNA Patterns and Evidence of *Mycobacterium Tuberculosis* Reactivation Occurring Decades after the Initial Infection". In: *The Journal of Infectious Diseases* 188.7 (2003), pp. 1032–1039. DOI: `10.1086/378240`.

[7] C. E. Barry et al. "The Spectrum of Latent Tuberculosis: Rethinking the Biology and Intervention Strategies". In: *Nature Reviews Microbiology* 7.12 (2009), pp. 845–855. DOI: `10.1038/nrmicro2236`.

[8] S. H. E. Kaufmann. "How Can Immunology Contribute to the Control of Tuberculosis?" In: *Nature Reviews Immunology* 1.1 (2001), pp. 20–30. DOI: `10.1038/35095558`.

[9] M. Woo et al. "Mycobacterium Tuberculosis Infection and Innate Responses in a New Model of Lung Alveolar Macrophages". In: *Frontiers in Immunology* 9 (2018), p. 438. DOI: `10.3389/fimmu.2018.00438`.

[10] K. A. McDonough, Y. Kress, and B. R. Bloom. "Pathogenesis of Tuberculosis: Interaction of Mycobacterium Tuberculosis with Macrophages." In: *Infection and Immunity* 61.7 (1993), pp. 2763–2773. ISSN: 0019-9567, 1098-5522.

[11] P. Narasimhan et al. "Risk Factors for Tuberculosis". In: *Pulmonary Medicine* 2013 (2013). DOI: `10.1155/2013/828939`.

[12] W. H. Organization. *WHO | Data and Statistics*. http://www.who.int/hiv/data/en/. 2020.

[13] Y. A. Melsew et al. "The Role of Super-Spreading Events in Mycobacterium Tuberculosis Transmission: Evidence from Contact Tracing". In: *BMC Infectious Diseases* 19.1 (2019), pp. 1–9. DOI: `10.1186/s12879-019-3870-1`.

[14] A. Handel et al. "Evidence for Supercoughers in an Analysis of Six Tuberculosis Cohorts from China, Peru, The Gambia and Uganda". In: *The International Journal of Tuberculosis and Lung Disease* Volume 23.Number 12 (2019), 1286–1292(7).

[15] X. Cen, Z. Feng, and Y. Zhao. “Emerging Disease Dynamics in a Model Coupling Within-Host and between-Host Systems”. In: *Journal of Theoretical Biology* 361 (2014), pp. 141–151. DOI: `10.1016/j.jtbi.2014.07.030`.

[16] K. P. Fennelly et al. “Variability of Infectious Aerosols Produced during Coughing by Patients with Pulmonary Tuberculosis”. In: *American Journal of Respiratory and Critical Care Medicine* 186.5 (2012), pp. 450–457. DOI: `10.1164/rccm.201203-0444OC`.

[17] L. S. Schlesinger. “Entry of Mycobacterium Tuberculosis into Mononuclear Phagocytes”. In: *Current Topics in Microbiology and Immunology* 215 (1996), pp. 71–96. DOI: `10.1007/978-3-642-80166-2_4`.

[18] D. G. Russell. “Mycobacterium Tuberculosis : Here Today, and Here Tomorrow”. In: *Nature Reviews Molecular Cell Biology* 2.8 (2001), pp. 569–578. DOI: `10.1038/35085034`.

[19] A. J. Verrall et al. “Early Clearance of Mycobacterium Tuberculosis Is Associated With Increased Innate Immune Responses”. In: *The Journal of Infectious Diseases* (2019). DOI: `10.1093/infdis/jiz147`.

[20] T. C. Zahrt. “Molecular Mechanisms Regulating Persistent Mycobacterium Tuberculosis Infection”. In: *Microbes and Infection* 5.2 (2003), pp. 159–167. DOI: `10.1016/S1286-4579(02)00083-7`.

[21] A. O’Garra et al. “The Immune Response in Tuberculosis”. In: *Annual Review of Immunology* 31.1 (2013), pp. 475–527. DOI: `10.1146/annurev-immunol-032712-095939`.

[22] J.-W. Ai et al. “Updates on the Risk Factors for Latent Tuberculosis Reactivation and Their Managements”. In: *Emerging Microbes & Infections* 5.1 (2016), pp. 1–8. DOI: `10.1038/emi.2016.10`.

[23] P.-J. Cardona. “A Dynamic Reinfection Hypothesis of Latent Tuberculosis Infection”. In: *Infection* 37.2 (2009), pp. 80–86. DOI: `10.1007/s15010-008-8087-y`.

[24] A. J. Verrall et al. “Early Clearance of Mycobacterium Tuberculosis: A New Frontier in Prevention”. In: *Immunology* 141.4 (2014), pp. 506–513. DOI: `10.1111/imm.12223`.

[25] J. L. Flynn, J. Chan, and P. L. Lin. “Macrophages and Control of Granulomatous Inflammation in Tuberculosis”. In: *Mucosal Immunology* 4.3 (2011), pp. 271–278. DOI: `10.1038/mi.2011.14`.

[26] C. Warrender, S. Forrest, and F. Koster. “Modeling Intercellular Interactions in Early Mycobacterium Infection”. In: *Bulletin of Mathematical Biology* 68.8 (2006), pp. 2233–2261. DOI: `10.1007/s11538-006-9103-y`.

[27] K. D. Mayer-Barber and D. L. Barber. “Innate and Adaptive Cellular Immune Responses to Mycobacterium Tuberculosis Infection”. In: *Cold Spring Harbor Perspectives in Medicine* 5.12 (2015), a018424. DOI: `10.1101/cshperspect.a018424`.

[28] N. W. Schluger and W. N. Rom. “The Host Immune Response to Tuberculosis”. In: *American journal of respiratory and critical care medicine* 157.3 (1998), pp. 679–691.

[29] R. Van Crevel, T. H. Ottenhoff, and J. W. Van Der Meer. “Innate Immunity to Mycobacterium Tuberculosis”. In: *Clinical microbiology reviews* 15.2 (2002), pp. 294–309.

[30] J. C. Leemans et al. “Depletion of Alveolar Macrophages Exerts Protective Effects in Pulmonary Tuberculosis in Mice”. In: *The Journal of Immunology* 166.7 (2001), pp. 4604–4611.

[31] J. C. Leemans et al. “Macrophages Play a Dual Role during Pulmonary Tuberculosis in Mice”. In: *The Journal of Infectious Diseases* 191.1 (2005), pp. 65–74. DOI: `10.1086/426395`.

[32] D. B. Young, H. P. Gideon, and R. J. Wilkinson. “Eliminating Latent Tuberculosis”. In: *Trends in Microbiology* 17.5 (2009), pp. 183–188. DOI: `10.1016/j.tim.2009.02.005`.

[33] P. C. Hopewell. “Overview of Clinical Tuberculosis”. In: *Tuberculosis*. Ed. by B. R. Bloom. Washington, DC, USA: ASM Press, 2014, pp. 25–46. DOI: `10.1128/9781555818357.ch3`.

[34] C. W. McMillen. *Discovering Tuberculosis: A Global History, 1900 to the Present*. 2016. ISBN: 978-0-300-21348-5.

[35] H. Levy et al. “Acute Respiratory Failure in Active Tuberculosis:” in: *Critical Care Medicine* 15.3 (1987), pp. 221–225. DOI: `10.1097/00003246-198703000-00008`.

[36] T. Röszer. “Understanding the Biology of Self-Renewing Macrophages”. In: *Cells* 7.8 (2018). DOI: `10.3390/cells7080103`.

[37] F. Ginhoux et al. “New Insights into the Multidimensional Concept of Macrophage Ontogeny, Activation and Function”. In: *Nature Immunology* 17.1 (2016), pp. 34–40. DOI: `10.1038/ni.3324`.

[38] A. V. Misharin et al. “Monocyte-Derived Alveolar Macrophages Drive Lung Fibrosis and Persist in the Lung over the Life Span”. In: *Journal of Experimental Medicine* 214.8 (2017), pp. 2387–2404. DOI: `10.1084/jem.20162152`.

[39] M. Schneemann and G. Schoeden. “Macrophage Biology and Immunology: Man Is Not a Mouse”. In: *Journal of Leukocyte Biology* 81.3 (2007), pp. 579–579. DOI: `10.1189/jlb.1106702`.

[40] G. Patterson et al. “Use of the Green Fluorescent Protein and Its Mutants in Quantitative Fluorescence Microscopy”. In: *Biophysical Journal* 73.5 (1997), pp. 2782–2790. DOI: `10.1016/S0006-3495(97)78307-3`.

[41] C. F. Craver. “When Mechanistic Models Explain”. In: *Synthese* 153.3 (2006), pp. 355–376. DOI: `10.1007/s11229-006-9097-x`.

[42] M. J. Keeling and P. Rohani. *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, 2008. DOI: `10.1515/9781400841035`.

[43] L. Fox. *Numerical Solution of Ordinary Differential Equations*. Springer Science & Business Media, 2012.

[44] C. Lobry and T. Sari. “Migrations in the Rosenzweig-MacArthur Model and the "Atto-Fox" Problem”. In: *Revue Africaine de la Recherche en Informatique et Mathématiques Appliquées* Volume 20 - 2015 - Special... (2015), p. 1990. DOI: `10.46298/arima.1990`.

[45] P. Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Methods* 17.3 (2020), pp. 261–272. DOI: `10.1038/s41592-019-0686-2`.

[46] P. N. Brown, G. D. Byrne, and A. C. Hindmarsh. “VODE: A Variable-Coefficient ODE Solver”. In: *SIAM Journal on Scientific and Statistical Computing* 10.5 (1989), pp. 1038–1051. DOI: `10.1137/0910062`.

[47] C. W. Gear. “The Numerical Integration of Ordinary Differential Equations”. In: *Mathematics of Computation* 21.98 (1967), pp. 146–156. DOI: `10.1090/S0025-5718-1967-0225494-5`.

[48] A. L. Bauer, C. A. A. Beauchemin, and A. S. Perelson. “Agent-Based Modeling of Host-Pathogen Systems: The Successes and Challenges”. In: *Information Sciences* 179.10 (2009), pp. 1379–1389. DOI: `10.1016/j.ins.2008.11.012`.

[49] D. T. Gillespie. “Exact Stochastic Simulation of Coupled Chemical Reactions”. In: *The Journal of Physical Chemistry* 81.25 (1977), pp. 2340–2361. DOI: `10.1021/j100540a008`.

[50] A. A. Markov. “Extension of the Limit Theorems of Probability Theory to a Sum of Variables Connected in a Chain”. In: *Dynamic probabilistic systems* 1 (1971), pp. 552–577.

[51] I. M. Sobol. “Global Sensitivity Indices for Nonlinear Mathematical Models and Their Monte Carlo Estimates”. In: *Mathematics and Computers in Simulation*. The Second IMACS Seminar on Monte Carlo Methods 55.1 (2001), pp. 271–280. DOI: `10.1016/S0378-4754(00)00270-6`.

[52] M. Renardy et al. “Global Sensitivity Analysis of Biological Multiscale Models”. In: *Current Opinion in Biomedical Engineering* 11 (2019), pp. 109–116. DOI: `10.1016/j.cobme.2019.09.012`.

[53] A. Saltelli. “Making Best Use of Model Evaluations to Compute Sensitivity Indices”. In: *Computer Physics Communications* 145.2 (2002), pp. 280–297. DOI: `10.1016/S0010-4655(02)00280-1`.

[54] A. Saltelli et al. *Global Sensitivity Analysis: The Primer.* John Wiley & Sons, 2008. ISBN: 978-0-470-72517-7.

[55] J. Herman et al. *SALib.* 2021.

[56] F. J. Massey. “The Kolmogorov-Smirnov Test for Goodness of Fit”. In: *Journal of the American Statistical Association* 46.253 (1951), pp. 68–78. DOI: `10.1080/01621459.1951.10500769`.

[57] F. W. Scholz and M. A. Stephens. “K -Sample Anderson–Darling Tests”. In: *Journal of the American Statistical Association* 82.399 (1987), pp. 918–924. DOI: `10.1080/01621459.1987.10478517`.

[58] M. Castro and R. J. de Boer. “Testing Structural Identifiability by a Simple Scaling Method”. In: *PLOS Computational Biology* 16.11 (2020). Ed. by M. P. Davenport, e1008248. DOI: `10.1371/journal.pcbi.1008248`.

[59] K. Godfrey and J. DiStefano. “Identifiability of Model Parameter”. In: *IFAC Proceedings Volumes* 18.5 (1985), pp. 89–114. DOI: `10.1016/S1474-6670(17)60544-5`.

[60] A. Re and P. Lecca. “Determining Structural Parameter Identifiability in Biological Dynamical Models by Analysing the Statistical Properties of the Likelihood Behaviour”. In: *2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB).* Siena, Italy: IEEE, 2019, pp. 1–8. DOI: `10.1109/CIBCB.2019.8791488`.

[61] S. T. Cole et al. “Deciphering the Biology of Mycobacterium Tuberculosis from the Complete Genome Sequence”. In: *Nature* 393.6685 (1998), pp. 537–544. DOI: `10.1038/31159`.

[62] T. Miyoshi-Akiyama et al. “Complete Annotated Genome Sequence of Mycobacterium Tuberculosis Erdman”. In: *Journal of Bacteriology* 194.10 (2012), pp. 2770–2770. DOI: `10.1128/JB.00353-12`.

[63] G. P. Kubica, T. H. Kim, and F. P. Dunbar. “Designation of Strain H37Rv as the Neotype of Mycobacterium Tuberculosis”. In: *International Journal of Systematic Bacteriology* 22.2 (1972), pp. 99–106. DOI: `10.1099/00207713-22-2-99`.

[64] T. R. Ioerger et al. “Variation among Genome Sequences of H37Rv Strains of Mycobacterium Tuberculosis from Multiple Laboratories”. In: *Journal of Bacteriology* 192.14 (2010), pp. 3645–3653. DOI: `10.1128/JB.00166-10`.

[65] B. W. James, A. Williams, and P. D. Marsh. “The Physiology and Pathogenicity of Mycobacterium Tuberculosis Grown under Controlled Conditions in a Defined Medium”. In: *Journal of Applied Microbiology* 88.4 (2000), pp. 669–677. DOI: `10.1046/j.1365-2672.2000.01020.x`.

[66] R. A. Cox. “Quantitative Relationships for Specific Growth Rates and Macromolecular Compositions of Mycobacterium Tuberculosis, Streptomyces Coelicolor A3(2) and Escherichia Coli B/r: An Integrative Theoretical Approach”. In: *Microbiology,* 150.5 (2004), pp. 1413–1426. DOI: `10.1099/mic.0.26560-0`.

[67] M. Lipsitch and B. R. Levin. “Population Dynamics of Tuberculosis Treatment: Mathematical Models of the Roles of Non-Compliance and Bacterial Heterogeneity in the Evolution of Drug Resistance”. In: *The International Journal of Tuberculosis and Lung Disease: The Official Journal of the International Union Against Tuberculosis and Lung Disease* 2.3 (1998), pp. 187–199. ISSN: 1027-3719.

[68] A. Jindani et al. “The Early Bactericidal Activity of Drugs in Patients with Pulmonary Tuberculosis”. In: *The American Review of Respiratory Disease* 121.6 (1980), pp. 939–949. DOI: `10.1164/arrd.1980.121.6.939`.

[69] R. J. North and A. A. Izzo. “Mycobacterial Virulence. Virulent Strains of Mycobacteria Tuberculosis Have Faster in Vivo Doubling Times and Are Better Equipped to Resist Growth-Inhibiting Functions of Macrophages in the Presence and Absence of Specific Immunity.” In: *Journal of Experimental Medicine* 177.6 (1993), pp. 1723–1733. DOI: `10.1084/jem.177.6.1723`.

[70] C. L. Sershen, S. J. Plimpton, and E. E. May. “Oxygen Modulates the Effectiveness of Granuloma Mediated Host Response to Mycobacterium Tuberculosis: A Multiscale Computational Biology Approach”. In: *Frontiers in Cellular and Infection Microbiology* 6 (2016). DOI: `10.3389/fcimb.2016.00006`.

[71] L. G. Wayne and L. G. Hayes. “An in Vitro Model for Sequential Study of Shiftdown of Mycobacterium Tuberculosis through Two Stages of Nonreplicating Persistence.” In: *Infection and Immunity* 64.6 (1996), pp. 2062–2069. ISSN: 0019-9567.

[72] P. L. Lin et al. “Sterilization of Granulomas Is Common in Active and Latent Tuberculosis despite Within-Host Variability in Bacterial Killing”. In: *Nature Medicine* 20.1 (2014), pp. 75–79. DOI: `10.1038/nm.3412`.

[73] E. R. Sumner and S. V. Avery. “Phenotypic Heterogeneity: Differential Stress Resistance among Individual Cells of the Yeast Saccharomyces Cerevisiae”. In: *Microbiology* 148.2 (2002), pp. 345–351. DOI: `10.1099/00221287-148-2-345`.

[74] N. Dhar, J. McKinney, and G. Manina. “Phenotypic Heterogeneity in *Mycobacterium Tuberculosis*”. In: *Microbiology Spectrum* 4.6 (2016). Ed. by W. R. Jacobs Jr. et al. DOI: `10.1128/microbiolspec.TBTB2-0021-2016`.

[75] H. P. Gideon and J. L. Flynn. “Latent Tuberculosis: What the Host “Sees”?” In: *Immunologic Research* 50.2-3 (2011), pp. 202–212. DOI: `10.1007/s12026-011-8229-7`.

[76] C. L. Hendon-Dunn et al. “A Flow Cytometry Method for Rapidly Assessing Mycobacterium Tuberculosis Responses to Antibiotics with Different Modes of Action”. In: *Antimicrobial Agents and Chemotherapy* 60.7 (2016), pp. 3869–3883. DOI: `10.1128/AAC.02712-15`.

[77] O. J. Billington, T. D. McHugh, and S. H. Gillespie. “Physiological Cost of Rifampin Resistance Induced In Vitro in Mycobacterium Tuberculosis”. In: *Antimicrobial Agents and Chemotherapy* 43.8 (1999), pp. 1866–1869. ISSN: 0066-4804.

[78] T. Cohen, B. Sommers, and M. Murray. “The Effect of Drug Resistance on the Fitness of Mycobacterium Tuberculosis”. In: *The Lancet Infectious Diseases* 3.1 (2003), pp. 13–21. DOI: `10.1016/S1473-3099(03)00483-3`.

[79] E. Pienaar, J. J. Linderman, and D. E. Kirschner. “Emergence and Selection of Isoniazid and Rifampin Resistance in Tuberculosis Granulomas”. In: *PLOS ONE* 13.5 (2018), e0196322. DOI: `10.1371/journal.pone.0196322`.

[80] M. L. Bastos, Z. Lan, and D. Menzies. “An Updated Systematic Review and Meta-Analysis for Treatment of Multidrug-Resistant Tuberculosis”. In: *European Respiratory Journal* 49.3 (2017). DOI: `10.1183/13993003.00803-2016`.

[81] D. J. Ordway et al. “Drug-Resistant Strains of Mycobacterium Tuberculosis Exhibit a Range of Virulence for Mice.” In: *Infection and Immunity* 63.2 (1995), pp. 741–743. ISSN: 0019-9567, 1098-5522.

[82] C. f. D. C. a. P. CDC. *Interim Laboratory Biosafety Guidance for Extensively Drug-Resistant (XDR) Mycobacterium Tuberculosis Strains.* https://www.cdc.gov/tb/topic/laboratory/biosafetyguidance_xdrtb.htm. 2012.

[83] “Extensively drug-resistant tuberculosis (XDR-TB): recommendations for prevention and control”. In: *Releve Epidemiologique Hebdomadaire* 81.45 (2006), pp. 430–432. ISSN: 0049-8114.

[84] G. P. Walsh et al. “The Philippine Cynomolgus Monkey (Macaca Fasicularis) Provides a New Nonhuman Primate Model of Tuberculosis That Resembles Human Disease”. In: *Nature Medicine* 2.4 (1996), pp. 430–436. DOI: `10.1038/nm0496-430`.

[85] S. V. Capuano et al. “Experimental Mycobacterium Tuberculosis Infection of Cynomolgus Macaques Closely Resembles the Various Manifestations of Human M. Tuberculosis Infection”. In: *Infection and Immunity* 71.10 (2003), pp. 5831–5844. DOI: `10.1128/IAI.71.10.5831-5844.2003`.

[86] E. Goldman and L. H. Green, eds. *Practical Handbook of Microbiology.* 2nd ed. Boca Raton: CRC Press, 2009. ISBN: 978-0-8493-9365-5.

[87] G. S. Dean et al. “Minimum Infective Dose of Mycobacterium Bovis in Cattle”. In: *Infection and Immunity* 73.10 (2005), pp. 6467–6471. DOI: `10.1128/IAI.73.10.6467-6471.2005`.

[88] C. R. Plumlee et al. “Ultra-Low Dose Aerosol Infection of Mice with Mycobacterium Tuberculosis More Closely Models Human Tuberculosis”. In: *Cell Host & Microbe* 29.1 (2021), 68–82.e5. DOI: `10.1016/j.chom.2020.10.003`.

[89] M. J. Evans et al. “Cell Division of Alveolar Macrophages in Rat Lung Following Exposure to NO2”. In: *The American Journal of Pathology* 70.2 (1973), pp. 199–208. ISSN: 0002-9440.

[90] P. B. Bitterman et al. “Alveolar Macrophage Replication. One Mechanism for the Expansion of the Mononuclear Phagocyte Population in the Chronically Inflamed Lung.” In: *The Journal of Clinical Investigation* 74.2 (1984), pp. 460–469. DOI: `10.1172/JCI111443`.

[91] R. G. Barbers et al. “Enhanced Alveolar Monocytic Phagocyte (Macrophage) Proliferation in Tobacco and Marijuana Smokers”. In: *The American Review of Respiratory Disease* 143.5 Pt 1 (1991), pp. 1092–1095. DOI: `10.1164/ajrccm/143.5_Pt_1.1092`.

[92] S. J. Jenkins et al. “Local Macrophage Proliferation, Rather than Recruitment from the Blood, Is a Signature of TH2 Inflammation”. In: *Science* 332.6035 (2011), pp. 1284–1288. DOI: `10.1126/science.1204351`.

[93] R. Van Furth, M. M. C. Diesselhoff-den Dulk, and H. Mattie. “Quantitative Study on the Production and Kinetics of Mononuclear Phagocytes During an Acute Inflammatory Reaction”. In: *The Journal of Experimental Medicine* 138.6 (1973), pp. 1314–1330. ISSN: 0022-1007.

[94] C. J. Cambier et al. “Mycobacteria Manipulate Macrophage Recruitment through Coordinated Use of Membrane Lipids”. In: *Nature* 505.7482 (2014), pp. 218–222. DOI: `10.1038/nature12799`.

[95] R. M. du Bois. “The Alveolar Macrophage.” In: *Thorax* 40.5 (1985), pp. 321–327. DOI: `10.1136/thx.40.5.321`.

[96] J. Murphy et al. “The Prolonged Life-Span of Alveolar Macrophages”. In: *American Journal of Respiratory Cell and Molecular Biology* 38.4 (2008), pp. 380–385. DOI: `10.1165/rcmb.2007-0224RC`.

[97] N. Joshi, J. M. Walter, and A. V. Misharin. “Alveolar Macrophages”. In: *Cellular Immunology* 330 (2018), pp. 86–90. DOI: `10.1016/j.cellimm.2018.01.005`.

[98] D. G. More and D. S. Nelson. “Antigen-Induced Mitosis in Liver Macrophages of Immunized Mice”. In: *Experientia* 28.5 (1972), pp. 566–567. DOI: `10.1007/BF01931880`.

[99] L. E. Lewis et al. “Candida Albicans Infection Inhibits Macrophage Cell Division and Proliferation”. In: *Fungal Genetics and Biology* 49.9 (2012), pp. 679–680. DOI: `10.1016/j.fgb.2012.05.007`.

[100] S. Dühring et al. “Modelling the Host–Pathogen Interactions of Macrophages and Candida Albicans Using Game Theory and Dynamic Optimization”. In: *Journal of The Royal Society Interface* 14.132 (2017), p. 20170095. DOI: `10.1098/rsif.2017.0095`.

[101] V. B. Antony et al. “Recruitment of Inflammatory Cells to the Pleural Space. Chemotactic Cytokines, IL-8, and Monocyte Chemotactic Peptide-1 in Human Pleural Fluids”. In: *Journal of Immunology (Baltimore, Md.: 1950)* 151.12 (1993), pp. 7216–7223. ISSN: 0022-1767.

[102] A. A. Nash, R. G. Dalziel, and J. R. Fitzgerald. *Mims’ Pathogenesis of Infectious Disease*. Academic Press, 2015. ISBN: 978-0-12-397781-6.

[103] L. S. Meena and n. Rajni. “Survival Mechanisms of Pathogenic Mycobacterium Tuberculosis H37Rv”. In: *The FEBS journal* 277.11 (2010), pp. 2416–2427. DOI: `10.1111/j.1742-4658.2010.07666.x`.

[104] J. Pieters. “Mycobacterium Tuberculosis and the Macrophage: Maintaining a Balance”. In: *Cell Host & Microbe* 3.6 (2008), pp. 399–407. DOI: `10.1016/j.chom.2008.05.006`.

[105] A. Aderem and D. M. Underhill. “MECHANISMS OF PHAGOCYTOSIS IN MACROPHAGES”. In: *Annual Review of Immunology* 17.1 (1999), pp. 593–623. DOI: `10.1146/annurev.immunol.17.1.593`.

[106] M. G. Bonecini-Almeida et al. “Induction of In Vitro Human Macrophage Anti-Mycobacterium Tuberculosis Activity: Requirement for IFN-γ and Primed Lymphocytes”. In: *The Journal of Immunology* 160.9 (1998), pp. 4490–4499. ISSN: 0022-1767, 1550-6606.

[107] I. E. Flesch and S. H. Kaufmann. “Activation of Tuberculostatic Macrophage Functions by Gamma Interferon, Interleukin-4, and Tumor Necrosis Factor”. In: *Infection and Immunity* 58.8 (1990), pp. 2675–2677. ISSN: 0019-9567.

[108] U. E. Schaible et al. “Cytokine Activation Leads to Acidification and Increases Maturation of Mycobacterium Avium-Containing Phagosomes in Murine Macrophages”. In: *Journal of Immunology (Baltimore, Md.: 1950)* 160.3 (1998), pp. 1290–1296. ISSN: 0022-1767.

[109] M. G. Gutierrez et al. “NF-κB Activation Controls Phagolysosome Fusion-Mediated Killing of Mycobacteria by Macrophages”. In: *The Journal of Immunology* 181.4 (2008), pp. 2651–2663. DOI: `10.4049/jimmunol.181.4.2651`.

[110] G. A. Rook et al. “Activation of Macrophages to Inhibit Proliferation of Mycobacterium Tuberculosis: Comparison of the Effects of Recombinant Gamma-Interferon on Human Monocytes and Murine Peritoneal Macrophages.” In: *Immunology* 59.3 (1986), pp. 333–338. ISSN: 0019-2805.

[111] G. S. Douvas et al. “Gamma Interferon Activates Human Macrophages to Become Tumoricidal and Leishmanicidal but Enhances Replication of Macrophage-Associated Mycobacteria.” In: *Infection and Immunity* 50.1 (1985), pp. 1–8. ISSN: 0019-9567, 1098-5522.

[112] A. Kahnert et al. “Alternative Activation Deprives Macrophages of a Coordinated Defense Program to Mycobacterium Tuberculosis”. In: *European Journal of Immunology* 36.3 (2006), pp. 631–647. DOI: `10.1002/eji.200535496`.

[113] M. Engele et al. “Induction of TNF in Human Alveolar Macrophages As a Potential Evasion Mechanism of Virulent *Mycobacterium Tuberculosis*”. In: *The Journal of Immunology* 168.3 (2002), pp. 1328–1337. DOI: `10.4049/jimmunol.168.3.1328`.

[114] L. E. DesJardin et al. “Mycobacterium Tuberculosis-Infected Human Macrophages Exhibit Enhanced Cellular Adhesion with Increased Expression of LFA-1 and ICAM-1 and Reduced Expression and/or Function of Complement Receptors, FcγRII and the Mannose Receptor”. In: *Microbiology* 148.10 (2002), pp. 3161–3171.

[115] J. J. Muldoon et al. “Macrophages Employ Quorum Licensing to Regulate Collective Activation”. In: *Nature Communications* 11.1 (2020), pp. 1–14. DOI: `10.1038/s41467-020-14547-y`.

[116] J. M. Tufariello, J. Chan, and J. L. Flynn. “Latent Tuberculosis: Mechanisms of Host and Bacillus That Contribute to Persistent Infection”. In: *The Lancet Infectious Diseases* 3.9 (2003), pp. 578–590. DOI: `10.1016/S1473-3099(03)00741-2`.

[117] F. Aldwell, D. Wedlock, and B. Buddle. “Sequential Activation of Alveolar Macrophages by IFNy and LPS Is Required for Enhanced Growth Inhibition of Virulent Mycobacterium Bovis but Not M. Bovis BCG”. In: *Immunology and cell biology* 75.2 (1997), pp. 161–166.

[118] S. Paul, P. Laochumroonvorapong, and G. Kaplan. “Comparable Growth of Virulent and Avirulent Mycobacterium Tuberculosis in Human Macrophages In Vitro”. In: *The Journal of Infectious Diseases* 174.1 (1996), pp. 105–112. DOI: `10.1093/infdis/174.1.105`.

[119] R. F. Silver, Q. Li, and J. J. Ellner. “Expression of Virulence of Mycobacterium Tuberculosiswithin Human Monocytes: Virulence Correlates with Intracellular Growth and Induction of Tumor Necrosis Factor Alpha but Not with Evasion of Lymphocyte-Dependent Monocyte Effector Functions”. In: *Infection and Immunity* 66.3 (1998), pp. 1190–1199. ISSN: 0019-9567, 1098-5522.

[120] M. Zhang et al. “Enhanced Capacity of a Widespread Strain of Mycobacterium Tuberculosis to Grow in Human Macrophages”. In: *The Journal of Infectious Diseases* 179.5 (1999), pp. 1213–1217. DOI: `10.1086/314738`.

[121] H. Bruns and S. Stenger. “New Insights into the Interaction of *Mycobacterium Tuberculosis* and Human Macrophages”. In: *Future Microbiology* 9.3 (2014), pp. 327–341. DOI: `10.2217/fmb.13.164`.

[122] J. L. Flynn and J. Chan. “Immunology of Tuberculosis”. In: *Annual Review of Immunology* 19.1 (2001), pp. 93–129. DOI: `10.1146/annurev.immunol.19.1.93`.

[123] V. Deretic et al. “Mycobacterium Tuberculosis Inhibition of Phagolysosome Biogenesis and Autophagy as a Host Defence Mechanism”. In: *Cellular Microbiology* 8.5 (2006), pp. 719–727. DOI: `10.1111/j.1462-5822.2006.00705.x`.

[124] J. A. Armstrong and P. D. Hart. “Response of Cultured Macrophages to Mycobacterium Tuberculosis, with Observations on Fusion of Lysosomes with Phagosomes”. In: *Journal of Experimental Medicine* 134.3 (1971), pp. 713–740. DOI: `10.1084/jem.134.3.713`.

[125] J. Keane, H. G. Remold, and H. Kornfeld. “Virulent Mycobacterium Tuberculosis Strains Evade Apoptosis of Infected Alveolar Macrophages”. In: *Journal of Immunology (Baltimore, Md.: 1950)* 164.4 (2000), pp. 2016–2020. DOI: `10.4049/jimmunol.164.4.2016`.

[126] D. H. Dockrell and M. K. B. Whyte. “Regulation of Phagocyte Lifespan in the Lung during Bacterial Infection”. In: *Journal of Leukocyte Biology* 79.5 (2006), pp. 904–908. DOI: `10.1189/jlb.1005555`.

[127] M. K. Balcewicz-Sablinska et al. “Pathogenic Mycobacterium Tuberculosis Evades Apoptosis of Host Macrophages by Release of TNF-R2, Resulting in Inactivation of TNF-α”. In: *The Journal of Immunology* 161.5 (1998), pp. 2636–2641. ISSN: 0022-1767, 1550-6606.

[128] J. Keane et al. “Infection by Mycobacterium Tuberculosis Promotes Human Alveolar Macrophage Apoptosis.” In: *Infection and Immunity* 65.1 (1997), pp. 298–304. ISSN: 0019-9567.

[129] M. Rojas et al. “TNF-alpha and IL-10 Modulate the Induction of Apoptosis by Virulent Mycobacterium Tuberculosis in Murine Macrophages”. In: *Journal of Immunology (Baltimore, Md.: 1950)* 162.10 (1999), pp. 6122–6131. ISSN: 0022-1767.

[130] C. Fratazzi et al. “Macrophage Apoptosis in Mycobacterial Infections”. In: *Journal of Leukocyte Biology* 66.5 (1999), pp. 763–764. DOI: `10.1002/jlb.66.5.763`.

[131] S. M. Behar, M. Divangahi, and H. G. Remold. “Evasion of Innate Immunity by Mycobacterium Tuberculosis : Is Death an Exit Strategy?” In: *Nature Reviews Microbiology* 8.9 (2010), pp. 668–674. DOI: `10.1038/nrmicro2387`.

[132] M. Oddo et al. “Fas Ligand-Induced Apoptosis of Infected Human Macrophages Reduces the Viability of Intracellular Mycobacterium Tuberculosis”. In: *The Journal of Immunology* 160.11 (1998), pp. 5448–5454.

[133] H. L. Collins and S. H. E. Kaufmann. “The Many Faces of Host Responses to Tuberculosis”. In: *Immunology* 103.1 (2001), pp. 1–9. DOI: `10.1046/j.1365-2567.2001.01236.x`.

[134] C. S. Hirsch et al. “Complement Receptor-Mediated Uptake and Tumor Necrosis Factor-Alpha-Mediated Growth Inhibition of Mycobacterium Tuberculosis by Human Alveolar Macrophages”. In: *Journal of Immunology (Baltimore, Md.: 1950)* 152.2 (1994), pp. 743–753. ISSN: 0022-1767.

[135] T. Repasy et al. “Intracellular Bacillary Burden Reflects a Burst Size for Mycobacterium Tuberculosis In Vivo”. In: *PLOS Pathogens* 9.2 (2013), e1003190. DOI: `10.1371/journal.ppat.1003190`.

[136] J. L. Segovia-Juarez, S. Ganguli, and D. Kirschner. “Identifying Control Mechanisms of Granuloma Formation during M. Tuberculosis Infection Using an Agent-Based Model”. In: *Journal of Theoretical Biology* 231.3 (2004), pp. 357–376. DOI: `10.1016/j.jtbi.2004.06.031`.

[137] A. M. Cooper. “Cell-Mediated Immune Responses in Tuberculosis”. In: *Annual Review of Immunology* 27.1 (2009), pp. 393–422. DOI: `10.1146/annurev.immunol.021908.132703`.

[138] B. Alberts. *Molecular Biology of the Cell*. 2017. ISBN: 978-1-315-73536-8.

[139] M. K. Jenkins et al. “In Vivo Activation of Antigen-Specific CD4 T Cells”. In: *Annual Review of Immunology* 19.1 (2001), pp. 23–45. DOI: `10.1146/annurev.immunol.19.1.23`.

[140] A. A. Chackerian et al. “Dissemination of Mycobacterium Tuberculosis Is Influenced by Host Factors and Precedes the Initiation of T-Cell Immunity”. In: *Infection and Immunity* 70.8 (2002), pp. 4501–4509. DOI: `10.1128/IAI.70.8.4501-4509.2002`.

[141] I. M. Orme and R. J. Basaraba. “The Formation of the Granuloma in Tuberculosis Infection”. In: *Seminars in Immunology*. Immunity to Mycobacterium Tuberculosis 26.6 (2014), pp. 601–609. DOI: `10.1016/j.smim.2014.09.009`.

[142] L. Ramakrishnan. “Revisiting the Role of the Granuloma in Tuberculosis”. In: *Nature Reviews Immunology* 12.5 (2012), pp. 352–366. DOI: `10.1038/nri3211`.

[143] J. M. Davis and L. Ramakrishnan. “The Role of the Granuloma in Expansion and Dissemination of Early Tuberculous Infection”. In: *Cell* 136.1 (2009), pp. 37–49.

[144] C.-L. Tran, A. D. Jones, and K. Donaldson. “Mathematical Model of Phagocytosis and Inflammation after the Inhalation of Quartz at Different Concentrations”. In: *Scandinavian Journal of Work, Environment & Health* 21 (1995), pp. 50–54. ISSN: 0355-3140.

[145] R. I. Kavet and J. D. Brain. “Phagocytosis: Quantification of Rates and Intercellular Heterogeneity”. In: *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology* 42.3 (1977), pp. 432–437. DOI: `10.1152/jappl.1977.42.3.432`.

[146] R. V. Carvalho et al. “Modeling Innate Immune Response to Early *Mycobacterium* Infection”. In: *Computational and Mathematical Methods in Medicine* 2012 (2012), pp. 1–12. DOI: `10.1155/2012/790482`.

[147] S. Marino et al. “TNF and IL-10 Are Major Factors in Modulation of the Phagocytic Cell Environment in Lung and Lymph Node in Tuberculosis: A next-Generation Two-Compartmental Model”. In: *Journal of Theoretical Biology* 265.4 (2010), pp. 586–598. DOI: `10.1016/j.jtbi.2010.05.012`.

[148] D. Gammack, C. Doering, and D. Kirschner. “Macrophage Response to Mycobacterium Tuberculosis Infection”. In: *Journal of Mathematical Biology* 48.2 (2003), pp. 218–242. DOI: `10.1007/s00285-003-0232-8`.

[149] J. E. Wigginton and D. Kirschner. “A Model to Predict Cell-Mediated Immune Regulatory Mechanisms during Human Infection with Mycobacterium Tuberculosis”. In: *The Journal of Immunology* 166.3 (2001), pp. 1951–1967.

[150] *Kirschner Lab*. http://malthus.micro.med.umich.edu/lab/index.html.

[151] S. Marino and D. E. Kirschner. “The Human Immune Response to Mycobacterium Tuberculosis in Lung and Lymph Node”. In: *Journal of Theoretical Biology* 227.4 (2004), pp. 463–486. DOI: `10.1016/j.jtbi.2003.11.023`.

[152] S. Marino et al. “Dendritic Cell Trafficking and Antigen Presentation in the Human Immune Response to Mycobacterium Tuberculosis”. In: *The Journal of Immunology* 173.1 (2004), pp. 494–506. DOI: `10.4049/jimmunol.173.1.494`.

[153] D. Kirschner and S. Marino. “Mycobacterium Tuberculosis as Viewed through a Computer”. In: *Trends in Microbiology* 13.5 (2005), pp. 206–211. DOI: `10.1016/j.tim.2005.03.005`.

[154] J. Day, A. Friedman, and L. S. Schlesinger. “Modeling the Immune Rheostat of Macrophages in the Lung in Response to Infection”. In: *Proceedings of the National Academy of Sciences* 106.27 (2009), pp. 11246–11251. DOI: `10.1073/pnas.0904846106`.

[155] D. Sud et al. “Contribution of CD8+ T Cells to Control of Mycobacterium Tuberculosis Infection”. In: *The Journal of Immunology* 176.7 (2006), pp. 4296–4314. DOI: `10.4049/jimmunol.176.7.4296`.

[156] S. Marino et al. “Differences in Reactivation of Tuberculosis Induced from Anti-TNF Treatments Are Based on Bioavailability in Granulomatous Tissue”. In: *PLoS Computational Biology* 3.10 (2007). Ed. by D. Wodarz, e194. DOI: `10.1371/journal.pcbi.0030194`.

[157] E. Pienaar and M. Lerm. “A Mathematical Model of the Initial Interaction between Mycobacterium Tuberculosis and Macrophages”. In: *Journal of Theoretical Biology* 342 (2013), pp. 23–32. DOI: `10.1016/j.jtbi.2013.09.029`.

[158] W. Hao, L. S. Schlesinger, and A. Friedman. “Modeling Granulomas in Response to Infection in the Lung”. In: *PloS One* 11.3 (2016), e0148738. DOI: `10.1371/journal.pone.0148738`.

[159] D. Gammack et al. “Understanding the Immune Response in Tuberculosis Using Different Mathematical Models and Biological Scales”. In: *Multiscale Modeling & Simulation* 3.2 (2005), pp. 312–345. DOI: `10.1137/040603127`.

[160] S. Marino et al. “A Methodology for Performing Global Uncertainty and Sensitivity Analysis in Systems Biology”. In: *Journal of Theoretical Biology* 254.1 (2008), pp. 178–196. DOI: `10.1016/j.jtbi.2008.04.011`.

[161] J. C. J. Ray, J. L. Flynn, and D. E. Kirschner. “Synergy between Individual TNF-Dependent Functions Determines Granuloma Performance for Controlling Mycobacterium Tuberculosis Infection”. In: *The Journal of Immunology* 182.6 (2009), pp. 3706–3717. DOI: `10.4049/jimmunol.0802297`.

[162] M. Fallahi-Sichani et al. “Multiscale Computational Modeling Reveals a Critical Role for TNF-α Receptor 1 Dynamics in Tuberculosis Granuloma Formation”. In: *The Journal of Immunology* 186.6 (2011), pp. 3472–3483. DOI: `10.4049/jimmunol.1003299`.

[163] N. A. Cilfone et al. “Multi-Scale Modeling Predicts a Balance of Tumor Necrosis Factor-α and Interleukin-10 Controls the Granuloma Environment during Mycobacterium Tuberculosis Infection”. In: *PLOS ONE* 8.7 (2013), e68680. DOI: `10.1371/journal.pone.0068680`.

[164] S. Marino et al. “Macrophage Polarization Drives Granuloma Outcome during Mycobacterium Tuberculosis Infection”. In: *Infection and Immunity* (2014), IAI.02494–14. DOI: `10.1128/IAI.02494-14`.

[165] D. Levin. *Cycells: Academic Release.* Zenodo. 2014. DOI: `10.5281/zenodo.13296`.

[166] R. Bowness et al. “Modelling the Effects of Bacterial Cell State and Spatial Location on Tuberculosis Treatment: Insights from a Hybrid Multiscale Cellular Automaton Model”. In: *Journal of Theoretical Biology* 446 (2018), pp. 87–100. DOI: `10.1016/j.jtbi.2018.03.006`.

[167] M. J. Pitcher et al. “A Spatially Heterogeneous Network-Based Metapopulation Software Model Applied to the Simulation of a Pulmonary Tuberculosis Infection”. In: *Applied Network Science* 3.1 (2018), p. 33. DOI: `10.1007/s41109-018-0091-2`.

[168] C. Aston et al. “Early Inhibition of Mycobacterial Growth by Human Alveolar Macrophages Is Not Due to Nitric Oxide”. In: *American Journal of Respiratory and Critical Care Medicine* 157.6 (1998), pp. 1943–1950. DOI: `10.1164/ajrccm.157.6.9705028`.

[169] N. Hens et al. “Seventy-Five Years of Estimating the Force of Infection from Current Status Data”. In: *Epidemiology and Infection* 138.6 (2010), pp. 802–812. DOI: `10.1017/S0950268809990781`.

[170] D. M. Mosser and J. P. Edwards. “Exploring the Full Spectrum of Macrophage Activation”. In: *Nature Reviews. Immunology* 8.12 (2008), pp. 958–969. DOI: `10.1038/nri2448`.

[171] N. S. Goel, S. C. Maitra, and E. W. Montroll. *On the Volterra and Other Nonlinear Models of Interacting Populations.* New York: Academic Press, 1971. ISBN: 978-0-12-287450-5.

[172] M. Iannelli and A. Pugliese. *An Introduction to Mathematical Population Dynamics.* Vol. 79. UNITEXT. Cham: Springer International Publishing, 2014. DOI: `10.1007/978-3-319-03026-5`.

[173] M. J. Pitcher. “In Silico Modelling of In-Host Tuberculosis Dynamics : Towards Building the Virtual Patient”. Thesis. University of St Andrews, 2020.

[174] J. M. Mouton et al. “Elucidating Population-Wide Mycobacterial Replication Dynamics at the Single-Cell Level”. In: *Microbiology* 162.6 (2016), pp. 966–978. DOI: `10.1099/mic.0.000288`.

[175] M. D. McKay. “Latin Hypercube Sampling as a Tool in Uncertainty Analysis of Computer Models”. In: *Proceedings of the 24th Conference on Winter Simulation - WSC ’92.* Arlington, Virginia, United States: ACM Press, 1992, pp. 557–564. DOI: `10.1145/167293.167637`.

[176] R. M. Wood, J. R. Egan, and I. M. Hall. “A Dose and Time Response Markov Model for the In-Host Dynamics of Infection with Intracellular Bacteria Following Inhalation: With Application to *Francisella Tularensis*”. In: *Journal of The Royal Society Interface* 11.95 (2014), p. 20140119. DOI: `10.1098/rsif.2014.0119`.

[177] E. Sherer et al. “Identification of Age-Structured Models: Cell Cycle Phase Transitions”. In: *Biotechnology and Bioengineering* 99.4 (2008), pp. 960–974. DOI: `10.1002/bit.21633`.

[178] G. G. Steel. *Growth Kinetics of Tumours: Cell Population Kinetics in Relation to the Growth and Treatment of Cancer.* Oxford [Eng.]: Clarendon Press, 1977. ISBN: 978-0-19-857388-3.

[179] C. A. Yates, M. J. Ford, and R. L. Mort. “A Multi-stage Representation of Cell Proliferation as a Markov Process”. In: *Bulletin of Mathematical Biology* 79.12 (2017), pp. 2905–2928. DOI: `10.1007/s11538-017-0356-4`.

[180] E. Gavagnin et al. “The Invasion Speed of Cell Migration Models with Realistic Cell Cycle Time Distributions”. In: *Journal of Theoretical Biology* 481 (2019), pp. 91–99. DOI: `10.1016/j.jtbi.2018.09.010`.

[181] H. X. Chao et al. “Evidence That the Human Cell Cycle Is a Series of Uncoupled, Memoryless Phases”. In: *Molecular Systems Biology* 15.3 (2019). DOI: `10.15252/msb.20188604`.

[182] A. Grami. *Probability, Random Variables, Statistics, and Random Processes: Fundamentals & Applications.* Hoboken, NJ, USA: John Wiley & Sons, Inc., 2019. DOI: `10.1002/9781119300847`.

[183] L. Gelens and S. D. Santos. “Eternal Sunshine of the Spotless Cycle”. In: *Molecular Systems Biology* 15.4 (2019). DOI: `10.15252/msb.20198864`.

[184] J. Carruthers et al. “A Novel Stochastic Multi-Scale Model of Francisella Tularensis Infection to Predict Risk of Infection in a Laboratory”. In: *Frontiers in Microbiology* 9 (2018), p. 1165. DOI: `10.3389/fmicb.2018.01165`.

[185] G. Mazo. “The Sobol Method in Sensitivity Analysis for Stochastic Computer Models”. hal- 02113448v1, 2019.

[186] J. Carruthers et al. “Stochastic Dynamics of Francisella Tularensis Infection and Replication”. In: *PLOS Computational Biology* 16.6 (2020). Ed. by J. O’Dwyer, e1007752. DOI: `10.1371/journal.pcbi.1007752`.

[187] E. D. Hawkins et al. “A Model of Immune Regulation as a Consequence of Randomized Lymphocyte Division and Death Times”. In: *Proceedings of the National Academy of Sciences* 104.12 (2007), pp. 5032–5037. DOI: `10.1073/pnas.0700026104`.

[188] H. Miao et al. “Evaluation of Multitype Mathematical Models for CFSE-Labeling Experiment Data”. In: *Bulletin of Mathematical Biology* 74.2 (2012), pp. 300–326. DOI: `10.1007/s11538-011-9668-y`.