



# Parameter estimation and sensitivity analysis in an agent-based model of *Leishmania major* infection<sup>☆</sup>

Garrett M. Dancik<sup>d,\*</sup>, Douglas E. Jones<sup>c</sup>, Karin S. Dorman<sup>a,b,d</sup>

<sup>a</sup> Department of Statistics, Iowa State University, Ames, IA 50011, USA

<sup>b</sup> Genetics, Development and Cell Biology, Rm 534, Science II, Iowa State University, Ames, IA 50011, USA

<sup>c</sup> Veterinary Pathology, 2764 Vet. Med., Iowa State University, Ames, IA 50011, USA

<sup>d</sup> Program in Bioinformatics and Computational Biology, Rm 533, Science II, Iowa State University, Ames, IA 50011, USA

## ARTICLE INFO

### Article history:

Received 19 January 2009

Received in revised form

6 October 2009

Accepted 6 October 2009

Available online 23 October 2009

### Keywords:

Calibration

Sensitivity analysis

*Leishmania major*

Agent-based model

## ABSTRACT

Computer models of disease take a systems biology approach toward understanding host–pathogen interactions. In particular, data driven computer model calibration is the basis for inference of immunological and pathogen parameters, assessment of model validity, and comparison between alternative models of immune or pathogen behavior. In this paper we describe the calibration and analysis of an agent-based model of *Leishmania major* infection. A model of macrophage loss following uptake of necrotic tissue is proposed to explain macrophage depletion following peak infection. Using Gaussian processes to approximate the computer code, we perform a sensitivity analysis to identify important parameters and to characterize their influence on the simulated infection. The analysis indicates that increasing growth rate can favor or suppress pathogen loads, depending on the infection stage and the pathogen's ability to avoid detection. Subsequent calibration of the model against previously published biological observations suggests that *L. major* has a relatively slow growth rate and can replicate for an extended period of time before damaging the host cell.

© 2009 Elsevier Ltd. All rights reserved.

## 1. Introduction

*Leishmania* are protozoan parasites that are transmitted by bites of infected sandflies. The macrophage is the primary host cell. Over 20 species of *Leishmania*, endemic in 88 countries, are capable of causing human disease. Disease is either cutaneous, where skin ulcers occur on exposed surfaces of the body, or visceral, with near certain mortality if left untreated. In mice that are able to control *Leishmania major* infection, the resistance mechanism is well understood: secretion of IL-12 by dendritic cells promotes a CD4<sup>+</sup> Th1 response, Th1 cells activate macrophages through IFN- $\gamma$  production, and activated macrophages clear the parasite. However, there are other species of the *Leishmania* parasite, such as *L. amazonensis*, which cause chronic disease in mice (McMahon-Pratt and Alexander, 2004), and the development of successful vaccines for any *Leishmania* species has remained elusive (Vanloubbeec and Jones, 2004).

Computer models of disease take a systems biology approach toward understanding adverse or inefficient immune responses by integrating multiple sources of knowledge about host–pathogen interactions and immune cell function in order to study the collective, emergent behavior of a population of immune cells, i.e., the immune response. Such computer models have been used to gain insight into a variety of diseases. For example, in a model of *Mycobacterium tuberculosis* infection, Segovia-Juarez et al. (2004) identify chemokine diffusion rates and the arrival time, location, and macrophage activation efficiency of T cells as important factors in granuloma formation. In a model of systemic inflammatory response and multiple organ failure, An (2002) reproduces outcomes of unsuccessful clinical trials involving blockage of proinflammatory mediators. In a model of influenza A infection, Beauchemin (2006) shows how infection dynamics depend on the spatial structure of initially infected cells. In a model of Epstein-Barr virus, Shapiro et al. (2008) identify lytic reactivation of B cells as an important parameter that determines disease outcome. Finally, in a model of antigen escape in HIV infection, Bernaschi and Castiglione (2002) find that escape mutants with low transcription rate can explain the long-term asymptomatic phase of disease. For a recent review of computational immune system models, see Forrest and Beauchemin (2007).

A challenge faced when working with computer models is the need to choose values for model parameters. Typically, plausible

<sup>☆</sup> Funded by NIH

\* Corresponding author at: Present address: Northwestern State University, Departments of Biological Sciences and Mathematics, 207 Bienvenu Hall, Natchitoches, LA 71497, USA. Tel.: +1 318 357 4485; fax: +1 318 357 4518.

E-mail addresses: [dancikg@nsula.edu](mailto:dancikg@nsula.edu) (G.M. Dancik), [jonesdou@iastate.edu](mailto:jonesdou@iastate.edu) (D.E. Jones), [kdorman@iastate.edu](mailto:kdorman@iastate.edu) (K.S. Dorman).

choices for parameters are determined through literature searches and expert consultation, yet often the best available information are plausible ranges for model parameters rather than single values. In order to validate the model, one must ultimately fit the computer output to field data, choosing parameter values that yield the best match between simulation output and biological observations. Model calibration has proven to be difficult in practice, especially since most computer models are high dimensional, non-linear, and resource-intensive. As a result, computer modelers traditionally employ *ad-hoc* approaches to parameter estimation (Kennedy and O'Hagan, 2001), where model validation is based on the qualitative comparison of model predictions with field data. In the field of population ecology, this approach is known as pattern-oriented modeling (reviewed in Grimm et al., 2005).

Nevertheless, data-driven model calibration is a fundamental and necessary step for making meaningful inference on model parameters, assessing the validity of a model, and performing model comparisons (Bayarri et al., 2002). The calibration process first requires the specification of the relationship between the computer model output and data that is observed in the field. When the computer model is expensive to run, the process then entails (1) approximation of the computer code through a surrogate statistical model (i.e., a Gaussian process) and (2) estimation of computer model parameters via the surrogate model. The Gaussian process approximation also facilitates sensitivity analysis which identifies important model parameters and the relationships between parameters and simulation output. Schonlau and Welch (2006), for example, use a Gaussian process emulator of a computer model of global human development to identify important parameters and characterize their effects. These sensitivity analysis tools are particularly useful for computer models, such as agent-based and cellular automata models, where it is not clear how the computer code itself can be exploited, without running the model, to quantitatively understand the system.

Computer model calibration through a Gaussian process intermediate has been demonstrated successfully in several disciplines. For example, Kennedy and O'Hagan (2001) calibrate a computer model of radionuclide exposure following a chemical accident; Bayarri et al. (2006, 2002) calibrate a vehicular suspension system model and a vehicle collision model; Higdon et al. (2004) calibrate a spot welding model; and Heitmann et al. (2006) calibrate a cosmological model. A recent formulation of the computer model calibration and validation approach is provided in Bayarri et al. (2007).

In this paper we describe the sensitivity analysis and calibration of an agent-based model of *Leishmania major* infection. A model of macrophage loss triggered by necrotic tissue production is proposed for explaining macrophage depletion after peak infection. We find that pathogen growth rate and host cell carrying capacity both affect macrophage levels early in infection, though not independently. Increasing parasite growth rate can both augment and paradoxically, suppress parasite loads, depending on the stage of infection and the ability of the pathogen to avoid detection. Furthermore, the ability of the pathogen to evade the adaptive immune response has a large effect on macrophage levels at 6.5 wpi. We verify that parameter estimation using the Gaussian process intermediate is accurate by calibrating the computer model using simulated field data, and then calibrate our model using field data from Belkaid et al. (2000). Parameter estimates suggest that intracellular pathogen replicates extensively before spreading to additional cells, a finding consistent with observations of *Leishmania* growth in cultured macrophages (Chang et al., 2002). Finally, model comparison supports our proposal of explicit macrophage loss in response to necrotic tissue production.

The rest of this paper is organized as follows: in Section 2 we describe an agent-based model of *Leishmania* infection in mice; in Section 3 we describe the statistical methods used in Gaussian process computer model approximation (Section 3.1), sensitivity analysis (Section 3.2), calibration (Section 3.3), implementation via Markov chain Monte Carlo (Section 3.4), and model comparison (Section 3.5); in Section 4 we report results obtained by applying these statistical methods to the agent-based model; and in Section 5 we discuss our findings.

## 2. An agent-based model of *Leishmania major* infection

Agent-based models (ABMs)<sup>1</sup> capture the dynamics of complex systems whose properties depend on the collective behavior of their interacting components. An ABM contains distinct entities, or *agents*, that inhabit a *spatial environment*. A simulation visualizes agents as they move and interact according to *update rules* executed at *discrete time steps* (Wooldridge, 2002).

We describe an ABM of the immune response to *L. major* infection in the ear of a C57BL/6 mouse, a system where experimental data are available (Belkaid et al., 2000). Our computer model is stochastic, so that multiple simulations will give different results even when the same model parameters and starting conditions are used. The current model is an extension of an earlier ABM of *L. major* infection (Dancik et al., 2006), and is largely based on the work of Segovia-Juarez et al. (2004), who explore granuloma formation during infection with another macrophage-tropic parasite, *Mycobacterium tuberculosis*. All model parameters are listed in Table C1 of the Supplement.

### 2.1. The model environment

We model a 2 mm × 2 mm cross section of the ear as a 100 × 100 lattice of square *micro-compartments*. This lattice represents an internal portion of a uniform infection area, and is therefore *toroidal*, so an object leaving the lattice will re-enter at the opposite end. We select four evenly distributed micro-compartments to serve as *source compartments* where new cells enter the system.

We label micro-compartments  $(i, j)$ , starting from (0,0) at the bottom left. Define a *Moore neighborhood* of length  $r$  at position  $(x_0, y_0)$  to be the local collection of points in the lattice given by

$$M_r(x_0, y_0) = \{(x, y) : |x - x_0| \leq r \text{ and } |y - y_0| \leq r\}.$$

Furthermore, define  $M_1(x_0, y_0)$  to be the *immediate Moore neighborhood* of a micro-compartment  $(x_0, y_0)$ .

### 2.2. Stages of infection

#### 2.2.1. Initial conditions

We take the starting point of our simulation to be 3.5 weeks post infection (wpi) of the experiment described in Belkaid et al. (2000), near the beginning of the dramatic phase of parasite growth. Data for the number of infected macrophages at this time point are not available. However, because pathogen load dynamics during the dramatic phase of parasite growth are similar to pathogen load dynamics following high dose infection (Lira et al., 2000), we assume that conditions at the start of our simulation match conditions just two days post high dose infection. Based on parasite and expected cell counts observed two days post high-

<sup>1</sup> Abbreviations used in the paper: Agent-based model (ABM), weeks post infection (wpi), Gaussian process (GP), sensitivity analysis (SA), principle components (PC), functional analysis of variance (FANOVA).

dose infection (Doug Jones, unpublished observations), we randomly place 105 macrophages on the lattice. A random number of parasites between 1 and  $P_m$  infect randomly selected resting macrophages until total pathogen load on the lattice exceeds  $50 - P_m$ . A final, random resting macrophage receives  $50 - P_m$  parasites, so the initial pathogen load is 50. All uninfected macrophages are given random life spans uniform between 0 and 100 days (Furth et al., 1973).

### 2.2.2. Infection of macrophages

Macrophages are the primary cells that *L. major* parasites infect (Belkaid et al., 2000) and the only host cells that support pathogen replication (Naderer et al., 2006). We use the term *resting* macrophage to refer to an uninfected, unactivated cell. We assume that intracellular parasites experience logistic growth at rate  $\alpha_i$  with carrying capacity  $k_i + 30$ . If the macrophage is not activated, intracellular parasites grow until their number exceeds the lethal parasite density  $k_i$ . Then, the macrophage enters a *dying* state, where it begins transferring parasite to macrophages in its length two Moore neighborhood. Since extracellular parasites are seldom observed in *Leishmania* infection (Chang et al., 2002), parasite transmission is thought to occur when nearby macrophages phagocytize dying host cells. Segovia-Juarez et al. (2004) use a similar model, but allow extracellular parasite and restrict parasite uptake to macrophages in their length 1 Moore neighborhood. All macrophages that are not dying, including those already infected, can take up parasite. A dying macrophage is removed from the system once all of its intracellular parasites are consumed by surrounding macrophages.

### 2.2.3. Chemokines and cell movement

Chemokines, chemical attractants that influence cell movement, play an important role in *Leishmania* infection (Roychoudhury and Roy, 2004). We include one generic chemokine as an attractant for both macrophages and T cells. Its diffusion and decay properties are based on interleukin-8, an important chemokine involved in early infection. Infected and activated macrophages release  $c_i$  units of chemokine per time step. In addition to directing cell movement toward infected areas (see below), chemokine also increases the rate of macrophage recruitment (Section 2.2.4).

Cell movement has been described as a biased-random walk in the presence of chemokine (Tranquillo et al., 1988), and we model this random walk on the ABM lattice. Let

$C_{i,j}$  = amount of chemokine in micro-compartment  $(i,j)$ ,

$C_{M_1}(i,j)$  = amount of chemokine in immediate Moore neighborhood  $M_1(i,j)$ .

A cell currently in micro-compartment  $(i,j)$  moves to micro-compartment  $(k,l) \in M_1(i,j)$ ,  $(k,l) \neq (i,j)$  with probability  $C_{k,l}/(C_{M_1}(i,j) - C_{i,j})$ . All cells in our model move in this way, but at different speeds which depend on cell type and state (Section 2.2.7 and Table C1).

### 2.2.4. Recruitment and the adaptive immune response

T cells and macrophages are recruited to infected areas during infection. In the ABM, source compartments represent blood vessels where recruited cells arrive. At each time-step, macrophages enter with probability based on source compartment chemokine levels (see below), and T cells, after a delay, enter with constant probability  $p_{Trecr}$ .

**Macrophage recruitment in response to chemokine:** Macrophages are actively recruited around two days post infection following high-dose *L. major* inoculation (Sunderkotter et al., 1993), and after a delay in low dose infection (Belkaid et al., 2000). This recruitment is modulated by chemokines such as macrophage chemotactic and activating factor (Badolato et al., 1996). Certain

chemokines trigger the activation of endothelial cells and upregulation of their monocyte receptors, which facilitate extravasation of cells from the circulatory system into infected tissue. Other chemokines, such as granulocyte-macrophage colony-stimulating factor, upregulate macrophage production from the bone marrow (Janeway et al., 2005). Therefore, chemokines collectively increase the amount of circulating macrophages as well as the rate of influx of these macrophages into the tissue.

We use a logistic function to model macrophage recruitment in response to the generic chemokine in our model. This function captures initial exponential increase in macrophage numbers due to macrophage production and influx, as well as the eventual loss of chemokine potency at high chemokine concentrations due to the saturation of chemokine receptors (Rajotte et al., 1997). Let

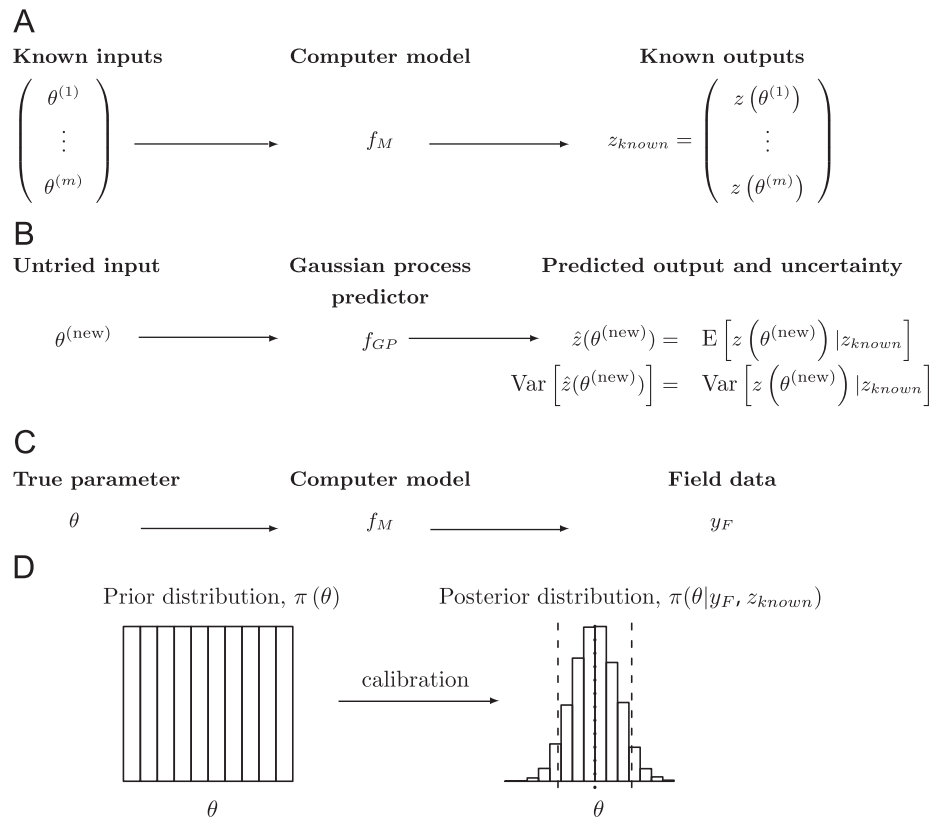
$$p_{Mrecr,i}(x) = \begin{cases} \frac{m_{Mrecr} \times \exp(\alpha_{Mrecr} + \beta_{Mrecr} \times x)}{1 + \exp(\alpha_{Mrecr} + \beta_{Mrecr} \times x)}, & x > C_M, \\ 1.74 \times 10^{-5}, & x \leq C_M, \end{cases} \quad (1)$$

where  $p_{Mrecr,i}(x)$  is the probability that a macrophage is recruited when there are  $x$  units of chemokine at source compartment  $i$ . The parameter  $m_{Mrecr}$  represents the maximum recruitment probability,  $\alpha_{Mrecr}$  determines the minimum probability of recruitment as  $x \rightarrow 0$ , and  $\beta_{Mrecr}$  determines how sharply recruitment increases in response to chemokine. This *active* macrophage recruitment mechanism is used as long as  $x$  is above a chemokine threshold,  $C_M$ ; otherwise we use a base level of recruitment that will maintain macrophage equilibrium in the absence of infection. In the experimental infection we are modeling, macrophage levels are initially constant (i.e., due to a base level of recruitment), then dramatically increase after 3.5 wpi as active recruitment begins (see Belkaid et al., 2000, Fig. 1B). Because the starting point of our simulation is 3.5 wpi of the experiment described in Belkaid et al. (2000), we set  $C_M$  dynamically to correspond to source compartment chemokine levels after one day of simulation. On average,  $C_M$  is assigned a value of approximately 12,000 chemokine units.

**The adaptive immune response:** During infection, antigen-presenting cells take up pathogen from the site of infection, migrate to the draining lymph node, and present antigen (processed pathogen) to naïve T cells. Naïve T cells, in response, proliferate and mature into several classes of T cells, including CD4+ Th1 cells. In *L. major* infection, CD4+ Th1 cells migrate to the infected area and activate infected macrophages. The initiation of the adaptive immune response requires a threshold level of antigen (Janeway et al., 2005) and occurs only after a prolonged period of parasite growth in low dose *L. major* infection (Belkaid et al., 2000). We assume that this threshold level of antigen is related to pathogen load at the infection site, and recruit T cells once a threshold pathogen level,  $T_{threshold}$ , is achieved. The threshold is a one way trigger so T cells will continue to enter at each time step and at each source compartment with probability  $p_{Trecr}$  as long as source compartment chemokine levels are greater than a threshold of  $C_T$  chemokine units.

### 2.2.5. Macrophage emigration following uptake of necrotic tissue

In *L. major* infection, there is a sharp decrease in macrophage levels near the time that the T cell response reaches its peak (Belkaid et al., 2000). In our model, this decrease cannot be explained merely from the death of infected and activated macrophages (data not shown). We propose a model of macrophage loss based on an experiment which shows that as the inflammatory response resolves, macrophages systematically migrate to the draining lymph node (Bellingan et al., 1996), an event known to occur following macrophage uptake of foreign antigen, apoptotic immune cells, and necrotic tissue (Winchester et al., 1984). In the ABM, macrophage emigration is triggered by



**Fig. 1.** Overview of Gaussian process-based computer model prediction and calibration. (A) The computer model is evaluated at  $m$  representative computer model parameters  $\theta^{(i)}, i = 1, \dots, m$ , to produce the output vector  $z_{known}$ . (B) Using the GP predictor, which is fit to  $z_{known}$ , the predicted value at untried input  $\theta^{(new)}$  is given by  $\hat{z}(\theta^{(new)})$  while  $\text{Var}[\hat{z}(\theta^{(new)})]$  quantifies uncertainty in the prediction. (C) Calibration is the process of estimating the computer model parameter (vector)  $\theta$  so that computer model output is consistent with the field data vector  $y_F$ . (D) Calibration of a single computer model parameter  $\theta$  is shown. The calibration process yields a posterior distribution  $\pi(\theta | y_F, z_{known})$  that reflects current knowledge about  $\theta$  based on the field data vector while considering various sources of uncertainty, including measurement error, computer model approximation using the GP predictor, and the prior distribution  $\pi(\theta)$ . The posterior median (dotted line) is used as a posterior estimate of  $\theta$  while 90% posterior credible intervals (dashed lines) quantify uncertainty in the posterior estimate.

the production and uptake of necrotic tissue, a side effect of macrophage activation (Billack, 2006). Subsequent to macrophage activation,  $A_{necr}$  units of non-diffusive necrotic tissue are released. Uninfected macrophages in the length two Moore neighborhood consume one unit of this tissue and disappear from the lattice, representing rapid migration to the draining lymph node.

#### 2.2.6. Macrophage activation

All T cells in our model are considered to be antigen specific CD4+ Th1 cells that are equally capable of activating infected macrophages. Macrophage activation is cell-mediated (Sypek et al., 1984), and T cells within the immediate Moore neighborhood of an infected macrophage activate it with probability  $T_{actm}$ . In *L. major* infection, macrophage activation is sufficient for elimination of intracellular parasite (Wei et al., 1995).  $M_{als}$  days after T cell activation, the macrophage finishes destroying all intracellular parasite, undergoes apoptosis, produces  $A_{necr}$  units of necrotic tissue, and is removed from the lattice.

#### 2.2.7. Time scales

Each time step in the ABM is equal to approximately six seconds of real time. Chemokine diffusion and decay as well as parasite growth occur each time step. Following Segovia-Juarez et al. (2004), T cells move every 200 time steps (20 min) and macrophages move on slower time scales that vary according to macrophage state (see Table C1). Additional update rules, such as the take up of parasite and activation of macrophages, occur every 10 min.

#### 2.3. Implementation

We program the ABM using C++ to encode the spatial environment, rules, and entities described above. All parameters are fixed or allowed to vary according to Table C1. Starting with the initial conditions described in Section 2.2.1, the ABM proceeds to simulate *L. major* infection according to the above rules. With initial conditions representing 3.5 wpi, we simulate infection until 8.5 wpi. Each simulation takes approximately 15 min on an Intel Xeon 3.00 GHz workstation.

### 3. Statistical methods

Our primary objective is to estimate the computer model parameter vector  $\theta = (\theta_1, \dots, \theta_p)$  ( $p$  parameters from Table C1) given a vector  $y_F$  of biological field observations. The calibration process (see Section 3.3) requires a large number of computer model evaluations to explore the parameter space. Although we limit our attention to the agent-based model described in Section 2, the methods we describe are appropriate for the analysis and calibration of a wide range of mathematical and computational models, including ordinary differential equation models. For models that are expensive, in money or time, the number of evaluations is limited, and calibration demands a fast and accurate emulator of the computer code. The type of emulator that we use is the Gaussian process (GP). In addition to its role in parameter calibration, the GP emulator is utilized in sensitivity analysis to explore the influence of parameters on the simulated response. A



schematic overview of GP prediction and calibration is given in Fig. 1, and a listing of available GP-based analyses is given in Table 1. In this section we give a brief, mostly conceptual overview of the statistical methods we use for GP prediction of scalar computer model output, sensitivity analysis, computer model calibration, and model comparison, with particular attention paid to notions relevant to our ABM. Throughout, we refer the reader to subsections in Supplement A for additional details. The reader is also encouraged to consult the cited references for more background and theory.

3.1. The Gaussian process as a computer model surrogate

GPs are flexible and useful tools for approximating computer models when the computer model response surface of interest is a smooth function of the parameter space (Sacks et al., 1989;

Kennedy and O'Hagan, 2001; Bayarri et al., 2002). We start by obtaining a set of  $m$  known scalar computer model outputs, for example macrophage counts or pathogen load at a particular day in the disease progression, by running the model at known parameter settings  $\theta^{(1)}, \dots, \theta^{(m)}$  selected by a maximin Latin Hypercube design (Morris and Mitchell, 1995). Call this data  $z_{\text{known}} = (z(\theta^{(1)}), \dots, z(\theta^{(m)}))$ . These are scalar data because there is a single output for each parameter vector input. The GP approximation is built by fitting a multivariate normal distribution to  $z_{\text{known}}$ , assuming the correlation structure depends on the parameter vector  $\theta$ . Computer model output at an untried parameter  $\theta^{(\text{new})}$ , conditioned on  $z_{\text{known}}$ , is predicted using standard multivariate normal distribution theory (Renchner, 2002). The predicted output  $\hat{z}(\theta^{(\text{new})})$  is given by  $E[z(\theta^{(\text{new})})|z_{\text{known}}]$  while  $\text{Var}[\hat{z}(\theta^{(\text{new})})] = \text{Var}[z(\theta^{(\text{new})})|z_{\text{known}}]$  quantifies uncertainty in the prediction (Figs. 1B, 2). A simple example of GP prediction of

Table 1  
Gaussian-process based analyses.

	Analysis methods	Method description
Sensitivity analysis	FANOVA decomposition	Section 3.2
	Main effects plots	Section 3.2, Supplement A.4
	Interaction effects plots	Section 3.2
Computer model validation	Calibration	Section 3.3, Supplement A.5
	Model comparison	Section 3.5

The fitted GP is used in sensitivity analysis, which identifies important parameters through FANOVA decomposition and characterizes parameter effects on model output through main and interaction effects plots; and in computer model validation, which involves calibration (i.e., parameter estimation) and model comparison. For each type of analysis, we provide references in our manuscript where methods are described.

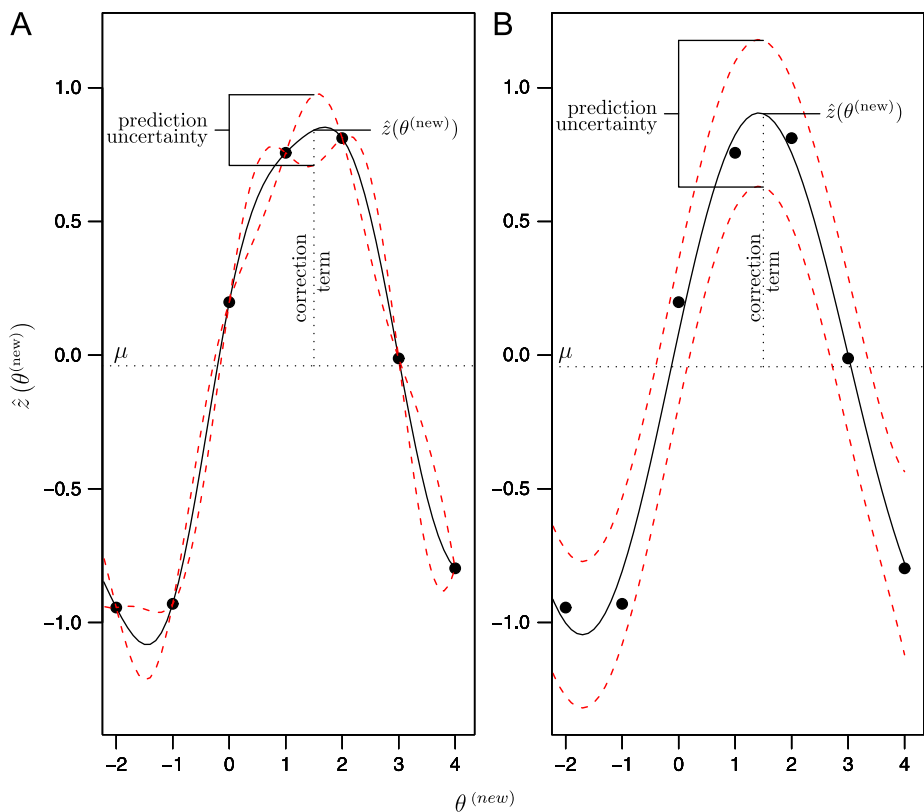


Fig. 2. Gaussian process prediction and uncertainty. Gaussian process prediction (solid black lines) of the function  $z(\theta^{(i)}) = \sin(\theta^{(i)}) + \varepsilon_i$ , where  $\varepsilon_i$  are iid  $N(0, 0.02)$  residuals. To prepare these plots, we used the fitted  $f_{GP}$  to make predictions at seven equally spaced values  $\theta^{(\text{new})}$  from  $-2$  to  $4$ . A. Without a nugget term, the GP predictor interpolates known computer model output (black circles). B. With a nugget term, the GP predictor interpolates and smooths the known data set. In either case, GP prediction of output at the parameter  $\theta^{(\text{new})}$  is given by  $\hat{z}(\theta^{(\text{new})}) = E[z(\theta^{(\text{new})})|z_{\text{known}}]$ , which takes the unconditional mean  $\mu$  (horizontal dotted line) and adds a correction term (vertical dotted line) that accounts for the correlation between outputs. Uncertainty in the GP prediction is quantified using  $\text{Var}[z(\theta^{(\text{new})})|z_{\text{known}}]$ . In the figure, uncertainty bounds (dashed lines) correspond to predicted output  $\pm 3$  standard deviations. For details about GP prediction, see Supplement A.1.

scalar output is provided in Fig. 2. If the computer model is *deterministic*, then the predictor will interpolate the known output  $z_{known}$  (Fig. 2A). If the computer model is *stochastic*, as is the case with our ABM, the GP includes an extra variance term, called a *nugget*, so the predicted output interpolates and smooths the known output  $z_{known}$  (Fig. 2B). For details see Supplement A.1. The accuracy of the GP model is assessed via a cross-validation procedure, described in Supplement A.2. More on the choice of parameter vector inputs  $\theta^{(i)}$  is discussed in Supplement A.7.

For models more complex than the sin wave of Fig. 2, computer output is typically multivariate. It is possible to fit one GP to each component of the data, but this solution becomes impractical for high dimensional functional data such as time series. For example, we can record the pathogen load every day of our simulation, a total of 35 measurements between 3.5 and 8.5 wpi. GP prediction of functional data involves using singular value decomposition to reduce the dimensionality of the output and fitting GPs to the most important principle component (PC) weights. In practice, we fit GPs to the most important PC weights that together account for at least 99% of the functional variance. For details on GP prediction of functional output, see Supplement A.3.

### 3.2. Sensitivity analysis using the Gaussian process

Sensitivity analysis (SA) includes a broad range of statistical techniques for measuring a parameter's ability to influence the output of a model. Typically, sensitivity analysis is used in *factor screening* methods for identifying the most important parameters of a system, as well as for assessing how the response of a model depends on its input parameters (Saltelli et al., 2000). Some of these methods have been adapted to ABMs (Marino et al., 2008), but here, we focus on sensitivity analysis techniques that capitalize on the functional GP approximation, following Schonlau and Welch (2006). Of interest is how the computer model parameter vector  $\theta = (\theta_1, \dots, \theta_p)$  influences computer model output. Inferences are based on a prior distribution  $\pi(\theta)$ , meant to capture lack of knowledge about model parameters, and the observed computer model output vector  $z_{known}$ . In our SA we use independent  $\text{Unif}(0,1)$  priors on all components of  $\theta$  scaled from the ranges given in Table C1.

For an appealing, but relatively qualitative visual assessment of computer model parameters on output, one can estimate the *main effect* of each parameter. First, we discuss main effects on scalar computer output. For  $\theta_k$ , the  $k$ th component of the parameter vector, the *main effect*  $g(\theta_k)$  is the expected computer model output,  $E(z(\theta)|z_{known}, \theta_k)$ , averaged over the joint prior distribution  $\pi(\theta_{-k})$  for all components of  $\theta$  except the  $k$ th. Two-way interaction effects  $g(\theta_k, \theta_l)$  are defined similarly, except integration is with respect to a joint prior on all parameter vector components but the  $k$ th and  $l$ th,  $k \neq l$ . A main effects graph plots  $g(\theta_k)$  against  $\theta_k$ , thereby illustrating the effect of varying the single parameter  $\theta_k$  on the simulation output. Two-way interaction effects are visualized through 3D surfaces based on the two effects and expected computer model output. For functional computer model output, like pathogen load or macrophage counts over time, a main effects plot displays expected *functional* computer model output for a fixed  $\theta_k$ . Varying  $\theta_k$  will produce multiple plots that can be compared to reveal the main effect of  $\theta_k$ . For details see Supplement A.4.

The GP is a functional approximation of the computer model and therefore lends itself to functional analysis of variance (FANOVA) decomposition when marginal priors on the individual components of  $\theta$  are independent (Schonlau and Welch, 2006). In FANOVA decomposition, the total functional variance of the GP is decomposed into variance due to the main and interaction effects of the parameters. The percentage of the total functional variance

accounted for by a particular effect provides a measure of the importance of that effect. We report the percent of total functional variance that is accounted for by main effects and two-way interactions. Details can be found in Schonlau and Welch (2006).

### 3.3. Computer model calibration using field observations

Computer model calibration is the process of estimating the unknown computer model parameter vector  $\theta$ , based on a vector of field observations  $y_F = (y_{F,1}, \dots, y_{F,n})$ , so that output from the calibrated computer model is (hopefully) consistent with field data (see the schematic provided in Fig. 1). Typically, the relationship between a computer model and reality is defined according to

$$y_{F,i} = y_M(\theta) + b + \varepsilon_i, \quad i = 1, \dots, n,$$

where  $y_{F,i}$  denotes a scalar field observation,  $y_M(\theta)$  the corresponding computer model output evaluated at the true value  $\theta$  of computer model parameters,  $b$  is a bias term, and  $\varepsilon_i$  are observation errors (Higdon et al., 2004). For simplicity, we assume that the computer model is an accurate, unbiased representation of reality, so  $b = 0$ . As is typical, we assume measurement errors are independent and identically distributed (iid) as  $N(0, \sigma_F^2)$ , so that all field observations are normally distributed. Finally, because  $y_M(\theta)$  is not readily available, we replace it with the GP approximation  $z(\theta)$ . Then the field data conditional on the computer output  $z_{known}$  follow a normal distribution with density  $f(y_F|\theta, z_{known})$ . Details are in Supplement A.5.

With the likelihood of the data  $y_F$  defined, estimation of  $\theta$  can proceed. Here we perform estimation of  $\theta$  in a Bayesian context. To reflect knowledge or uncertainty before data collection, a prior distribution  $\pi(\theta)$  is specified, in our case the same product of independent uniforms used for SA. After observing the field observations  $y_F$ , calibration yields updated knowledge about  $\theta$  in the form of a posterior distribution

$$\pi(\theta|y_F, z_{known}) \propto f(y_F|\theta, z_{known})\pi(\theta).$$

The posterior distribution is estimated using Markov chain Monte Carlo (see next section). It incorporates various sources of uncertainty, including measurement error, uncertainty in the GP approximation to the computer model, and prior uncertainty about  $\theta$ . For dealing with multiple scalar computer outputs and other details, see Supplement A.5.

### 3.4. Markov chain Monte Carlo

Markov chain Monte Carlo (MCMC) is a sampling method used to sample from a probability distribution that is known up to a multiplicative constant (Gilks et al., 1998). In computer model calibration, the goal is to obtain the posterior distribution  $\pi(\theta|y_F, z_{known})$ , for the computer model parameter vector  $\theta$ , given a prior  $\pi(\theta)$  on computer model parameters. Using the MCMC sample, we estimate the marginal posterior distributions of each  $\theta_k$ ,  $k = 1, \dots, p$ , by constructing a histogram. We take the posterior median as a point estimate and use the 90% posterior credible interval to quantify uncertainty in the estimate. The 90% posterior credible interval of the parameter  $\theta_k$  is the interval  $(a, b)$ , where  $a$  is the 5% quantile of  $\theta_k$  and  $b$  is the 95% quantile of  $\theta_k$  from the posterior distribution; 90% prior credible intervals are defined similarly, but using quantiles of the prior distribution. An example of a 90% posterior credible interval is given in Fig. 1D. For details about sampling from the posterior distribution, see Supplement A.6.

### 3.5. Bayesian model comparison

Bayes factors (BFs) are commonly used for assessing the evidence in favor of one model against an alternative (Kass and Raftery, 1995). In our case, we are interested in assessing the evidence in favor of model  $M_1$ , which allows necrotic tissue parameter  $A_{necr}$  to be 1, 2, 3, 4, 5, or 6, against the null model  $M_0$ , where  $A_{necr}$  is identically 0. The full parameter vector for both models includes additional parameters, but  $A_{necr}$  is the only parameter altered in the comparison.

For observed data  $y_F$ , the Bayes factor  $BF_{10}$  in favor of model  $M_1$  against  $M_0$  is the ratio of the marginal likelihood of the data under  $M_1$  to the marginal likelihood of the data under  $M_0$ . Equivalently, the Bayes factor is the ratio of the posterior odds in favor of  $M_1$  to its prior odds,

$$BF_{10} = \frac{\pi(M_1|y_F, z_{known})}{\pi(M_0|y_F, z_{known})} = \frac{\pi(M_1)}{\pi(M_0)},$$

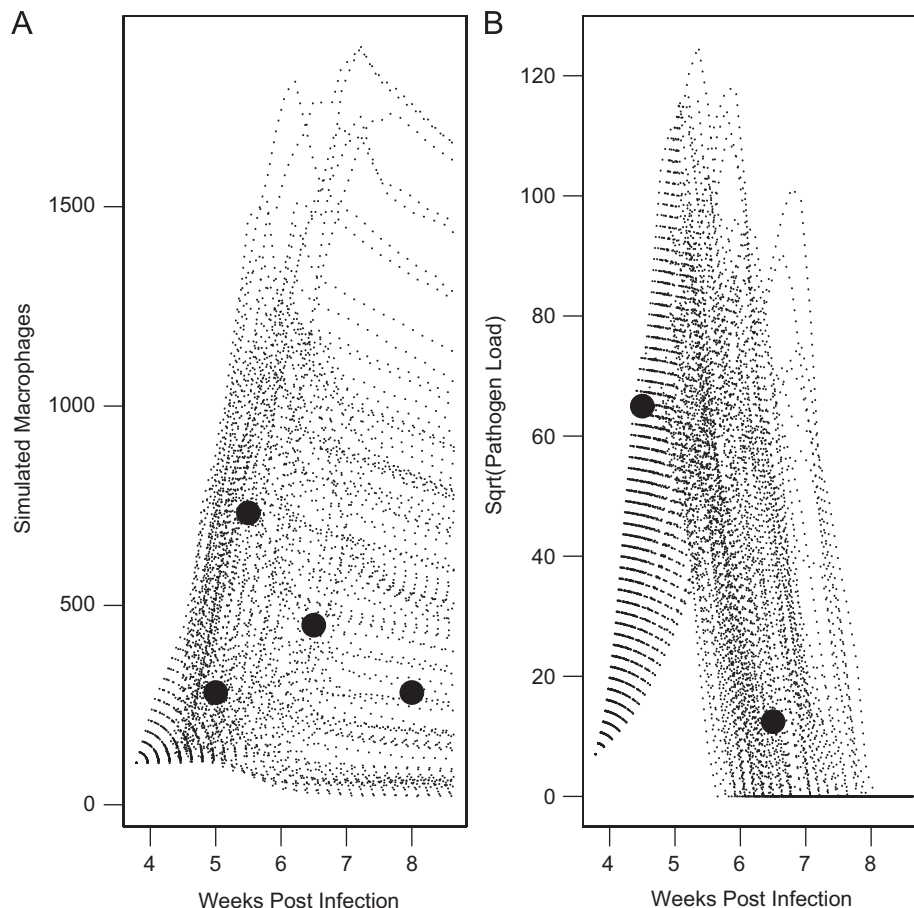
where  $\pi(M_i)$  is the *a priori* probability assigned to  $M_i$ ,  $i = 0, 1$ , with  $p(M_0) = 1 - p(M_1)$ . If the two models are *a priori* equally likely, which is what we assume, then the Bayes factor is the posterior odds in favor of  $M_1$ . In our analysis, we estimate the probabilities  $\pi(M_i|y_F, z_{known})$ ,  $i = 1, 2$  from the posterior distribution of the computer model parameter vector  $\theta$ , which includes  $A_{necr}$ , that we obtain using MCMC.

## 4. Results

### 4.1. Experimental design and computer model output

Ideally we would like to calibrate all 25 parameters in our computer model, but because field data is limited we will focus our attention on the five parameters we believe to be the most interesting and the most relevant given the field data available. Thus, we focus on parameter vector  $\theta = (\alpha_I, k_I, \beta_{Mrecr}, p_{Trecr}, T_{threshold})$  whose true value we assume exists somewhere in the hypercube formed by allowing each parameter to vary in the ranges specified in Table C1 of the Supplement. The remaining parameters are fixed at values given in the table. Note that the parameter  $A_{necr}$  is currently fixed at 2; we vary this parameter in Section 4.5. The parameters  $\alpha_I$ ,  $k_I$ , and  $T_{threshold}$  are pathogen-dependent parameters which determine growth rate, the lethal parasite density in infected macrophages, and the threshold level of pathogen that triggers the adaptive immune response.  $\beta_{Mrecr}$  is a macrophage recruitment parameter and  $p_{Trecr}$  determines the rate of T cell recruitment. We choose all parameter values and ranges to be consistent with *L. major* infection, and display references for these choices in Table C1.

In order to construct the GP approximation to the computer model, we must obtain computer model output for  $m$  representative choices of the parameter vector  $\theta$ . In the absence of informative prior information about the true value of  $\theta$ , the desired approach is to choose a space-filling design that adequately covers the parameter space. We choose 50 design points using a maximin Latin Hypercube



**Fig. 3.** Simulated output over time (black dashed lines) and the field data (solid black circles) used in calibration. (A) Simulated macrophages; field data for time points  $t_1$  through  $t_4$ . (B) Square root of simulated pathogen load; field data for time points  $t_5$  and  $t_6$ . Simulated data was obtained from a 50 point Latin Hypercube design with two replicates per design point over the parameters  $\alpha_I$ ,  $k_I$ ,  $\beta_{Mrecr}$ ,  $p_{Trecr}$ ,  $T_{threshold}$ . Outputs from all 100 simulations are plotted.

design (Morris and Mitchell, 1995) and run two replicates at each design point. Output from the simulation consists of total macrophage and parasite counts on the lattice over time (Fig. 3, black dashed lines). Variation in macrophage counts and pathogen load is evident in the sizes and locations of the peaks. Peak macrophage levels range from 112 to 1903 macrophages, with peak locations ranging from 4.2 to 6.9 wpi. Peak pathogen load ranges from 3263 to 15610 (57.1 and 125 on the scale of Fig. 3B), with peak locations ranging from 4.5 to 7.0 wpi. In all of our simulations, infection is cleared by approximately 8 wpi.

#### 4.2. Fitting the Gaussian processes

Our goal is to fit a GP approximation to the computer model output given in Fig. 3. The calibration will consider the following field data: macrophage counts at  $t_1 = 5$ ,  $t_2 = 5.5$ ,  $t_3 = 6.5$ , and  $t_4 = 8$  wpi; and parasite load at  $t_5 = 4.5$  and  $t_6 = 6.5$  wpi (Fig. 3, solid black circles). Therefore, our GP emulators must be able to predict computer model output at these six time points. Let  $z_j^*(\theta^{(i)})$  be the *untransformed* computer model output for simulation  $i$  at time point  $t_j$ , and define

$$z_{known,j}(\theta^{(i)}) = \begin{cases} \log z_j^*(\theta^{(i)}), & j = 1, \dots, 5, \\ \log(z_j^*(\theta^{(i)}) + 1), & j = 6. \end{cases} \quad (2)$$

The simulation output is adjusted at  $t_6$  to prevent numerical issues for simulations where the infection is cleared by 6.5 wpi. We fit independent GPs to each vector  $z_{known,j}$  as described in Section 3.1 for the six time points using the software PERK (Santner et al., 2003). Using independent GPs is a significant assumption, especially for time series data on single subjects, but the calibration data we use are independent because distinct mice were sacrificed for each observation (Belkaid et al., 2000).

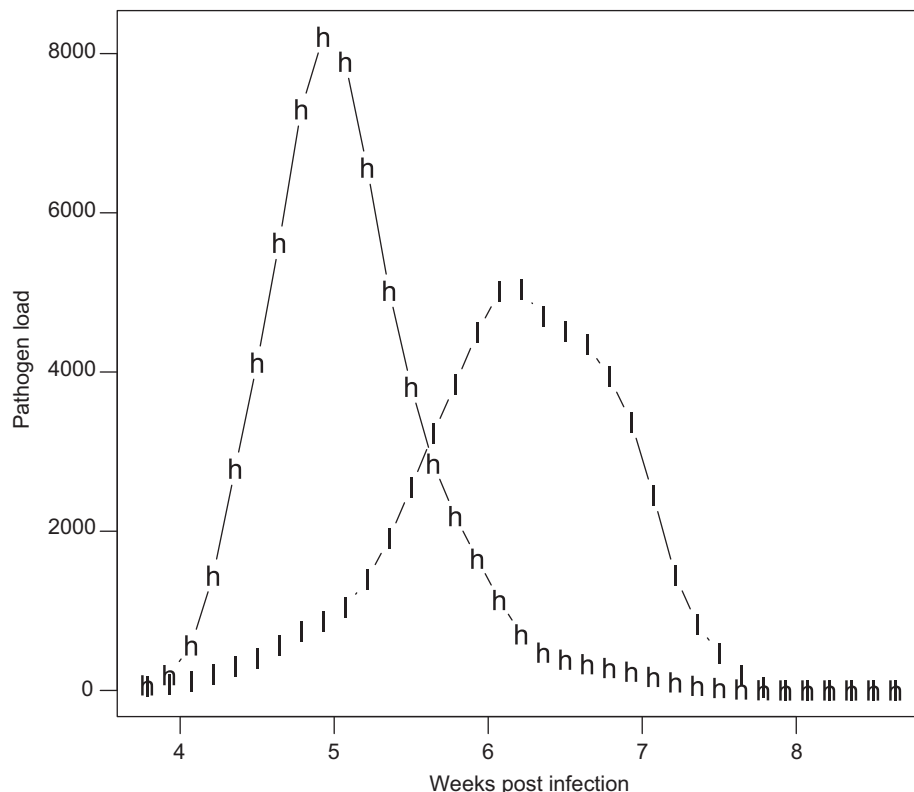
Accuracy, as determined by cross-validated predictions and standardized residuals of the GP predictors, is good (see Supplement B.1); subsequent calibration uses these fitted GPs.

#### 4.3. Sensitivity analysis

In order to quantify the importance of computer model parameters on simulation output and to characterize their effects, we perform sensitivity analysis as described in Section 3.2. While we performed many different SAs, we report here three results that both illustrate the GP-based techniques and important dynamics of the model. Specifically, we study the main effects of model parameters on the time course of pathogen load, use FANOVA to find important parameters and interactions at all time points where data are available, and study these effects through main and interaction effects plots.

To characterize the qualitative effect of each parameter on pathogen load over time, we fit GP emulators to the six most important PC weights of the pathogen load time course. These six components explain over 99% of the variation in output. Fig. 4 illustrates the main effect of growth rate, ( $\alpha_I$ ), the most important computer model parameter on pathogen load over time, where we assess parameter importance by calculating the variability of each main effect as described in Supplement A.4. Interestingly, high growth rate favors pathogen load early in infection but not at later time points. When growth rate is high, there is a sharp increase in pathogen load and a higher peak, but a faster resolution of infection. Infection is prolonged when growth rate is low, and the peak is lower, but the infection is still eventually cleared.

Because we are interested in calibrating the computer model to time points  $t_1, \dots, t_6$ , we perform sensitivity analysis on the GP predictors for those time points. The percentage contributions of main and important two-factor interaction effects are provided in



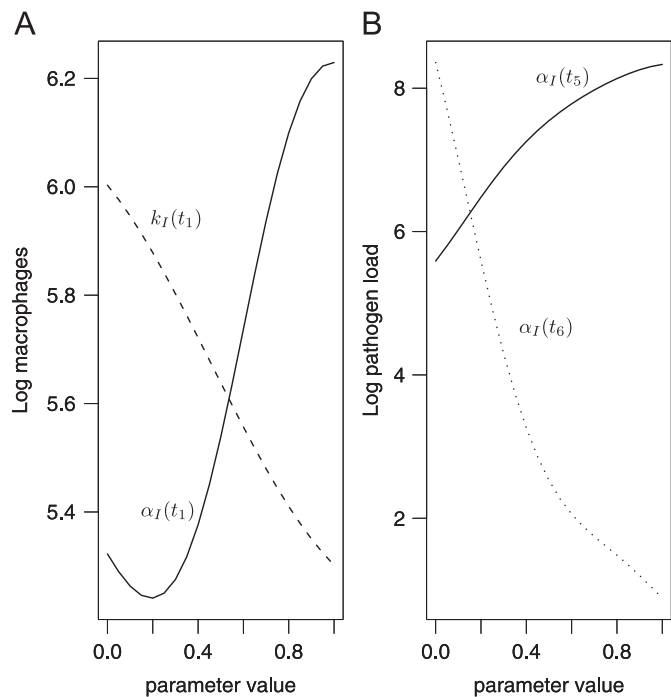
**Fig. 4.** Main effects of growth rate ( $\alpha_I$ ) on simulated pathogen load over time. Main effects are calculated using low (l) and high (h) values of  $\alpha_I$ , with (l,h) equal to  $(2.41, 9.63) \times 10^{-4}/\text{min}$ .



**Table 2**  
FANOVA decomposition.

Effect	Time point					
	Macrophage (wpi)			Pathogen (wpi)		
	$t_1$ (5)	$t_2$ (5.5)	$t_3$ (6.5)	$t_4$ (8)	$t_5$ (4.5)	$t_6$ (6.5)
$\alpha_I$	54.13	24.74	1.32	3.08	93.47	54.61
$k_I$	18.68	29.73	36.65	33.53	5.09	12.44
$\beta_{Mrecr}$	9.76	6.48	6.09	8.44	–	1.68
$p_{Trecr}$	–	–	–	–	–	2.46
$T_{threshold}$	4.15	17.48	39.25	32.08	–	23.32
$\alpha_I \times k_I$	11.27	11.30	5.42	5.04	1.31	–
$\alpha_I \times T_{threshold}$	1.43	5.48	–	–	–	4.47
$k_I \times T_{threshold}$	–	2.50	–	15.12	–	–
Total	99.42	97.72	97.30	97.40	99.88	98.98

For each time point, the percentage contribution of main effects and two-way factor interaction effects to the total functional variance of the Gaussian process predictor. Only effects accounting for > 1% of the total variance for at least one time point are shown, with ‘–’ indicating < 1%.



**Fig. 5.** Main effects of selected parameters. (A) Log macrophages at 5.0 ( $t_1$ ) as a function of  $\alpha_I$  and  $k_I$ . (B) Similarly, log pathogen load at 4.5 ( $t_5$ ) and 6.5 ( $t_6$ ) wpi as functions of  $\alpha_I$ . The biological ranges of model parameters, corresponding to (0,1), are as follows:  $\alpha_I$ ,  $(2.41, 9.63) \times 10^{-4}$ /min;  $k_I$ , (50, 400) parasites.

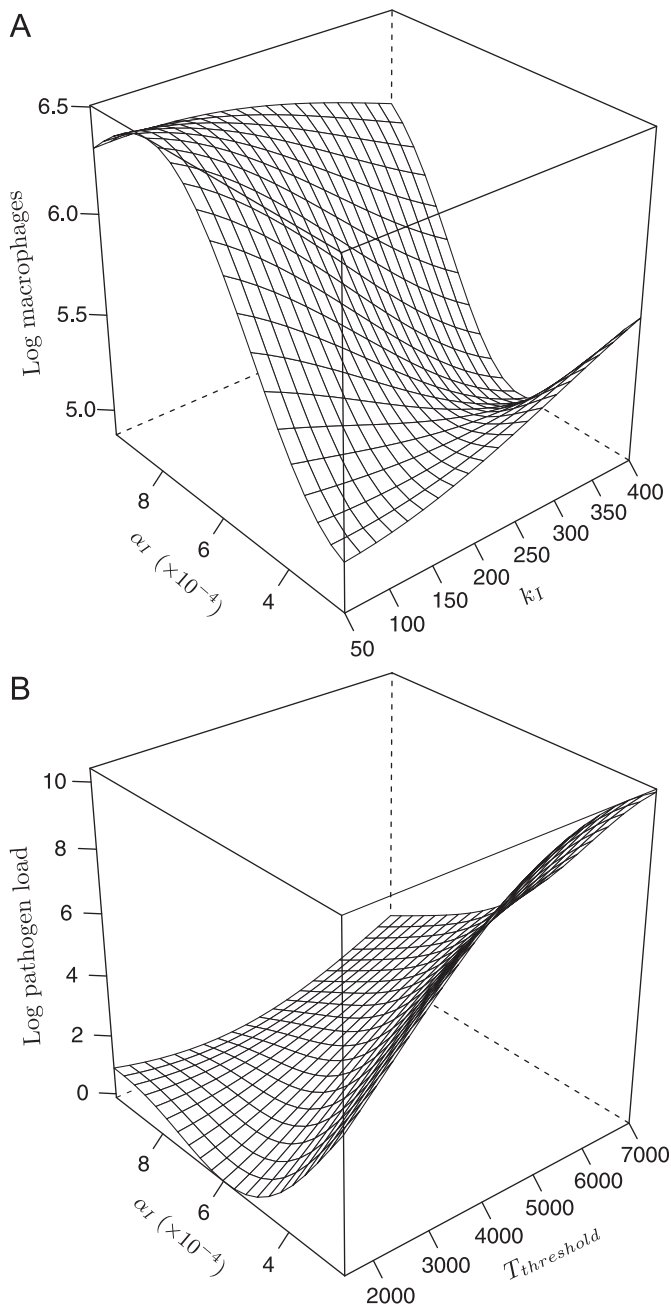
**Table 2.** Together, main effects and two-factor interactions account for more than 97% of the total functional variance of the GP emulators for all time points. The most important main effects at  $t_1$ , for parameters  $\alpha_I$  and  $k_I$ , are plotted in Fig. 5A. The most important main effect at  $t_5$  and  $t_6$ , which are both for  $\alpha_I$ , are plotted in Fig. 5B. The two largest interaction effects,  $\alpha_I \times k_I$  and  $\alpha_I \times T_{threshold}$ , are plotted in Fig. 6.

The most influential parameter for any time point is pathogen growth rate ( $\alpha_I$ ), which accounts for 93.5% of the total functional variance for the GP predictor at  $t_5$ . Growth rate is also the most important parameter at  $t_1$  and  $t_6$  (Table 2). For (scaled) values of  $\alpha_I > 0.2$ , the relationship between pathogen growth rate and macrophage levels at 5.5 wpi ( $t_1$ ) is positive, whereas lethal parasite density ( $k_I$ ), the second most important parameter at  $t_1$ , is negatively related to the response (Fig. 5A). However, these effects are not independent; the  $\alpha_I \times k_I$  interaction accounts for 11.3% of the total variance of the  $t_1$  GP predictor (Table 2), with  $k_I$  showing

a slight positive effect when  $\alpha_I$  is below  $4 \times 10^{-4}$ /min, but a negative effect otherwise (Fig. 6A). In agreement with the  $\alpha_I$  main effect plot (Fig. 4), the effect of  $\alpha_I$  on pathogen load changes over time (Fig. 5B). The parameter  $\alpha_I$  has a positive effect on pathogen load at 4.5 wpi but a negative effect on pathogen load at 6.5 wpi ( $t_5$  and  $t_6$ , respectively). At  $t_6$ , there is an interaction between  $\alpha_I$  and  $T_{threshold}$ , which accounts for 4.5% of the total functional variance of the GP predictor (Table 2).  $T_{threshold}$  is positively and strongly correlated with pathogen load, but only when  $\alpha_I$  is low (Fig. 6B).

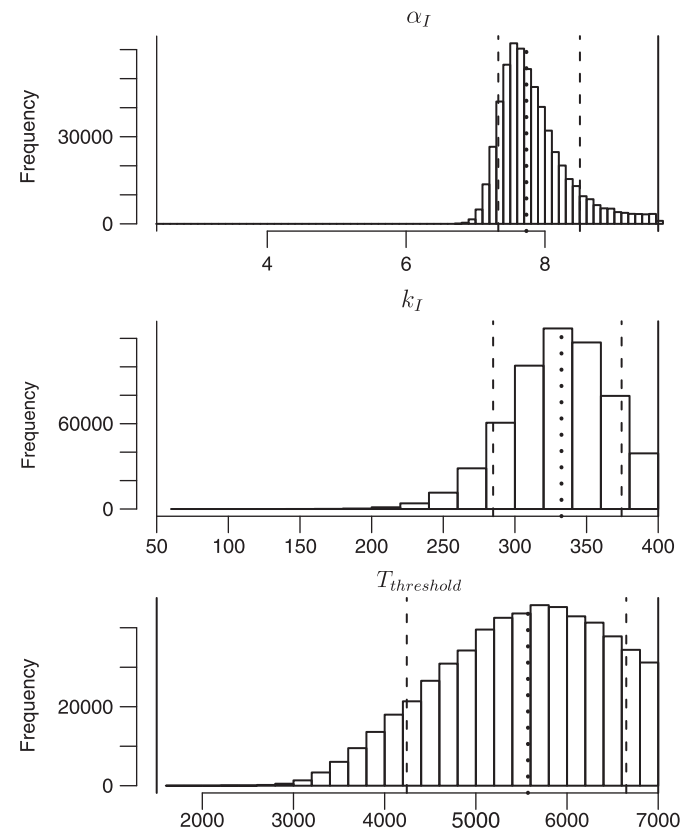
4.4. Calibration using field data

Belkaid et al. (2000) provide data plots for parasite and macrophage counts over time following low dose *L. major* infection in the mouse ear, but the amount of field data available



**Fig. 6.** Two-way interaction plots. (A) Log macrophages at 5 wpi ( $t_1$ ) for parameters  $\alpha_I$  and  $k_I$ ; (B) Log pathogen load at 6.5 wpi ( $t_6$ ) for parameters  $\alpha_I$  and  $T_{threshold}$ . Units are rate/min for  $\alpha_I$  and parasites for  $k_I$  and  $T_{threshold}$ .

for calibration is limited (e.g., macrophage data are *representative* of 10 experiments; Belkaid et al., 2000), and we do not have accurate information about field error. Therefore, in order to verify that accurate calibration is possible using the type of data available, we first apply the calibration approach to simulated field data, generated assuming a randomly chosen parameter vector  $\theta^{(sim)}$ . In the absence of field-estimated error, we assume that  $\sigma_F^2$  is equal to the sample variance of observed computer model output, which we calculate by using the ABM to generate three independent observations for each time point and fixing



**Fig. 7.** Posterior distributions of  $\alpha_I$ ,  $k_I$ , and  $T_{threshold}$  following calibration using field data. Frequency histograms representing posterior distributions of  $\alpha_I$ ,  $k_I$ , and  $T_{threshold}$  are obtained from three chains of 200,000 samples each, including a 10,000 sample burn-in. Dashed black lines correspond to 90% posterior credible intervals; the median value of each parameter is represented by a thick black dotted line. The solid black lines that surround the histograms indicate the *a priori* parameter ranges, according to independent prior distributions:  $\alpha_I \sim \text{Unif}(2.407, 9.632) \times 10^{-4}/\text{min}$ ,  $k_I \sim \text{Unif}(50, 400)$  parasites, and  $T_{threshold} \sim \text{Unif}(1500, 7000)$  parasites.

**Table 3**

Posterior summary of computer model parameters (with  $A_{neer} = 2$ ) following calibration using field data.

Parameter	Posterior estimates		Units	% decrease in credible interval
	Median	90% credible interval		
$\alpha_I$	7.73	(7.24, 8.88)	/min. ( $\times 10^{-4}$ )	74.73
$k_I$	1.50	(1.30, 1.60)	doubling time ( $\times 10$ h)	
	3.32	(2.70, 3.84)	parasites ( $\times 10^2$ )	63.79
$\beta_{Mrecr}$	2.31	(2.03, 2.81)	/unit chemokine ( $\times 10^{-4}$ )	13.58
$p_{Trecr}$	1.93	(1.54, 2.43)	%/time step ( $\times 10^{-2}$ )	0.79
$T_{threshold}$	5.57	(3.92, 6.82)	parasites ( $\times 10^{-4}$ )	41.46

All parameters are scaled back to their biological values and are reported in the units indicated;  $\alpha_I$  is reported on two scales, as a rate of growth and as a doubling time. The last column reports the percent decrease in the width of the 90% posterior credible interval relative to the 90% prior credible interval.

$\theta = \theta^{(\text{sim})}$ . The simulated field data yielded substantial and accurate information about the five calibration parameters, particularly  $\alpha_I$  and  $k_I$ . See Supplement B.2 for details and discussion.

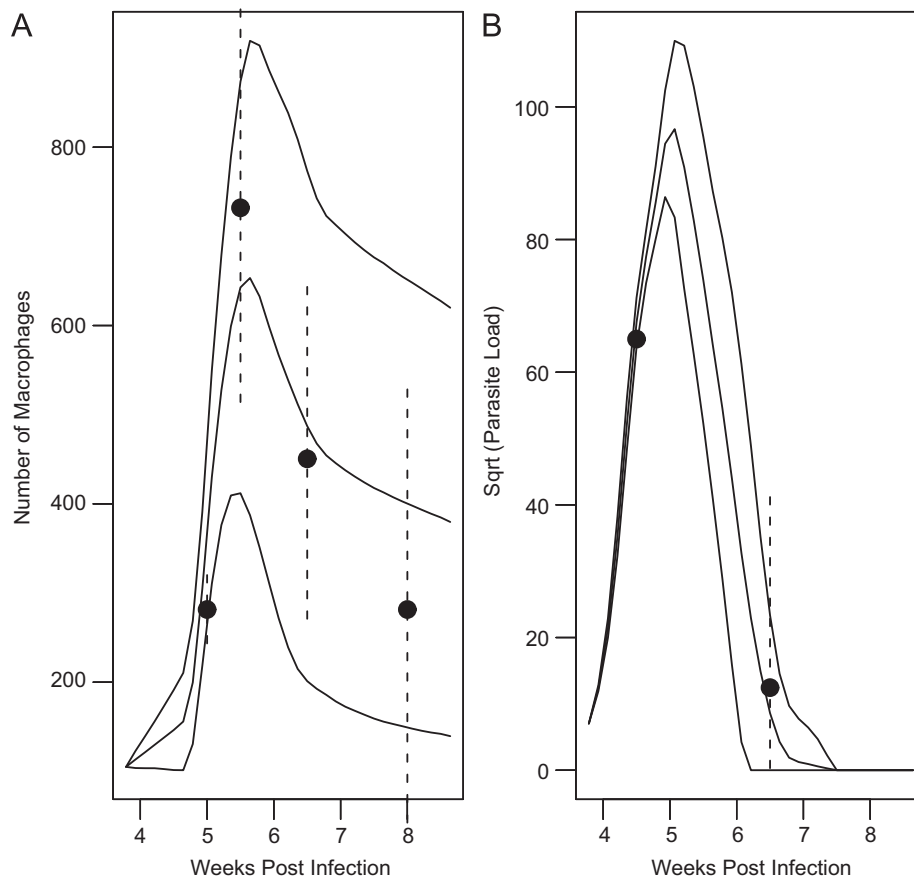
We next calibrate the model using the mouse data (Belkaid et al., 2000). Because we are modeling the center of the infection, and not the entire ear, we scale field observations by a factor of  $71^{-1}$ , based on expected macrophage counts at two days post high dose infection (i.e., the start of our simulation), and log transform these observations following Eq. (2). We continue to assume the field variance  $\sigma_F^2$  estimated during testing with simulated field data. A summary of the resulting marginal posterior distributions is given in Table 3, and the distributions for  $\alpha_I$ ,  $k_I$ , and  $T_{\text{threshold}}$  are plotted in Fig. 7. Using posterior medians as point estimates, we estimate a pathogen growth rate ( $\alpha_I$ ) of  $7.73 \times 10^{-4}/\text{min}$  (a doubling time of 15 h), and a lethal parasite density ( $k_I$ ) of 332 parasites. The posterior distributions of  $\beta_{\text{Mrecl}}$  and  $p_{\text{Trecl}}$  are relatively non-informative (Table 3 and data not shown). To determine whether our calibrated model could accurately reproduce field observations, we ran 100 simulations using parameter vectors drawn from the posterior distribution of  $\theta$ . Graphs of mean computer model output,  $\pm 1$  standard deviation, along with the observed field data, are given in Fig. 8. All field observations are within one standard deviation of mean computer model output from the calibrated computer model.

#### 4.5. Comparison of alternative models of macrophage behavior

In the previous model, we arbitrarily fixed the parameter  $A_{\text{necr}}$  at a value of two in an attempt to capture the decrease in

macrophage counts between 5.5 and 8 wpi that Belkaid et al. (2000) observe. We next want to estimate the value of  $A_{\text{necr}}$  and to assess whether or not a model with necrotic tissue fits the data better than a model without necrotic tissue. Formally, we consider the alternative model  $M_1$ , where  $A_{\text{necr}}$  is 1, 2, 3, 4, 5, or 6, against the null model  $M_0$ , where  $A_{\text{necr}}$  is identically 0. Because the posterior distributions for  $\alpha_I$ ,  $k_I$ , and  $T_{\text{threshold}}$  had posterior distributions concentrated at the high end of their ranges (Fig. 7), we update our prior ranges for the parameter vector  $\theta = (\alpha_I, k_I, \beta_{\text{Mrecl}}, p_{\text{Trecl}}, T_{\text{threshold}})$  by shifting the range of each parameter such that its 99.99% highest posterior density credible interval is centered in the new range. This shift did not effect the parameters  $\beta_{\text{Mrecl}}$  and  $p_{\text{Trecl}}$ . The width of the ranges were kept the same in order to maintain the same level of uncertainty in our priors. Updated posterior ranges for  $\alpha_I$ ,  $k_I$ , and  $T_{\text{threshold}}$  and the range for  $A_{\text{necr}}$ , are given in Table C2 of the Supplement.

We use a 70 point maximin Latin Hypercube design to vary the updated computer model parameter vector  $\theta = (A_{\text{necr}}, \alpha_I, k_I, \beta_{\text{Mrecl}}, p_{\text{Trecl}}, T_{\text{threshold}})$ , again with two replicate runs per design point. Because  $A_{\text{necr}}$  is discrete, we map its scaled values from the Latin Hypercube design  $\{1/70, \dots, 1\}$  to the discrete values  $\{0, \dots, 6\}$  so that each possible value of  $A_{\text{necr}}$  is observed 10 times in the design. For time points  $t_j$ ,  $j = 1, \dots, 6$ , we fit independent GPs to  $Z_{\text{known},j}$  and calibrate our computer model as before. A sensitivity analysis shows the sharpest decrease in macrophage levels after 5.5 wpi occurs when  $A_{\text{necr}}$  is high, while the other computer model parameters mostly affect the peak height, rather than steepness of the decline (Fig. B3 of the Supplement and data not shown).

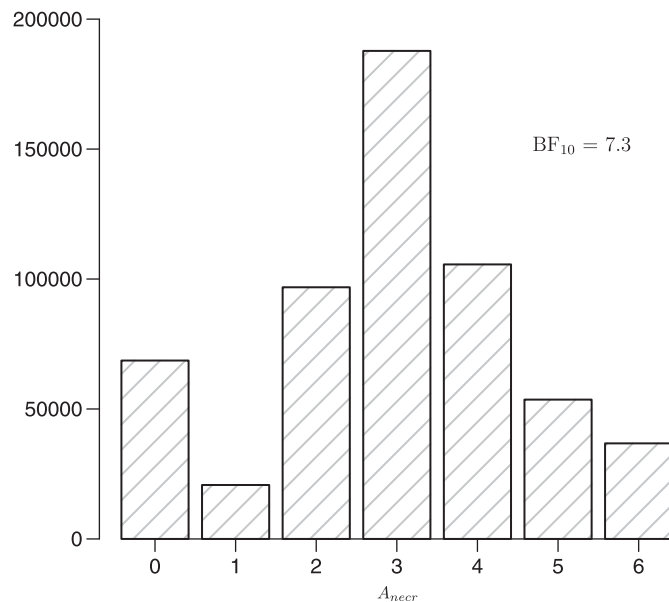


**Fig. 8.** Posterior predictive distribution of macrophage counts (A) and parasite load (B). One hundred simulations were run using randomly selected parameter vectors from the posterior distribution of  $\theta$ . Solid black lines, mean computer model output  $\pm 1$  standard deviation; solid black circles, field data ((A), time points  $t_1$  through  $t_4$ ; (B), time points  $t_5$  and  $t_6$ ) used during calibration; vertical dashed black lines indicate  $\pm 1$  standard deviation corresponding to estimated field variance.

**Table 4**Posterior summary of computer model parameters, including  $A_{necr}$ , following calibration using field data.

Parameter	Posterior estimates		Units	% decrease in credible interval
	Median	90% credible interval		
$\alpha_l$	7.38	(6.97, 8.07)	/min ( $\times 10^{-4}$ )	83.05
	1.57	(1.43, 1.66)	doubling time (10 h)	
$k_l$	3.17	(2.26, 3.89)	parasites ( $\times 10^2$ )	48.38
$\beta_{Mrecr}$	2.58	(2.06, 2.94)	/unit chemokine ( $\times 10^{-4}$ )	2.38
$p_{Trecr}$	1.86	(1.53, 2.41)	%/time step ( $\times 10^{-2}$ )	2.81
$T_{threshold}$	5.27	(2.53, 6.80)	parasites ( $\times 10^{-4}$ )	13.61
$A_{necr}$	3.00	(0.00, 6.00)	units of necrotic tissue	0.00

All parameters are scaled back to their biological values and are reported in the units indicated;  $\alpha_l$  is reported on two scales, as a rate of growth and as a doubling time. The last column reports the percent decrease in the width of the 90% posterior credible interval relative to the 90% prior credible interval.



**Fig. 9.** Posterior distribution of  $A_{necr}$  following calibration using field data. A frequency histogram representing the posterior distribution of  $A_{necr}$  obtained from three chains of 200,000 samples each, including a 10,000 sample burn-in. Prior distribution of  $A_{necr}$  is  $\pi(A_{necr}) = \frac{1}{2}$  if  $A_{necr} = 0$  and  $\pi(A_{necr}) = \frac{1}{12}$  if  $A_{necr} \in \{1, \dots, 6\}$ .  $BF_{10}$ , Bayes factor of  $M_1 : A_{necr} \in \{1, \dots, 6\}$  against  $M_0 : A_{necr} = 0$ .

To calibrate  $A_{necr}$  and other parameters to the available data, we assume that the two models are *a priori* equally likely (i.e.,  $\pi(M_0) = \pi(M_1) = \frac{1}{2}$ ), specifically

$$\pi(A_{necr}) = \begin{cases} \frac{1}{2} & \text{if } A_{necr} = 0, \\ \frac{1}{12} & \text{if } A_{necr} = 1, \dots, 6 \end{cases} \quad (3)$$

and  $A_{necr}$  is independent of the other parameters. The remaining (scaled) parameters are again given independent Unif(0,1) priors. A summary of each marginal posterior distribution is given in Table 4, and the posterior distribution of  $A_{necr}$  is given in Fig. 9. Using posterior medians as point estimates, we estimate a pathogen growth rate ( $\alpha_l$ ) of  $7.38 \times 10^{-4}$ /min (a doubling time of 15.7 h), a transfer threshold ( $k_l$ ) of 317 parasites, and the production of three units of necrotic tissue ( $A_{necr}$ ) following macrophage activation. The Bayes factor for  $M_1$  against  $M_0$  is  $BF_{10} = 7.3$  which indicates ‘substantial’ evidence in favor of the

model with necrotic tissue-induced macrophage loss, using the scale of Kass and Raftery (1995).

## 5. Discussion

In this paper, we have described an agent-based model of *L. major* infection, fit Gaussian process approximations to computer model outputs of pathogen load and macrophage counts at various time points, performed sensitivity analysis on the GP emulators to identify and characterize the effects of important model parameters, and calibrated the computer model using the data of Belkaid et al. (2000). GP diagnostics confirm the accuracy of the GP emulators (details in Supplement B.1), and tests with simulated field data indicate accurate model parameter estimates are possible despite the limited field data (details in Supplement B.2). Our most interesting findings are: (1) pathogen growth rate naturally enhances pathogen load early in infection but counterintuitively suppresses pathogen load later, (2) the macrophage recruitment parameter is *not* the greatest determinant of macrophage levels when these cells first infiltrate the infection site (5 wpi), (3) *L. major* can replicate extensively before destroying the host cell, indicative of an immune hiding phenotype, and (4) accelerated macrophage loss in the presence of necrotic tissue provides a substantially better fit to the data. Sensitivity analysis, calibration with field data, and model comparison are discussed in detail in the following sections.

### 5.1. Sensitivity analysis

The FANOVA decomposition (Table 2) summarizes how variability in computer model output is attributed to uncertainty in computer model input, specified through prior distributions. Using the language of Kennedy and O’Hagan (2001), the prior distribution may represent (1) current belief about the true value of a parameter (*parameter uncertainty*), or (2) environmental variability in a parameter that cannot be controlled (*parametric variability*). Our choices of priors are primarily based on uncertainty, although we also expect variability to be present. We also expect calibration to provide the most information about parameters with large FANOVA contributions. Indeed, the credible interval for  $\alpha_l$ , which was the greatest across-the-board contributor to variation (Table 2), decreased the most after calibration (Table 3). In Supplement B.3 we discuss the idea of using sensitivity analysis to guide data collection so that model calibration is more efficient. Here, we discuss how the estimated main and interaction effects (Figs. 4–6) help illustrate the complex relationship between pathogen behavior and host response.



*Pathogen persistence is associated with slow pathogen growth and immune avoidance.* A fast-growing pathogen peaks early and quickly disappears, while a slow-growing pathogen peaks late and typically lower (Fig. 4). These patterns are also apparent in the main effects of growth rate ( $\alpha_l$ ), which has a positive slope for pathogen load at 4.5 wpi (time point  $t_5$ ) but a negative slope at 6.5 wpi (time point  $t_6$ ; Fig. 5B). The paradox that high growth rate favors early survival but is long-term detrimental to the pathogen results because a rapidly replicating pathogen risks early detection by the adaptive immune response. In the model, the adaptive response is triggered once pathogen load crosses a threshold ( $T_{\text{threshold}}$ ). Looking at the last pathogen time point  $t_6$  (Fig. 6B), expected pathogen load decreases as growth rate increases, but the loss of pathogen can be partially abrogated by increasing  $T_{\text{threshold}}$ . Together, pathogens that grow slowly and without early detection by the adaptive immune response persist at the highest levels late in infection.

There have been biological and model-based observations about pathogen immune evasion and growth rate related to our findings. Parasites use many mechanisms to escape immune detection (reviewed in Damian, 1997). For example, *L. major* is known to inhibit antigen presentation in infected macrophages (Fruth et al., 1993), and *Toxoplasma gondii* downregulates MHC class II expression, which is required for antigen presentation (Luder et al., 1998). Both are slow-growing pathogens that produce latent or persistent infection. Others have noted that high growth rates are not always favorable to pathogens. In their model of *Mycobacterium tuberculosis*, Segovia-Juarez et al. (2004) report that the sign of the partial rank correlation coefficient between growth rate and extracellular bacteria changes from positive to negative over time. From an epidemiological perspective, Ewald (1994) notes that virulent pathogens cannot survive if they kill or immobilize their hosts before transmission to a new host occurs, so high growth rate is detrimental without efficient transmission, for example via vectors.

Many pathogen and host parameters influence the amount of macrophages at 5.5 and 6.5 wpi (time points  $t_1$  and  $t_3$ , respectively). One surprising result is that the macrophage recruitment parameter  $\beta_{\text{Mrecr}}$  is not the most influential at either time point (Table 2), despite being the only parameter that directly influences macrophage levels. The large uncertainty ranges for some input parameters, such as  $\alpha_l$ , explain some of the corresponding variation in macrophage counts, but there are also unanticipated system dynamics at work. At macrophage time point  $t_1$ , growth rate ( $\alpha_l$ ) is the most important parameter. As the doubling time increases above 30 h ( $\alpha_l > 0.2$  in Fig. 5A), the number of macrophages is expected to increase sharply, probably because rapid growth promotes spread of the infection among macrophages, which each release macrophage-recruiting chemokine. The upward curve in the main effect of  $\alpha_l$  near its lower range is likely an artifact of the stochasticity of the ABM and our use of only two replicates per design point when fitting the GP.  $T_{\text{threshold}}$ , the pathogen level that triggers the T cell response, accounts for the most functional variance of the GP predictor at time point  $t_3$ . Although the primary role of T cells is macrophage activation, activated macrophages are short-lived and leave behind necrotic tissue that triggers the departure of additional resting macrophages. Also, as T cells successfully resolve the infection, cytokine levels decrease, reducing macrophage recruitment. In combination, pathogen and T cell actions, and therefore  $\alpha_l$  and  $T_{\text{threshold}}$ , have a significant effect on macrophage numbers.

Like  $\alpha_l$ , lethal parasite density  $k_l$  in macrophages should affect infection spread and consequently, macrophage recruitment. As the value of  $k_l$  decreases, increased pathogen transmission leads to increased macrophage recruitment and higher macrophage counts (Fig. 5A). However the role of parameters  $\alpha_l$  and  $k_l$  on

macrophage recruitment is complex, and their two-way interaction accounts for 11.3% of the total functional variance of the  $t_1$  GP predictor. Interestingly,  $k_l$  has a positive effect on macrophage counts when  $\alpha_l$  is low but a negative effect when  $\alpha_l$  is high (Fig. 6A). We suspect that low  $\alpha_l$  prevents much spread of infection by  $t_1$  regardless of  $k_l$ , so large  $k_l$  simply means more infected macrophages have survived until  $t_1$ . However, when  $\alpha_l$  is high, rapid spread of infection, enhanced as  $k_l$  declines, induces massive macrophage recruitment.

## 5.2. Calibration using field data

In Supplement B.2, we show that calibration with simulated data, designed to mimic the available field data, yields accurate and informative estimates of computer model parameters. Nevertheless, we cautiously interpret calibrated model parameters (for an important discussion of this issue, see Kennedy and O'Hagan, 2001). The essential point is that since no model is correct, the 'true' value of a parameter almost certainly differs from its calibrated value. Our intent is to use the calibrated computer model as a measure of our systems-level understanding of *L. major* infection. If our prior parameter ranges are reasonable and the model is correct, then the posterior distributions reflect, to the best of our knowledge, the 'true' values of model parameters. Surprising or interesting posterior estimates may warrant special attention and should be further analyzed through biological experiments. Lethal parasite density  $k_l$  is such a parameter in our model. Relatively little is known about its *in vivo* value. What follows is a discussion of the calibration results of Section 4.4, where we let  $\theta$  vary according to Table C1 and fix  $A_{\text{necr}}$  at 2.

Despite the limited amount of field data used, we were able to learn information about most of the computer model parameters we estimated (Table 3). The high posterior median for lethal parasite density ( $k_l = 332$ ) suggests that *L. major* is capable of extensive intracellular replication without destroying the host cell, while the median growth rate (15 h doubling time), is relatively slow. As discussed earlier, these properties lead to longer term pathogen persistence and lower peak pathogen loads. They are further believed critical to the microbial virulence of *L. amazonensis* (Chang et al., 2002), and may partially explain the silent phase of *L. major* growth observed by Belkaid et al. (2000), although our calibration results did not consider this slow phase.

In the calibration process, we have assumed that field measurements are unbiased and, because the raw replicated field data were not available, that the field variance equals the computer model variance. Importantly, the assumption of unbiased field measurements may not hold for certain types of biological data, such as parasite counts, that are estimated through serial dilution assays. Typically, biologists use *ad hoc* methods to estimate concentrations in this manner, but more precise estimates with corresponding standard errors can be obtained through formal statistical treatment of the data (Gelman et al., 2004; Mehrabi and Matthews, 1995). Although Belkaid et al. (2000) do not provide the raw data, they do provide error bars for the two pathogen time points in their Fig. 1. Not surprisingly, since variability observed by Belkaid et al. (2000) includes measurement error and biological variation (Wong et al., 2005), the computer variance appears to underestimate the plotted error bars. Thus, underestimation of field variance implies our posterior distributions may be misleadingly narrow. With replicate field observations available, field variance could be estimated directly, or handled in a Bayesian context as part of a more accurate calibration process. In fact, because our model is a stochastic ABM, if measurement error is assumed constant, replicate observations would allow us to calibrate the model not only by matching the field data with model output means, but also the field variance with model output variances.

Our model does not capture the late, persistent stage of *L. major* infection, and our calibrated  $k_i$  seems to contradict the low pathogen loads measured during persistence. Although Belkaid et al. (2000) provide data for up to 10.5 wpi, we have chosen not to use any data after 8 wpi in our analysis. After 8 wpi, despite the fact that the infection is controlled, low levels of parasite persist (Belkaid et al., 2000). It is now known that regulatory T cells are required to maintain this low-level parasite persistence, which also provides host immunity to re-infection (Belkaid et al., 2003). Without modeling regulatory T cells, the computer model achieves complete parasite elimination (Fig. 3B), an outcome consistent with regulatory T cell depletion studies following *L. major* infection (Belkaid et al., 2003). We expect that inclusion of regulatory T cell function into our model will decrease the rate of parasite clearance, achieve a better fit at the last parasite time point  $t_6$ , and capture the low level parasite persistence. Nevertheless, the nature of pathogen persistence is curious in light of our high estimate for lethal parasite density  $k_i$ . Belkaid et al. (2000) observe the persistence of 100–10,000 parasites in the skin, which combined with our  $k_i$  estimate suggests that pathogens survive in a very small population of infected macrophages (between 1 and 30) at the site of infection. Although Bogdan et al. (2000) report pathogen persists in lymph node fibroblasts, it remains perplexing to us how parasites survive at the infection site, where it appears that only macrophages are infected. The *in vivo* value of  $k_i$  is not known, and it is unclear whether  $k_i$  changes as the infection progresses, but Belkaid et al. (2000) observe 'heavily infected macrophages' during the persistent stage of infection.

### 5.3. Model comparison and macrophage loss

When mean computer model behavior is considered, predictions from the calibrated computer model are accurate for most time points, but appear biased low at 5.5 wpi and high at 8 wpi for macrophages (time points  $t_2$  and  $t_4$ ) (Fig. 8). Together, these results suggest that at least one component of macrophage loss following infection resolution is missing or incomplete. Indeed, *in vivo* experiments involving peritonitis and the adoptive transfer of macrophages show that these macrophages are no longer detectable 96 h following the resolution of inflammation (Bellingan et al., 1996). The exact signals and mechanisms for this rapid macrophage loss are unclear, but macrophage loss could be triggered by long-range signals, such as decreasing cytokine levels, in addition to local phagocytotic events. In an earlier version of the model that assumed a constant macrophage recruitment rate, we modeled this behavior by arbitrarily assigning macrophages a short life span during the adaptive immune response (Dancik et al., 2006).

In our current model, we have proposed that macrophage activation results in the production of  $A_{necr}$  units of necrotic tissue, and that macrophages leave the site of infection following uptake of necrotic tissue. This mechanism of macrophage loss is supported by the high Bayes factor ( $BF_{10} = 7.3$ ), indicating 'substantial' evidence in favor of a model where macrophage behavior is altered in the presence of necrotic tissue. Our posterior estimate of  $A_{necr}$  is three (Fig. 9), which is one unit higher than the value of  $A_{necr}$  used in the original calibration. The large amount of uncertainty in our posterior estimate of  $A_{necr}$  (Table 4) is the result of assuming that both models are *a priori* equally likely, i.e.,  $\pi(A_{necr} = 0) = \pi(A_{necr} \in \{1, \dots, 6\}) = \frac{1}{7}$ . Using a uniform prior  $\pi(A_{necr}) = \frac{1}{7}, A_{necr} \in \{0, \dots, 6\}$  gives a 90% posterior credible interval of (1.4) and has little effect on the Bayes factor (data not shown).

Our proposed model of macrophage loss following uptake of necrotic tissue is one possible mechanism for explaining the

observed decrease in macrophages. Although it is possible to use other macrophage recruitment functions than Eq. (1), the decrease in macrophage numbers cannot be explained even by *entirely ceasing* recruitment after the macrophage peak (data not shown). While we imagine macrophage loss is via emigration to the draining lymph, loss could also be through death. Another explanation is that altered cytokine environments bias monocyte differentiation toward the dendritic cell type (Mohammadzadeh et al., 2001; Chomarat et al., 2003). However, if this differentiation takes place prior to monocyte influx into the tissue, it would be another recruitment effect. Furthermore, it cannot explain the rapid decrease in macrophages observed by (Belkaid et al., 2000), who distinguish between macrophages and dendritic cells in their collection assays. Interestingly, macrophage take-up and presentation of necrotic tissue (i.e., self-antigen) may partially explain the increase in natural regulatory T cells observed at the infection site during resolution of the infection (Cozzo et al., 2003; Mendez et al., 2004). Although *L. major*-specific natural regulatory T cells have been observed (Suffia et al., 2006), it is unclear why regulatory T cell levels do not increase until the infection is nearly resolved (Mendez et al., 2004).

Mathematical models of disease generally assume that the average lifespan of host cells is constant. A notable exception are models attempting to explain CD4+ T cell depletion in HIV (Perelson et al., 1992; Bernaschi and Castiglione, 2002; Yates et al., 2007). Our results indicate that in *L. major* infection, the average lifespan of a resting macrophage at the infection site is *not* constant, and our proposed mechanism of macrophage loss is not specific to *L. major* (Bellingan et al., 1996). We would expect similar macrophage behavior in other infections where necrotic tissue or, perhaps more generally, self antigen is present. It may be important to consider this non-constant macrophage behavior when modeling infection of other macrophage-tropic organisms, such as *M. tuberculosis* and *T. gondii*.

## 6. Conclusions and future work

In this work we described an agent-based model of *L. major* infection, used a GP to approximate and analyze the computer model, and estimated six immunological and pathogen-related parameters using field data. Simulations from the calibrated model are generally consistent with biological data, although more field data is needed in order to test the effects of the 19 parameters we did not investigate. In addition, our current model neither captures the long, slow, early phase of low dose infection nor the long-term persistence of *L. major* (Belkaid et al., 2003). We await proposals for specific biological mechanisms that explain these observations. Working within the dynamics of mid-stage infection, there is no question that our findings are influenced by the limited set of parameters we chose to vary. Nevertheless, our findings motivate interesting questions that may yield testable hypotheses. Some questions can be addressed *in silico*, for example whether the presence of necrotic tissue can explain the timing of regulatory T cell recruitment. Others require further biological investigation. For example, the high estimate of lethal pathogen density  $k_i$  within infected macrophages may imply that non-macrophage cellular reservoirs (e.g., fibroblasts) fuel the persistence of pathogen at the infection site.

## Acknowledgments

GMD and KSD were partially supported by PHS Grant GM068955, USDA IFAFS Multidisciplinary Graduate Education Training Grant (2001-52100-11506), and an ISU CIAG Research Support Grant.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:[10.1016/j.jtbi.2009.10.007](https://doi.org/10.1016/j.jtbi.2009.10.007).

## References

- An, G., 2002. Agent-based computer simulation and SIRS: building a bridge between basic science and clinical trials. *Shock* 16, 266–273.
- Badolato, R., Sacks, D.L., Savoia, D., Musso, T., 1996. *Leishmania major*: infection of human monocytes induces expression of IL-8 and MCAF. *Exp. Parasitol.* 82, 21–26.
- Bayarri, M., Berger, J.O., Cafeo, J., Garcia-Donato, G., Liu, F., Palomo, J., Parthasarathy, R.J., Paulo, R., Sacks, J., Walsh, D., 2006. Computer model validation with functional output. Technical Report Number 165, vol. 70, National Institute of Statistical Sciences, p. 102.
- Bayarri, M., Berger, J.O., Higdon, D., Kennedy, M., Kottas, A., Paulo, R., Sacks, J., Cafeo, J., Cavendish, J., Tu, J., 2002. A framework for the validation of computer models. In: Pace, D., Stevenson, S. (Eds.), *Proceedings of the Workshop on Foundations for V&V in the 21st Century*. Society for Modeling and Simulation International, San Diego, CA.
- Bayarri, M., Berger, J.O., Paulo, R., Sacks, J., Cafeo, J.A., Cavendish, J., Lin, C.-H., Tu, J., 2007. A framework for validation of computer models. *Technometrics* 49, 138–154.
- Beauchemin, C., 2006. Probing the effects of the well-mixed assumption on viral infection dynamics. *J. Theor. Biol.* 242, 464–477.
- Belkaid, Y., Mendez, S., Lira, R., Kadambi, N., Milon, G., Sacks, D., 2000. A natural model of *Leishmania major* infection reveals a prolonged “silent” phase of parasite amplification in the skin before onset of lesion formation and immunity. *J. Immunol.* 165, 969–977.
- Belkaid, Y., Piccirillo, C.A., Méndez, S., Shevach, E.M., Sacks, D.L., 2003. CD4+CD25+ immunoregulatory T cells control *Leishmania major* persistence and the development of concomitant immunity. *Nature* 420, 502–507.
- Bellingan, G.J., Caldwell, H., Howie, S.E.M., Dransfield, I., Haslett, C., 1996. In vivo fate of the inflammatory macrophage during resolution of inflammation. *J. Immunol.* 157, 2577–2585.
- Bernaschi, M., Castiglione, F., 2002. Selection of escape mutants from immune recognition during HIV infection. *Immunol. Cell. Biol.* 80, 307–313.
- Billack, B., 2006. Macrophage activation: role of Toll-like receptors, nitric oxide, and nuclear factor kappa B. *Am. J. Pharm. Educ.* 70, 102.
- Bogdan, C., Donhauser, N., Döring, R., Röllinghoff, M., Diefenbach, A., Rittig, M.G., 2000. Fibroblasts as host cells in latent leishmaniasis. *J. Exp. Med.* 191 (12), 2121–2129.
- Chang, K., Reed, S.G., McGwire, B.S., Soong, L., 2002. *Leishmania* model for microbial virulence: the relevance of parasite multiplication and pathoantigenicity. *Acta Trop.* 85, 375–390.
- Chomarat, P., Dantin, C., Bennett, L., Banchereau, J., Palucka, A.K., 2003. TNF skews monocyte differentiation from macrophages to dendritic cells. *J. Immunol.* 171, 2262–2269.
- Cozzo, C., Larkin, J.I., Caton, A.J., 2003. Cutting edge: self-peptides drive the peripheral expansion of CD4+CD25+ regulatory T cells. *J. Immunol.* 171, 5678–5682.
- Damian, R.T., 1997. Parasite immune evasion and exploitation: reflections and projections. *Parasitology* 115, 169–175.
- Dancik, G.M., Jones, D.E., Dorman, K.S., 2006. An agent-based model for *Leishmania* infection. *Interjournal Complex Syst.* 1853.
- Ewald, P., 1994. *Evolution of Infectious Disease*. Oxford University Press, USA.
- Forrest, S., Beauchemin, C., 2007. Computer immunology. *Immunol. Rev.* 216, 176–197.
- Fruth, U., Solioz, N., Louis, J.A., 1993. *Leishmania major* interferes with antigen presentation by infected macrophages. *J. Immunol.* 150, 1857–1864.
- Furth, R.V., Dulk, M.D.D., Mattie, H., 1973. Quantitative study on the production and kinetics of mononuclear phagocytes during an acute inflammatory reaction. *J. Exp. Med.* 138, 1314–1330.
- Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. *Bayesian Data Analysis*. Chapman & Hall, London.
- Gilks, W.R., Richards, S., Spiegelhalter, D.J., 1998. *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London.
- Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S.F., Thulke, H.-H., Weiner, J., Wiegand, T., DeAngelis, D.L., 2005. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *Science* 310, 987–991.
- Heitmann, K., Higdon, D., Nakhleh, C., Habib, S., 2006. Cosmic calibration. *Astrophys. J.* 646, L1–L4.
- Higdon, D., Kennedy, M., Cavendish, J.C., Cafeo, J.A., Ryne, R.D., 2004. Combining field data and computer simulations for calibration and prediction. *SIAM J. Sci. Comput.* 26, 448–466.
- Janeway, C.A., Travers, P., Walport, M., Shlomchik, M.J., 2005. *Immunobiology: The Immune System in Health and Disease*, sixth ed. Garland Science Publishing, New York.
- Kass, R.E., Raftery, A.E., 1995. Bayes factors. *J. Am. Stat. Assoc.* 90 (430), 773–795 <[citeseer.ist.psu.edu/539880.html](https://citeseer.ist.psu.edu/539880.html)>.
- Kennedy, M.C., O'Hagan, A., 2001. Bayesian calibration of computer models. *J. R. Stat. Soc. B* 63, 425–464.
- Lira, R., Doherty, M., Modi, G., Sacks, D., 2000. Evolution of lesion formation, parasitic load, immune response, and reservoir potential in C57BL/6 mice following high- and low-dose challenge with *Leishmania major*. *Infect. Immun.* 68, 5176–5182.
- Luder, C.K., Lang, T., Beuerle, B., Gross, U., 1998. Down-regulation of MHC class II molecules and inability to up-regulate class I molecules in murine macrophages after infection with *Toxoplasma gondii*. *Clin. Exp. Immunol.* 112, 308–316.
- Marino, S., Hogue, I.B., Ray, C.J., Kirschner, D.E., 2008. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *J. Theor. Biol.* 254 (1), 178–196.
- McMahon-Pratt, D., Alexander, J., 2004. Does the *Leishmania major* paradigm of pathogenesis and protection hold for New World cutaneous leishmaniasis or the visceral disease? *Immunology* 201, 206–224.
- Mehrabi, Y., Matthews, J.N.S., 1995. Likelihood-based methods for bias reduction in limiting dilution assays. *Biometrics* 51, 1543–1549.
- Mendez, S., Reckling, S.K., Piccirillo, C.A., Sacks, D., Belkaid, Y., 2004. Role for CD4+CD25+ regulatory T cells in reactivation of persistent leishmaniasis and control of concomitant immunity. *J. Exp. Med.* 200, 201–210.
- Mohamadzadeh, M., Berard, F., Essert, G., Chalouni, C., Pulendran, B., Davoust, J., Bridges, G., Palucka, A.K., Banchereau, J., 2001. Interleukin-15 skews monocyte differentiation into dendritic cells with features of langerhans cells. *J. Exp. Med.* 193, 1013–1020.
- Morris, M.D., Mitchell, T.J., 1995. Exploratory designs for computational experiments. *J. Stat. Plann. Infer.* 43, 381–402.
- Naderer, T., Ellis, M.A., Sernee, M.F., De Souza, D.P., Curtis, J., Handman, E., McConville, M.J., 2006. Virulence of *Leishmania major* in macrophages and mice requires the gluconeogenic enzyme fructose-1,6-bisphosphatase. *PNAS* 103, 5502–5507.
- Perelson, A.S., Kirschner, D.E., Boer, R.D., 1992. Dynamics of HIV infection of CD4+ T cells. *Math. Biosci.* 114, 81–125.
- Rajotte, D., Cadieux, C., Haman, A., Wilkes, B.C., Clark, S.C., Hercus, T., Woodcock, J.A., Lopez, A., Hoang, T., 1997. Crucial role of the residue R280 at the F-G loop of the human granulocyte/macrophage colony-stimulating factor receptor chain for ligand recognition. *J. Exp. Med.* 185, 1939–1950.
- Rencher, A.C., 2002. *Methods of Multivariate Analysis*, second ed. Wiley, New York.
- Roychoudhury, K., Roy, S., 2004. Role of chemokines in *Leishmania* infection. *Curr. Mol. Med.* 4, 691–696.
- Sacks, J., Welch, W.J., Mitchell, T.J., Wynn, H.P., 1989. Design and analysis of computer experiments. *Stat. Sci.* 4 (4), 409–435.
- Saltelli, A., Chan, K., Scott, E.M., 2000. *Sensitivity Analysis*. Wiley, New York.
- Santner, T.J., Williams, B.J., Notz, W., 2003. *The Design and Analysis of Computer Experiments*. Springer, New York.
- Schonlau, M., Welch, W., 2006. Screening the input variables to a computer model via analysis of variance and visualization. In: Dean, A., Lewis, S. (Eds.), *Screening: Methods for Experimentation in Industry, Drug Discovery, and Genetics*. Springer, New York, pp. 308–327.
- Segovia-Juarez, J.L., Ganguli, S., Kirschner, D., 2004. Identifying control mechanisms of granuloma formation in *M. tuberculosis* infection using an agent-based model. *J. Theor. Biol.* 231, 357–376.
- Shapiro, M., Duca, K.A., Lee, K., Delgado-Eckert, E., Hwkins, J., Jarrah, A.S., Laubenbacher, R., Polys, N.F., Hadinoto, V., Thorley-Lawsome, D.A., 2008. A virtual look at Epstein-Barr virus infection: simulation mechanism. *J. Theor. Biol.* 252, 633–648.
- Suffia, I.J., Reckling, S.K., Piccirillo, C.A., Goldszmid, R.S., Belkaid, Y., 2006. Infected site-restricted Foxp3+ natural regulatory T cells are specific for microbial antigens. *J. Exp. Med.* 203 (3), 777–788.
- Sunderkotter, C., Kunz, M., Steinbrink, G., Meinardus-Hager, M., Goebeler, H.B., Song, C., 1993. Resistance of mice to experimental leishmaniasis is associated with more rapid appearance of mature macrophages in vitro and in vivo. *J. Immunol.* 151, 4891–4901.
- Sypek, J.P., Panosian, C.B., Wyler, D.J., 1984. Cell contact-mediated macrophage activation for antileishmanial defense. *J. Immunol.* 133, 3358–3365.
- Tranquillo, R.T., Lauffenburger, D.A., Zigmond, S.H., 1988. A stochastic model for leukocyte random motility and chemotaxis based on receptor binding fluctuation. *J. Cell. Biol.* 106, 303–309.
- Vanloubbeek, Y., Jones, D.E., 2004. The immunology of *Leishmania* infection and the implications for vaccine development. *Ann. N. Y. Acad. Sci.* 1026, 267–272.
- Wei, X.Q., Charles, I.G., Smith, A., Feng, G.H., Huang, F.P., Xu, D., Muller, W., Moncada, S., Liew, F.Y., 1995. Altered immune response in mice lacking inducible nitric oxide synthase. *Nature* 375, 408–411.
- Winchester, G., Sunshine, G., Nardit, N., Mitchison, A., 1984. Antigen-presenting cells do not discriminate between self and nonself. *Immunogenetics* 19, 487–491.
- Wong, A., Gottesman, I., Petronis, A., 2005. Phenotypic differences in genetically identical organisms: the epigenetic perspective. *Hum. Mol. Genet.* 14 (Review Issue 1), R11–R18.
- Wooldridge, M., 2002. *Introduction to Multiagent Systems*. Wiley, New York.
- Yates, A., Stark, J., Klein, N., Antia, R., Callard, R., 2007. Understanding the slow depletion of memory CD4+ T cells in HIV infection. *Plos Med.* 4 (5), e177.