BCG vaccination reduces the rate of Mycobacterium tuberculosis dissemination between murine lungs

Dipanjan Chakraborty[1*] and Vitaly V. Ganusov[1,2*]

[1]Host-Pathogen Interactions (HPI) program, Texas Biomedical Research Institute, San Antonio, TX 78227, USA

[2]Molecular Microbiology and Immunology department, University of Texas at San Antonio, San Antonio, TX 78249

[*]Corresponding authors: dipanjan.cu13@gmail.com, vitaly.ganusov@gmail.com

March 10, 2026

Abstract

The BCG vaccine remains the only licensed vaccine against tuberculosis (**TB**), yet the mechanisms behind BCG-induced protection remain poorly understood. Plumlee et al. (PLOS Pathogens 2023) infected over 1,000 mice, half of which were vaccinated with BCG, with an ultra-low dose (**ULD**) of Mycobacterium tuberculosis (**Mtb**); the authors found that BCG vaccination resulted in fewer infected mice, lower CFU lung burden, and more frequent unilateral lung infection. We have developed several mathematical models of Mtb dynamics and dissemination between murine right and left lungs and fit these models to the CFU data from unvaccinated or BCG-vaccinated mice. Alternative mathematical models incorporating either direct (lung-to-lung) or indirect (lung-intermediate-tissue-lung) dissemination pathways fit the unvaccinated data equally well, suggesting multiple plausible routes of Mtb spread. Yet, irrespective of the dissemination route, the models predicted rapid Mtb replication during early infection, transient control within 1–2 months after infection, and continued bacterial growth in the chronic phase. Fitting models to the data from BCG-vaccinated animals revealed that BCG reduces the rate of Mtb dissemination between the lungs by 89% while having a more modest effect on the replication rate within the lung, reducing it by 9%. We found that the dominant effect of BCG in curbing lung dissemination arises from its ability to reduce Mtb replication resulting in fewer infected mice, lower lung CFU, and decreased bilateral infection of the lung. We used our parameterized mathematical models to calculate the number of mice needed to detect the efficacy of a hypothetical vaccine on the probability of Mtb clearance or dissemination between murine lungs that extends previously provided estimates. Taken together, our novel mathematical modeling-based framework provides a rigorous way of quantifying vaccine efficacy in ULD-infected mice, paving the way for the pre-clinical evaluation of next-generation TB vaccines.

Abbreviations: CFU — colony forming unit, Mtb — *Mycobacterium tuberculosis*, TB — tuberculosis, MDR – multi-drug resistant, ODE — ordinary differential equation, SSA — stochastic simulation algorithm, ULD – ultra-low dose, DD – direct dissemination, ID – indirect dissemination, LOD – limit of detection, CI – Confidence interval.

Keywords: Tuberculosis, Mycobacterium tuberculosis, mathematical model, within-host dynamics, BCG vaccine, disseminated disease.

Author summary

The Bacillus Calmette–Guérin (BCG) vaccine remains the only licensed vaccine against tuberculosis (TB), a disease caused by Mycobacterium tuberculosis (Mtb) bacteria. BCG is clearly protective for several years when given at birth but the mechanisms by which it provides protection remain incompletely understood. By combining analysis of Mtb dynamics in over 1,000 control and BCG-vaccinated mice, infected with an ultra low dose of Mtb, with mathematical models we show that BCG reduces both the rate of Mtb replication in the lung and the rate of Mtb dissemination between the lungs. Importantly, our model suggests that BCG may thus reduce a chance of being infected and that BCG-mediated prevention of disseminated TB in humans could be due vaccine's ability to block Mtb dissemination out of the lung. Our novel framework thus will allow to predict efficacy of the next generation TB vaccines in settings of infection at doses that humans are thought to be typically exposed to.

Introduction

Tuberculosis (**TB**), a disease caused by *Mycobacterium tuberculosis* (**Mtb**), remains one of the prevalent infectious diseases in humans[1]. Each year, approximately 10 million indivduals develop TB, and many individuals, particularly children and immunocompromised adults, develop extrapulmonary disease, in which Mtb disseminates beyond its primary site of infection, the lung[2,3]. Even though drug-susceptible TB can be effectively treated within 4-6 months in most individuals and relatively short and effective treatments are now available for patients with multi-drug resistant (**MDR**) TB[4], controlling TB prevalence in humans will require additional tools such as an effective vaccine[5].

The Bacillus Calmette–Guérin (**BCG**) vaccine, first used to immunize a human volunteer in 1921, remains the only licensed vaccine against TB in humans[6–8]. Initial randomized, placebo-controlled trials conducted in the UK and Scandinavia demonstrated substantial protection of BCG vaccine against pulmonary and extrapulmonary TB especially in children in 2-3 years of the follow up[9,10]; however, later studies found more variable efficacy of BCG with some studies, such as the Chingleput study in India, reporting no protection against TB[11]. A more recent meta-analysis of 18 clinical trials confirmed that BCG provides consistent and robust protection against severe childhood TB manifestations including miliary and meningeal forms[12–15]. Yet, mechanisms by which BCG induces protection of children against TB remain poorly understood and typically are difficult to study in humans[16–18].

Multiple animal species have been used to quantify BCG-mediated protection against Mtb infection and to identify mechanisms of such protection[18–20]. In general, BCG vaccination induces measurable protection of mice, rats, guinea pigs, rabbits, and macaques against Mtb challenge; vaccinated animals typically have reduced bacterial burdens in the lungs and/or extrapulmonary sites (e.g., spleen), reduced lung pathology, and longer survival[21–39]. In particular, in typical mouse strains such as C57BL/6 and BALB/c, BCG vaccination reduces lung bacterial loads by $0.5-1\log_{10}$, and guinea pigs (species highly susceptible to Mtb infection) after BCG vaccination have up to $2\log_{10}$ reduced lung CFU as compared to unvaccinated controls, measured several weeks after infection[21,31]. Interestingly, BCG vaccination can induce sterilizing protection in rhesus macaques[27,32]. Overall, previous studies have suggested that BCG is likely to mediate protection via a variety of different mechanisms, including generation of trained innate immunity, T cell memory, and antibodies[18,19,23,33,37,38].

It has been well recognized that Mtb dose has a strong impact on the course of infection in animals (e.g., lung CFU numbers, lung pathology, and survival time)[28,40]; yet, whether degree of protection conferred by BCG vaccine depends on the Mtb exposure dose has remained relatively unexplored. Plumlee *et al.*[39] found that exposing BCG-vaccinated B6 mice to a conventional dose (**CD**) of Mtb H37Rv (about 100 CFU/mouse) resulted in $\sim 1\log_{10}$ reduction in lung CFU as compared to unvaccinated controls at day 42 post infection; however, over time (at day 120) the difference in lung CFU between control and BCG-vaccinated mice became reduced and statistically non-significant. Interestingly, exposing control and BCG-vaccinated mice to an ultra-low dose (**ULD**) of Mtb H37Rv (with predicted effective dose of $\sim$ 1 CFU/mouse) resulted in a smaller proportion of mice with detectable CFU (i.e., infected), lower lung CFU (by about $1\log_{10}$ for 120 days), and less frequent bilateral infection of both right and left lungs[39]. However, this study did not investigate how BCG vaccination-impacted metrics such as infection rate, lung CFU, bilateral infection may be related; for example, whether reduction in the frequency of bilateral infection is simply a result of a lower lung CFU or if BCG reduces the per capita dissemination rate of bacteria between the lungs was not determined.

Here we used mathematical modeling to study impact of BCG vaccination on the probability of infection, Mtb growth in the lung and dissemination between right and left lungs in settings of ULD infection[41]. Interestingly, we found that mathematical models that assume direct (lung to lung) or indirect (lung to intermediate tissue to lung) dissemination pathways fit the data from unvaccinated animals with similar quality suggesting that existing data are insufficient to discriminate between the alternatives[42]. Yet, independently of the model of Mtb dissemination we found that BCG vaccination reduced both the rate of Mtb replication in the lung and the rate of Mtb dissemination between the lungs; alternative models in which BCG increased the Mtb clearance rate or had no impact on dissemination did not fit the data well. Importantly, model predicted BCG-induced reduction in Mtb replication rate naturally leads to higher infection clearance probability (by 9%) and reduced (by about 1 $\log_{10}$) total CFU in the lung in vaccinated animals. We used our mathematical modeling-based framework to evaluate impact of growth- or migration-reducing vaccine on Mtb dynamics and dissemination and provided power analyses on the number of mice that would be required to detect efficacy of a hypothetical vaccine. Taken together, our work establishes a predictive and mechanistic lens for interpreting efficacy of the BCG vaccine and serves as a foundation for evaluating next-generation TB vaccines in ULD infection setting.

Materials and Methods

Experimental data

Plumlee *et al.*[41] have developed a protocol in which a standard culture used for conventional dose (**CD**) infections is diluted $\sim$ 50-fold, filtered, and used in a nebulizer to produce ultra-low dose (**ULD**) aerosol to infect mice in the infection chamber. In a typical ULD experiment, about 60-80% of the mice get infected (as determined by CFU measured at later time points for all mice in the experiment), and the number of founding strains detected in individual mice using barcoded bacteria after ULD infection is consistent with a Poisson distribution[41]. In a recent study, Plumlee *et al.*[39] used the same protocol to infect over 1,000 of B6 mice with an ULD of H37Rv Mtb ($\sim$ 1 CFU/mouse, **Suppl. Fig. S1**A); approximately half of mice ($n = 531$) were subcutaneously immunized 8 weeks earlier with 10^6 BCG-Pasteur and half were left unimmunized ($n = 521$). In total, the authors performed 21 experiments.

For most of our analyses, we included only infected mice (i.e., mice with CFU > 0 in both lungs); excluding uninfected mice or mice with no CFU data for right and left lungs yielded CFU data for $n = 300$ unvaccinated and $n = 256$ BCG-vaccinated mice (**Suppl. Fig. S2**). In the experiments, Plumlee *et al.*[39] measured Mtb CFU at different days post-infection in the right (**RL**) and left (**LL**) lungs separately (**Suppl. Fig. S2**); our previous studies suggested that a bias in the initial infection towards the right lung most likely arises due to its larger volume/weight[41] (**Suppl. Fig. S1**B). While infection could be initiated in any of the lungs (or in some cases in both lungs if there are 2+ bacteria entering the lung), it was difficult to envision one deterministic model that would incorporate such stochastic outcomes of the initial infection step. Therefore, to allow fitting ordinary differential equation(ODE)-based mathematical models to CFU data in the RL and LL we re-arranged the data by assigning larger CFU values (between RL and LL in individual mice) as Lung 1 (L_1) and the smaller value as Lung 2 (L_2), with the assumption that bacteria initiate the infection in the Lung 1 (with CFU = 1), and that the infection disseminates over time to from Lung 1 to Lung 2 (**Suppl.**

Figs. S2 and S3 and **Fig. 1**).

Mathematical models

To study Mtb infection dynamics in the lungs, we use a compartmental ODE-based modeling framework that focuses directly on predicting bacterial burden (CFU) over time[41,43,44]. Building on this approach, we explored several alternative mathematical models to describe Mtb dynamics in the initially infected lung (Lung 1) and its dissemination to the collateral lung, Lung 2 (**Fig. 1**A&B). We re-organized the data as Lung 1 and Lung 2 (see above) and assume that the infection is initiated in Lung 1 (with CFU = 1) but over time Mtb disseminates to Lung 2 (**Fig. 1** and see above). We consider two alternative pathways of Mtb dissemination between Lung 1 and Lung 2: (i) In the direct dissemination (**DD**) model (**Fig. 1**A), Mtb directly disseminates from the initially infected lung (Lung 1) to the collateral lung (Lung 2); (ii) in the indirect dissemination (**ID**) model (**Fig. 1**B), Mtb reaches the collateral lung (Lung 2) via an intermediate tissue (e.g., blood or spleen).

Direct dissemination (DD) model. We have used a simple system of linear ODEs to model the dynamics of Mtb (**Fig. 1**A&B):

$$\frac{\mathrm{d}L_1(t)}{\mathrm{d}t} = (1-\epsilon_r)r_L(t)L_1(t) - \frac{\delta(t)}{1-\epsilon_\delta}L_1(t) - (1-\epsilon_m)mL_1(t), \tag{1}$$

$$\frac{\mathrm{d}L_2(t)}{\mathrm{d}t} = (1-\epsilon_m)mL_1 \; + \; \theta(L_2-1)\Big[(1-\epsilon_r)r_L L_2 - \frac{\delta(t)}{1-\epsilon_\delta}L_2(t)\Big], \tag{2}$$

where $L_1(t)$ and $L_2(t)$ represent the Mtb numbers in Lung 1 and Lung 2 at time t; $r_L(t)$ denote the replication rate of Mtb; $\delta(t)$ is the death rate of Mtb; m is the migration rate of Mtb from Lung 1 to Lung 2 and ϵ_r, ϵ_m, and ϵ_δ quantify the effects of BCG vaccination on replication rate, migration rate, and death rate of Mtb, respectively. The Heaviside theta function $\theta(x)$ in **eqn. (2)** indicates that replication and death of Mtb begin in Lung 2 only when the number of disseminated Mtb exceeds 1 CFU. This modification was done to prevent rapid exponential growth of bacteria in Lung 2 when only a tiny fraction of a bacterium ($\ll 1$) is disseminated from Lung 1.

In this model formulation $\epsilon_r = 0$, $\epsilon_m = 0$, and $\epsilon_\delta = 0$ correspond to unvaccinated mice. For non-zero values of ϵ_r, ϵ_m, and ϵ_δ, we can explore different models by choosing different combinations of ϵ_r, ϵ_m, ϵ_δ. The death rate, $\delta(t)$, is time-dependent but is kept the same for both Lung 1 and Lung 2. We adopted the death rate estimates reported by McDaniel *et al.*[43], where these parameters are obtained by fitting multiple alternative models to experimental data[45]:

$$\delta(\mathrm{t}) = \begin{cases} 0.40/\mathrm{day}, & 0 \le \mathrm{t} \le 16 \text{ days}, \\ 0.13/\mathrm{day}, & 16 < \mathrm{t} \le 26 \text{ days}, \\ 0.13/\mathrm{day}, & 26 < \mathrm{t} \le 42 \text{ days}, \\ 0.13/\mathrm{day}, & 42 < \mathrm{t} \le 125 \text{ days}. \end{cases} \tag{3}$$

To maintain consistency with the time-dependent formulation of the death rates, we modeled the Mtb replication rates as time-dependent functions, denoted by $r_L(t)$:

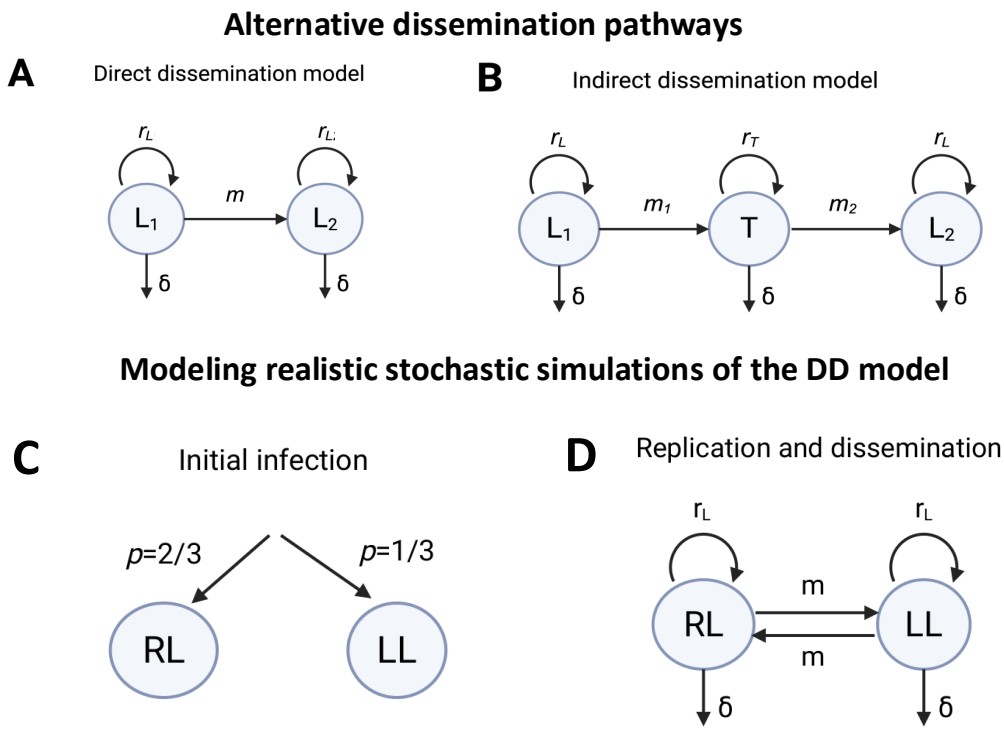


Figure 1: Schematics of alternative dissemination pathways of Mtb between the lungs of ULD-infected mice. In our main ODE model-based analyses we assume that infection with one bacterium (**Suppl. Fig.** S1A) starts in Lung 1 (L_1) where bacteria replicate and die at rates r_L and δ, respectively. In the direct dissemination (**DD**) model (panel **A** and **eqns.** (1)–(2)) Mtb disseminates from Lung 1 to Lung 2 at a rate m, and in the indirect dissemination (**ID**) model (panel **B** and **eqns.** (5)–(7)), Mtb first disseminates from Lung 1 to an intermediate tissue T at a rate m_1 and then from tissue T disseminates to Lung 2 at a rate m_2. Mtb replicates and dies in both models in the Lung 2 at rates r_L and δ, respectively. In the ID model, Mtb replicates and dies in the intermediate tissue at rate r_T and δ, respectively. In our stochastic simulations (dubbed as "realistic" simulations, panels **C-D**) we allow the infection to occur with a Poisson-distributed number of bacteria (with average CFU = 1/mouse, **Suppl. Fig.** S1A) that infect the right lung (**RL**) or left lung (**LL**) with a probability 2/3 or 1/3, respectively (**Suppl. Fig.** S1B and **C**). Mtb then replicates and dies in the lungs at rates r_L and δ, respectively, and disseminates between the lungs at a rate m (**D**).

$$r_L(t) = \begin{cases} r_1, & 0 \leq \mathrm{t} \leq 16 \text{ days}, \\ r_2, & 16 < \mathrm{t} \leq 26 \text{ days}, \\ r_3, & 26 < \mathrm{t} \leq 42 \text{ days}, \\ r_4, & 42 < \mathrm{t} \leq 125 \text{ days}, \end{cases} \tag{4}$$

where we assume that the replication rate of Mtb at Lung 1 and 2 is the same. The migration rate of Mtb from Lung 1 to Lung 2 (m) is assumed to be independent of time. Each of the key terms in the model equations (**eqns.** (1)–(2)), the replication-driven growth term, the immune-mediated decay term, and the migration term representing dissemination is proportional to the bacterial load at a given time point, reflecting the assumption that these processes scale with the number of Mtb present. Since this DD model mimics the ULD Mtb infection in mice, we started our simulation with 1 bacterium (as the average effective dose is 1 CFU, **Suppl. Fig.** S1) in Lung 1 and 0 bacterium in

Lung 2.

Indirect dissemination (ID) model. In the indirect dissemination model, the spread of Mtb to the collateral lung occurs through an intermediate tissue, e.g., blood or spleen (**Fig. 1**B). Therefore, the ID model has another equation for the Mtb dynamics in the intermediate compartment. If the intermediate tissue is the spleen, that allows for Mtb replication, the equations describing Mtb dynamics in Lung 1, tissue, and Lung 2 will be the following:

$$\frac{dL_1(t)}{dt} = (1-\epsilon_r)r_L(t)L_1(t) - \frac{\delta(t)}{1-\epsilon_\delta}L_1(t) - (1-\epsilon_m)m_1L_1(t), \tag{5}$$

$$\frac{dT(t)}{dt} = \theta(T-1)\big[(r_T(t)T(t) - \delta(t)T(t))\big] + (1-\epsilon_{m_1})m_1L_1(t) - m_2(1-\epsilon_m)T(t), \tag{6}$$

$$\frac{dL_2(t)}{dt} = \theta(L_2-1)\big[(1-\epsilon_r)r_L(t)L_2(t) - \frac{\delta(t)}{1-\epsilon_\delta}L_2(t)\big] + m_2(1-\epsilon_m)T(t), \tag{7}$$

where $r_T(t)$ indicates the Mtb replication rate at the intermediate tissue T, m_1 is the migration rate of Mtb from Lung 1 to tissue, and m_2 is the migration rate of Mtb from tissue to Lung 2. If the tissue is blood, then there will be no replication and death of Mtb within the blood, i.e., $r_T = \delta_T = 0$; only the migration term would exist, signifying blood as a carrier of Mtb but not the place of extensive replication (**eqns. (S.1)–(S.3)**). The initial condition for starting the simulation is kept the same as described in the DD model. Further, the replication and death of Mtb in the tissue and Lung 2 would only occur when the number of disseminated Mtb crosses 1 CFU, as indicated by the Heaviside function $\theta(x)$ in **eqns. (6)–(7)**.

Fitting the direct dissemination (DD) and indirect dissemination (ID) models to data

To fit both the DD and ID models to the CFU data in mouse lungs we used nonlinear least-squares minimization with the **Levenberg–Marquardt (Marq)** optimization algorithm implemented in the `modFit` routine of the `FME` package in `R`. We log-transformed model predictions and the data and minimized the sum of squared residuals to find the best fit. Prior to analysis, mice with no detectable Mtb burden in whole lung were removed from the dataset. Consequently, a CFU value of zero occurred only when one lung was infected, and the contralateral lung had not yet acquired bacteria. Because these zero measurements cannot be log-transformed, all data were shifted by adding one CFU, such that

$$\text{CFU}^* = 1 + \text{CFU}, \tag{8}$$

so that an experimental observation such as ($L_1 = 5$, $L_2 = 0$) becomes ($L_1 = 6$, $L_2 = 1$) prior to log transformation. This shift ensures that $\log_{10}(\text{CFU}^*)$ remains well-defined while maintaining the biological ordering of the observed infection levels.

To initialize the ODE solver in R consistently with the experimental data, a data point at $t = 0$ was introduced which is the initial condition that we assumed but the associated 0 residual was later removed in calculating various additional characteristics (e.g. AIC)[46]. To compare the DD and ID

models, we used the Akaike Information Criterion (**AIC**) and the associated Akaike weights (w)[47]. The model with the lowest AIC (and thus highest Akaike weight) was selected as the most plausible among alternative models under consideration.

Stochastic simulations of the $L_1 \rightarrow L_2$ version of the DD model

To model Mtb dynamics stochastically we converted the ODE version of a model (e.g., DD model) into discrete probabilistic events where Mtb replication, death, and migration occur with time-dependent transition rates derived from the best-fitted parameters of the ODE model (**Table. 1**). These transition rules (**Table. 1**) define how the system evolves through random events rather than via continuous population changes. For each simulation, the system was initialized with a single bacterium in Lung 1 and none in Lung 2 (**Suppl. Fig. S1**A). Because stochastic realizations yield only one possible outcome of a random process, multiple independent trajectories (typically 100-1,000) were generated to assess whether the variability captured by the model matches the experimental CFU distributions across lungs. Simulations were performed for 125 days using Gillespie's exact stochastic simulation algorithm, implemented in the `GillespieSSA2` package in `R`[48,49] and recorded with a step size of $\Delta t = 1.0$ day.

Reaction	Rate of Reaction
$L_1 \rightarrow L_1 + 1$	$r_L L_1$
$L_1 \rightarrow L_1 - 1$	δL_1
$L_1 \rightarrow L_1 - 1,\ L_2 \rightarrow L_2 + 1,$	$m L_1$
$L_2 \rightarrow L_2 + 1$	$r_L L_2$
$L_2 \rightarrow L_2 - 1$	δL_2

Table 1: List of reactions and their rates of the stochastic simulations of the $L_1 \rightarrow L_2$ version of the DD model. In the model, increase or decrease in the number of bacteria in Lung 1 occurs at the rate $r_L L_1$ or δL_1, respectively, and migration of bacteria from Lung 1 to Lung 2 occurs at the rate $m L_1$ (**eqns. (1)–(2)**). Note that the rates of Mtb replication and death in the Lung 1 and 2 are time-dependent (**eqns. (3)–(4)**).

Realistic stochastic simulations of the RL $\rightleftarrows$ LL version of the DD model

In the $L_1 \rightarrow L_2$ version of the DD model we assumed a unidirectional dissemination pattern where bacterial infection was initiated in Lung 1, and it subsequently spread to Lung 2 (**Table. 1**). However, in reality, the primary site of infection may arise in either left or right lung, followed by dissemination to the contralateral lung (**Fig. 1**C&D), and the initial number of inhaled bacteria could be higher than 1 CFU due to random nature of infection (**Suppl. Fig. S1**A). To incorporate these additional biological details, we modified stochastic simulations of the DD model to allow for randomness of lung infection and bidirectional migration between the two lungs (**Table. 2**).

First, we modeled the initial bacterial inoculum as a random draw from a Poisson distribution with mean $\lambda = 1$ (or any other dose), consistent with experimental measurements of infection dose variability (**Suppl. Fig. S1**A). Second, based on previous experimental observations[41], we introduced an asymmetry in the initial infection probability, assigning a 2/3 chance that a bacterium

Reaction	Rate of Reaction
$RL \to RL + 1$	$r_L RL$
$RL \to RL - 1$	δRL
$RL \to RL - 1$, $LL \to LL + 1$,	mRL
$LL \to LL + 1$	$r_L LL$
$LL \to LL - 1$	δLL
$LL \to LL - 1$, $RL \to RL + 1$,	mLL

Table 2: List of reactions and their rates of the realistic stochastic version of the DD model. In this model, Mtb infection is initiated in the right lung (RL) or the left lung (LL) with a probability $2/3 : 1/3$ (**Fig.** 1C). Thereafter, the increase or decrease in the number of bacteria in RL occurs at a rate $r_L RL$ or δRL, respectively. Migration of bacteria from RL to LL and LL to RL occurs at a rate $m.RL$ or mLL respectively. Replication rates and death rates of Mtb in the RL and LL are assumed to be the same as those of Lung 1 and Lung 2 (**eqns. (3)–(4)**).

lands in the right lung and a 1/3 chance for it to land in left lung (**Fig.** 1C and **Suppl. Fig.** S1B). Finally, we used parameters for replication (r_L), death (δ), and migration (m) as estimated from fitting ODE version of the DD model to data to simulate Mtb dynamics stochastically in a similar fashion as was described above (e.g., **Fig.** 1C&D).

Quantifying efficacy of hypothetical vaccines in restricting bacterial dissemination

Using realistic stochastic simulations of the DD model (**Fig.** 1C&D and **Table.** 2) we evaluated how a hypothetical vaccine influences the spread of Mtb between the lungs. We quantified the dissemination outcomes under two vaccine efficacy parameters: ϵ_r, which reduces the rate of bacterial replication, and ϵ_m, which reduces the rate of inter-lung dissemination. We assessed the probability that a mouse develops a bilateral infection, a clinically relevant marker of dissemination severity. We initialized simulations with a virtual cohort of 100 mice, using parameter values derived from the best-fitted model. For each parameter setting (e.g., varying ϵ_r while fixing ϵ_m), we calculated the fraction of mice that developed bilateral infections at day 14 post infection relative to the total number of infected mice (with limit of detection LOD = 1). Because stochastic models produce variability across realizations, we performed 100 independent simulations for each combination of ϵ_r and ϵ_m. The results were then averaged to obtain estimate of bilateral infection probability in different levels of vaccine efficacy.

Quantitative assessment of stochastic model fit using the Kolmogorov-Smirnov (KS) test

To evaluate the degree of agreement between stochastic simulation trajectories and experimental data, we propose to use the Kolmogorov–Smirnov (**KS**) test and KS statistics[50]. This non-parametric test quantifies the maximum distance between the cumulative distribution functions (**CDFs**) of the simulated and experimental data. In this context, the KS statistic (denoted as $D \in (0, 1)$) serves as a measure of the goodness of fit: smaller values indicate closer correspondence between model predictions and observed outcomes. We calculated an average KS distance $\overline{D}$ across multiple

measurement days to account for temporal variability in bacterial counts as follows

$$\overline{D_{L_j}} = \frac{1}{n}\sum_{i=1}^{n} D_i, \tag{9}$$

where i indexes the measurement days, n is the total number of days, and j denotes the lung under consideration (e.g., $j = 1, 2$ for Lung 1 and Lung 2 or RL and LL). A lower value of $\overline{D_{L_j}}$ implies improved model fidelity. In the idealized case, $\overline{D_{L_j}} = 0$ would indicate perfect agreement; however, due to intrinsic stochasticity and biological variability, such an outcome is not expected.

Model-based power analysis for detecting vaccine effects

Statistical power analysis determines the minimum number of animals needed to detect a treatment effect with a desired confidence. Plumlee *et al.*[39] estimated the sample sizes required to distinguish impact of a vaccine on probability of infection at different levels of statistical power. Building on this idea, we performed a model-based power analysis using our realistic stochastic simulations of the DD model to assess how vaccine efficacy influences these sample size requirements. We examined two types of vaccines: vaccine reducing Mtb replication rate (ϵ_r) or vaccine reducing Mtb dissemination rate (ϵ_m). In the first case, we computed the infection probabilities in unvaccinated and vaccinated groups of mice by varying the efficacy of vaccination in replication (ϵ_r). For a given sample size, these probabilities were summarized using a 2×2 contingency table, and statistical significance of group differences was evaluated using the χ^2 test ($p \leq 0.05$). By iterating across sample sizes, we identified the minimum cohort size required to achieve 80% power. Similarly, to assess dissemination control, we performed power analysis by examining the frequencies of bilateral lung infections in unvaccinated and vaccinated mice, under different levels of vaccine efficacy in blocking dissemination rate (ϵ_m). The outcomes were organized into contingency tables, and χ^2 testing was again applied to assess statistical significance across unvaccinated and vaccinated groups. In the calculation of power analysis, while exploring the effect of ϵ_r, we set $\epsilon_m = 0$, and similarly, when varying ϵ_m, we fixed $\epsilon_r = 0$.

Results

Mathematical modeling suggests multiple plausible routes of Mtb dissemination

Plumlee *et al.*[39] exposed control or BCG-vaccinated B6 mice to an ultra-low dose (**ULD**) of Mtb and measured CFU in the right and left lungs at different times after infection; in total of 21 experiments, over 1,000 mice were used. Early after the exposure, bacteria were found primarily in one of the lung but over time, both lungs became infected. It is well understood that in mice, Mtb typically disseminates to extrapulmonary sites (e.g., lymph nodes or spleen)[51–53]; however, whether systemic dissemination is required for infection of both lungs especially in settings of ULD infection has not been determined.

Mtb dynamics at ULD infections is intrinsically stochastic due to a low number of founding bacteria[41] and we initially could not come up with a simple deterministic (ODE-based) mathematical

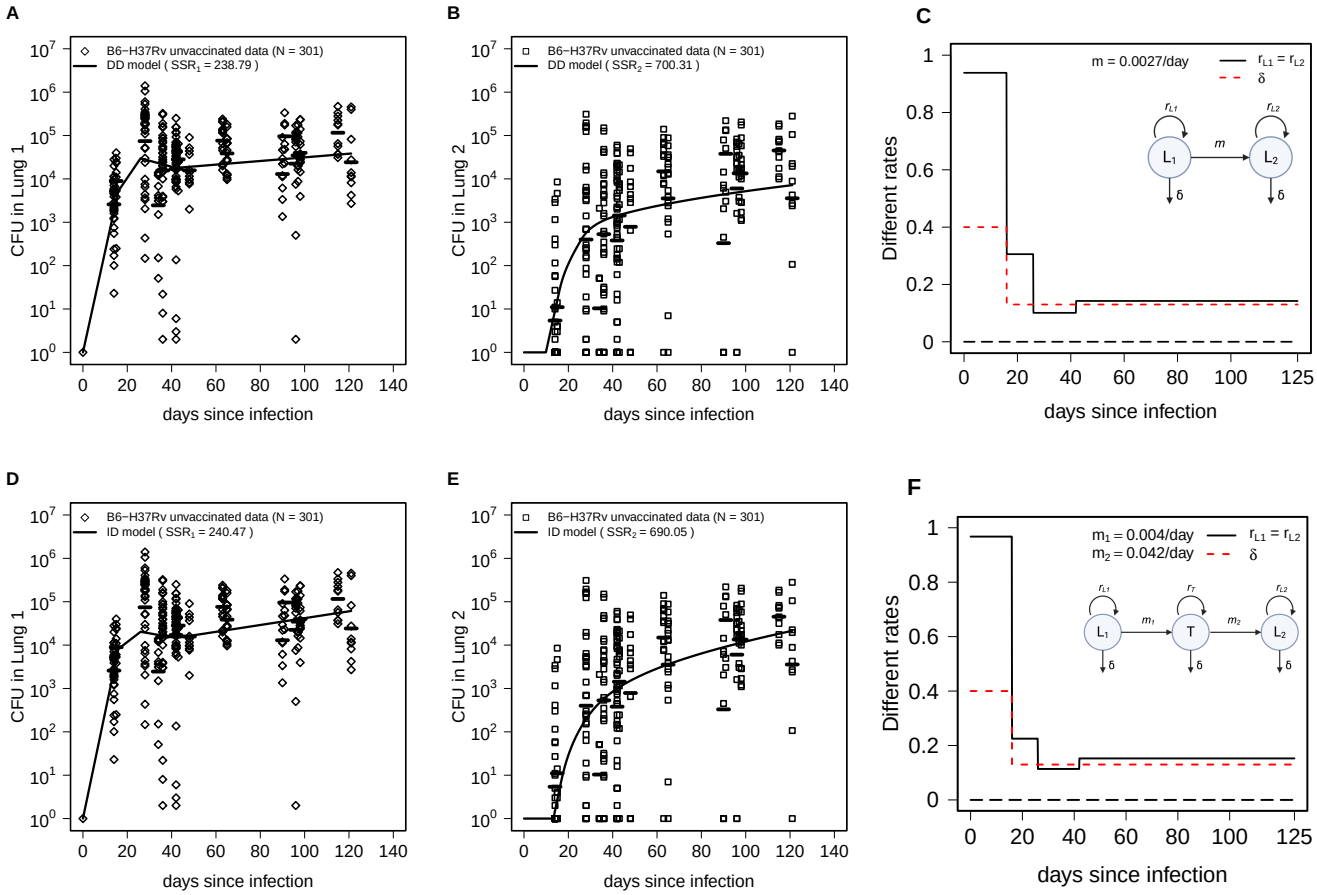


Figure 2: Alternative mathematical models assuming direct or indirect Mtb dissemination between murine lungs describe the data on Mtb dynamics in unvaccinated mice with similar quality. We fit the direct dissemination model (**eqns. (1)–(2)**, **Fig. 1**A, **A-C**) or indirect dissemination model (**eqns. (5)–(7)**, **Fig. 1**B, **D-F**) to CFUs found in lungs of unvaccinated mice (after re-arranging the data as Lung 1 and Lung 2, see Materials and methods for detail). We show the data as markers and model fits as lines in A-B and D-E along with estimated rates of Mtb replication in the lung (C&F) or in the intermediate tissue (**Suppl. Figs. S4 and S5** and **Suppl. Table. S1**). In the model fits we assumed $L_1(0) = 1$ and $L_2(0) = 0$. The quality of the model fit was evaluated by sum of squared residuals (**SSR**) or AIC (noted on individual panels; see also **Suppl. Table. S2**).

model to describe Mtb dynamics and dissemination between the right and left lungs. Because in Plumlee *et al.*[39] experiments, the average effective dose was 1 CFU/mouse (**Suppl. Fig. S1**A), per Poisson distribution most infections ($\sim 60\%$) would be initiated by a single bacterium landing in the right or left lung. Therefore, we opted for a simplified mathematical model that assumes that infection is initiated by a single bacterium in one of the lungs (denoted as Lung 1) and over time, the bacteria disseminate to the collateral Lung 2. The dissemination may be direct or indirect (**Fig. 1**A-B): the direct dissemination (**DD**) model assumes Mtb migration directly from Lung 1 to Lung 2, while the indirect dissemination (**ID**) model assumes that Mtb passes via an intermediate tissue; Mtb may be replicating in the intermediate tissue (e.g., spleen) or just being transported between the lungs (e.g., via blood).

To fit the models to data we re-classified CFU in the right or left lung as CFU in Lung 1 or 2 depending on the CFU value in the two lungs for each individual mouse (i.e., larger CFU value was

assigned to Lung 1 and the smaller value to Lung 2, **Suppl. Figs. S2 and S3**). This procedure is likely accurate when CFU in one of the lungs is much larger than in another but may introduce noise when CFU numbers in both lungs are similar. Thereafter, we fit the DD model (**eqns. (1)–(2)**) to the modified data on CFU in Lung 1 and Lung 2 (see Materials and methods for details); the model accurately captured rapid rise in CFU in Lung 1 followed by bacterial control and decline in CFU after the peak, and then slow increase in CFU over time (**Fig. 2**A). The model also accurately predicted continuous accumulation of CFU in the collateral lung (Lung 2) over time (**Fig. 2**B). In the model, CFU dynamics was determined by changes in the Mtb replication rate that had relatively high values early in the infection ($r_L \approx 0.94$/day), declined to low levels ($r_L \approx 0.1$/day), but then increased in the chronic phase (after 40 days $r_L \approx 0.14$/day) remaining slightly higher than the death rate leading to slow increase in CFU number over time (**Fig. 2**C & **Suppl. Table. S1**). Estimated Mtb per capita dissemination rate $m = 2.7 \times 10^{-3}$/day appears to be quite small yet sufficient to cause Mtb dissemination to the collateral lung by 2-3 weeks post infection (**Fig. 2**B). Interestingly, allowing for the Mtb replication rate to differ between Lung 1 and Lung 2 did not improve the model fit (F-test for nested models, $p = 0.24$), suggesting that changes in the rate of Mtb replication with time in both lungs may be set at the time of infection.

We next fit the ID model (**eqns. (5)–(7)**) that assumes that Mtb is able to replicate in an intermediate tissue (**Fig. 1**B), to the data. Interestingly, the model could also accurately describe CFU in the Lung 1 and 2 at somewhat similar rates of Mtb replication (**Fig. 2**D-F & **Suppl. Table. S1**); the model predicted relatively slow migration rate from Lung 1 to the intermediate tissue but then rapid migration from the tissue to Lung 2 (**Fig. 2**F). The model fits also predicted relatively rapid Mtb replication in the intermediate tissue (**Suppl. Fig. S4**A); interestingly, the model-predicted CFU in the tissue was consistent (but slightly lower than the mean) with the independently measured bacterial numbers in the spleens of ULD-infected mice at day 63 post infection (**Suppl. Fig. S4**B). Importantly, an alternative version of the ID model that assumes that dissemination between murine lungs occurs by passive transport (without replication), e.g., via blood, also fit the data well (**Suppl. Fig. S5**). Thus, our analysis suggests that both direct and indirect dissemination routes of Mtb spread between individual lungs are possible, and additional data would be required to discriminate between alternatives[42]. Statistically, however, the DD model fit the data with slightly better quality (based on AIC and Akaike weights) in part because of the smaller number of model parameters (**Suppl. Table. S2**).

BCG vaccination reduces Mtb replication and dissemination rates

We next turned to evaluating impact of BCG vaccination on Mtb dynamics and dissemination in ULD-infected mice. Visual inspection of the data indicated lower CFU in both right and left lung (or Lung 1 and 2, **Suppl. Figs. S2 and S3**) suggesting impact on net Mtb replication rate. To minimize the number of parameters fitted to data (and thus to reduced the likelihood of overfitting) we opted for the approach in which we fixed estimates of Mtb replication and dissemination rates to values estimated from unvaccinated mice but allowed the overall vaccine efficacy at reducing Mtb replication rate ϵ_r or dissemination rate ϵ_m (see **eqns. (1)–(2)**), and fitted the model by varying only ϵ_r and/or ϵ_m. Interestingly, we found that in order to accurately fit the data both DD and ID models required that BCG vaccination reduces both Mtb replication and dissemination rates (**Fig. 3**): the DD model predicted 9% and 89% reduction in the Mtb replication and dissemination, respectively, while the ID model (with lymph node/spleen being intermediate tissue) predicted rate 9% and 65% reduction in replication and dissemination rate, respectively (**Fig. 3**). The ID model more accurately

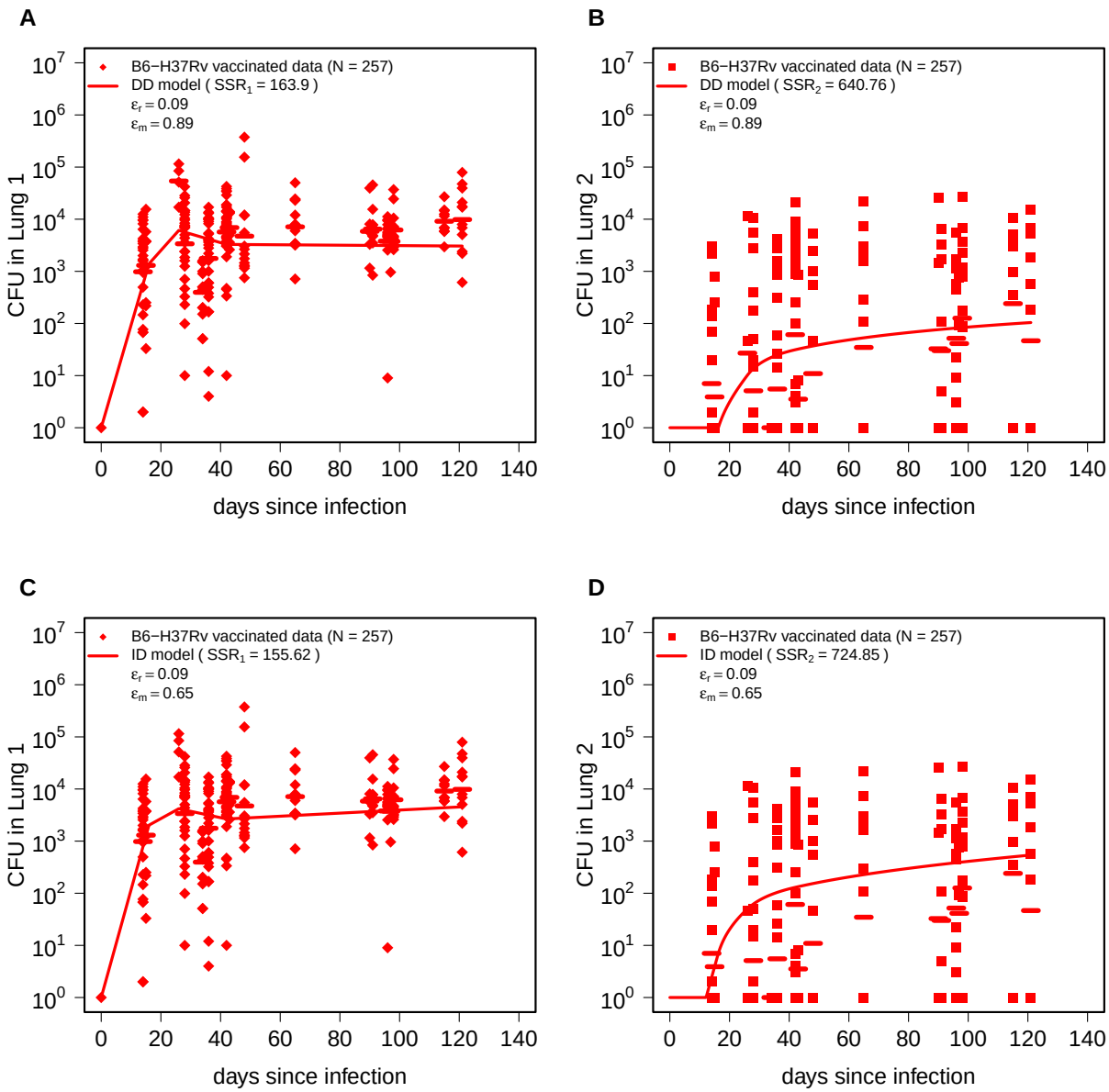


Figure 3: Both direct and indirect dissemination models predict that BCG vaccination significantly reduces Mtb dissemination rate and slightly reduces Mtb replication rate. Using the estimates of Mtb replication (r_L) and dissemination (m) rates for unvaccinated mice (**Fig. 2**) we fit the DD (**eqns. (1)–(2), A-B**) or ID (**eqns. (5)–(6), C-D**) models to CFU data on Mtb dynamics in Lung 1 (**A&C**) and Lung 2 (**B&D**) in BCG-vaccinated mice. The data are shown by markers, and best-fit model predictions are shown by lines. The estimated efficacy of BCG vaccine at reducing the rate of Mtb replication (ϵ_r) and Mtb dissemination rate (ϵ_m) is shown on individual panels (see also **Suppl. Table. S2**).

fit the data on Lung 1, while the DD model better fit the data for Lung 2 (note SSR values in **Fig. 3**). Statistically, the DD model fit the data slightly better (based on AIC values, **Suppl. Table. S2**).

It was notable that while the DD model accurately predicted average CFU in Lung 1 in the first 45 days post infection, the model consistently predicted lower average lung CFU than that observed experimentally (**Fig. 3**A&C). We wondered if this could arise if efficacy of BCG vaccine at reducing the rate of Mtb replication may be time-dependent. We therefore re-fitted the DD model by allowing ϵ_r to change between different time intervals (i.e., see **eqn. (3)**). Surprisingly, the model fits predicted decline of BCG vaccine efficacy with time since infection being 8-15% in the first 4 weeks of infection but then declining to 0.2-3% at later times (**Suppl. Fig. S7**). Thus, ULD infection of mice also recapitulates decline in efficacy of BCG vaccine with time[54].

Our results show that, independently of the assumed mode of Mtb dissemination, BCG vaccination reduces the Mtb replication rate in the lung; this finding is important for two reasons. First, it naturally explains lower CFU numbers observed in BCG-vaccinated mice. While the average 9% reduction in replication rate may appear small, in about 25 days such a difference in replication rates will result in 10 fold difference in CFU numbers as observed experimentally (**Suppl. Fig. S2**). Second, $\epsilon_r = 9\%$ reduction in the rate of Mtb replication naturally results in an identical increase in the probability of infection clearance when exposed to a dose of 1 CFU (infection extinction probability for 1 CFU is $p_0 = \delta/r$); thus, our analysis of the data from only infected mice made a similar inference of reduced infection probability of BCG-vaccinated mice as was done by Plumlee *et al.*[39] by looking at the proportion of infected mice.

Along with the models that fit the data well (e.g., **Fig. 3**), it is also important to note that several alternative versions of the DD or ID models did not fit the data (**Suppl. Table. S2**). For example, the model assuming that BCG vaccination only reduces the rate of Mtb replication poorly fit the data (**Suppl. Fig. S6** and **Suppl. Table. S2**); this important result suggests that reduced bilateral infection of the lungs of BCG-vaccinated mice is not a simple consequence of reduced CFU numbers (arising due to slower Mtb replication rate) and that BCG vaccination somehow impairs the ability of Mtb to disseminate in the lung (or outside of the lung). The model assuming that BCG vaccination increases Mtb death rate (with ϵ_δ, **eqns. (1)–(2)**) also did not fit the data well (**Suppl. Table. S2**). This result is somewhat consistent with our previous observation that in CD infection, the rate of Mtb replication changes with time since infection more often than death rate does[43].

Stochastic simulations of $L_1 \rightarrow L_2$ version of the DD model fails to reproduce observed CFU variability

One of the key features of CFU data from ULD-infected mice is their large variability that typically is not observed in CD-infected mice[41]. When fitting ODE-based models to such variable data we make an implicit assumption that the data variability comes from measurement noise; however, given that in these experiments infection starts with very few bacteria (**Suppl. Fig. S1**A), variability in cell numbers may also arise due to stochasticity of Mtb replication, death, and migration (i.e., process noise)[41]. We therefore sought to investigate whether stochastic simulation of Mtb dynamics using best fit parameters of a model would match CFU variability measured experimentally.

To simulate Mtb dynamics stochastically we 1) converted the deterministic, ODE-based $L_1 \rightarrow L_2$ version of the DD model into a set of transition rules (**Table. 1**); 2) used parameters from the best

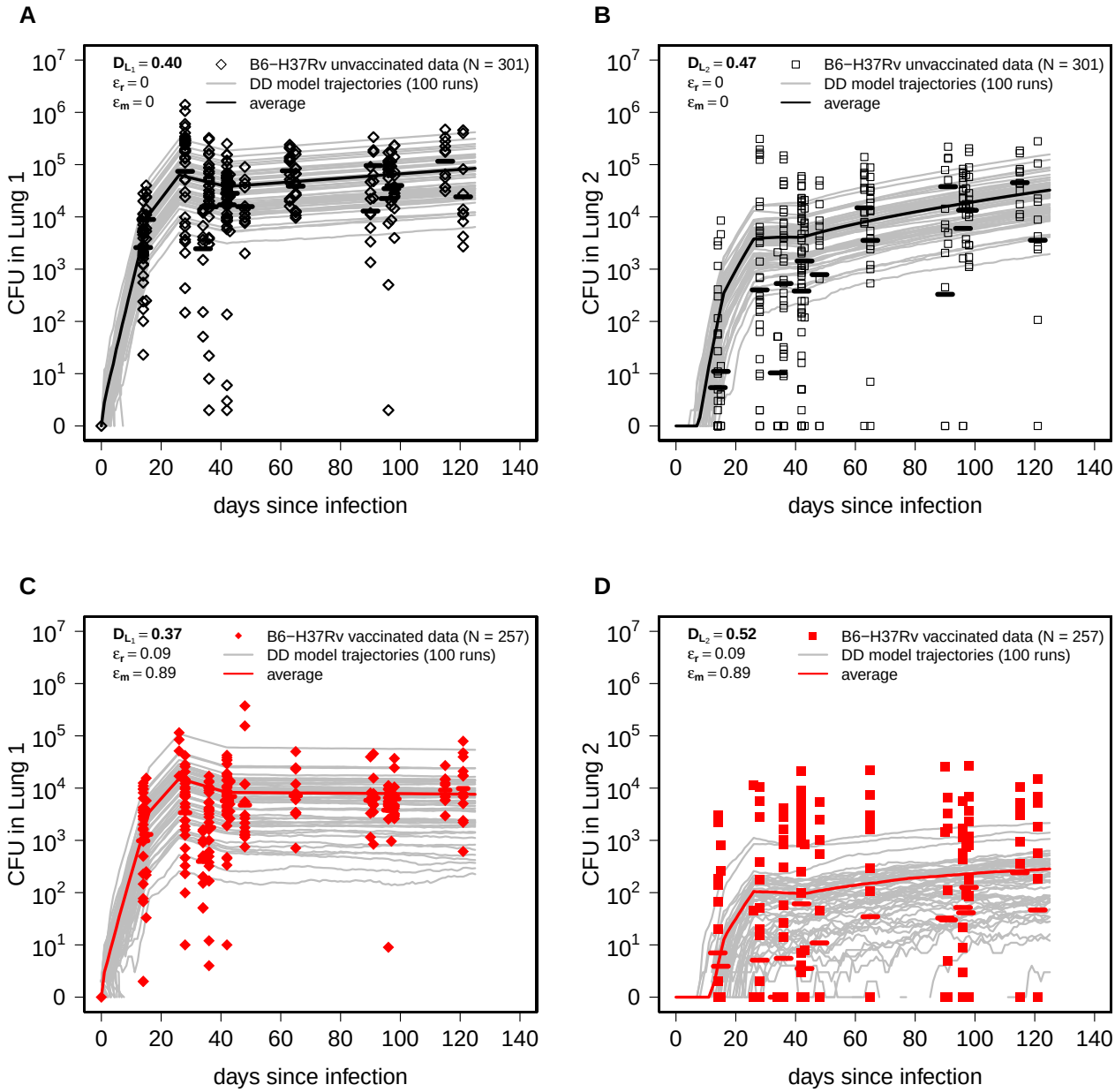


Figure 4: Stochastic simulations of the DD model fail to capture full variability of CFUs in the data. We translated the ODE-based DD model (**eqns. (1)–(2)**) into a corresponding set of transition rules (see Materials and methods for detail) and performed stochastic simulations using the best-fit parameter estimates (**Figs. 2 and 3**) for both unvaccinated (panels **A-B**) and BCG-vaccinated (panels **C-D**) mice. Gray lines denote individual runs of the model (total $n = 100$ runs) and solid lines denote the average of all simulated trajectories. To compare how well stochastic simulations match experimental data we calculated the Kolmogorov–Smirnov (KS) test statistic D (shown on individual panels) denoting the distance between cumulative distribution of CFUs predicted by the model and that observed in the data (**eqn. (9)**, see Materials and methods for detail).

fits of the DD model to the data from unvaccinated or BCG-vaccinated mice (**Fig. 2**A-B and **Fig. 3**A-B); 3) set the initial conditions as $L_1(0) = 1$ and $L_2(0) = 0$, and 4) ran 100 Gillespie simulations (and see Materials and methods for detail) for each of a set of model parameters for unvaccinated and BCG-vaccinated mice (**Fig. 4**). As expected, stochastic simulations predicted divergent trajectories for CFU for individual runs with the average predicted CFU somewhat matching experimentally observed values; however, it was clear that there were some data points that stochastic trajectories did not describe. In particular, while the model relatively well match variability in CFU numbers in the Lung 1, it failed to accurately predict larger CFU values in Lung 2 (**Fig. 4** B&D).

To more rigorously evaluate the match between stochastic simulations and experimental data we used Kolmogorov-Smirnov (**KS**) test that measures distance between two distributions (**eqn. (9)** and see Materials and methods for detail). Similarly to the visual comparison, KS test metric was higher for model predictions in the Lung 2 as compared to Lung 1, and in all cases D values were relatively high ($D = 0.37 - 0.52$) suggesting that stochastic simulations of the $L_1 \rightarrow L_2$ version of the DD model were not able to fully capture variability of experimental data.

Realistic stochastic simulations of infection of right and left lungs better capture CFU variability

The mismatch between stochastic realizations of the DD model and experimental data may have arisen because of simplifying assumptions in the modeling framework. First, in the DD model we assumed that infection starts with a single bacterium; however, while this is probably correct for most ULD infections with an effective dose of 1 CFU/mouse, there is a substantial chance that two or three bacteria may initiate the infection[41]. Second, the DD model assumes that infection always starts in Lung 1; however, we have previously established that when the number of founding bacteria is greater than one, both lungs may be infected with the right murine lung having about 2 fold higher probability of infection (**Suppl. Fig. S1**B). The latter assumption is likely to result in under-predicted numbers of bacteria in the Lung 2 (**Fig. 4**B&D).

We therefore extended stochastic simulations of the DD model to include additional biological detail of ULD infection. First, we introduced random sampling of the actual number of infecting bacteria from the Poisson distribution assuming that the average dose is 1 CFU/mouse (**Suppl. Fig. S1**A). In case when the sampled number is above zero, we partitioned the sampled bacteria between right or left lung with $2/3 : 1/3$ probability (**Fig. 1**C). We then modelled Mtb growth in the right or left lung with estimated replication (r_L) and death (δ) rates, and possibility of migration between the lungs at rate m (**Figs. 2 and 3**) assuming direct dissemination between the lungs (**Fig. 1**D). We dubbed this model as "realistic stochastic simulations" of the DD model.

Importantly, the realistic simulations of Mtb dynamics provided a larger range of predictions of bacterial numbers for both right and left lungs and more accurately predicted both large and small CFU numbers in BCG-vaccinated mice (**Fig. 5**). The improved visual match of the data was confirmed by lower values of KS statistics (compare **Fig. 4** and **Fig. 5**). Yet, D values remained relatively high suggesting that this model still did not fully capture CFU variability observed experimentally. Another interesting observation was that the average CFU predicted by the model did not accurately match averages observed in experiments (**Fig. 5**) suggesting that parameter estimates generated in the DD model may not be fully accurate for the realistic stochastic simulations.

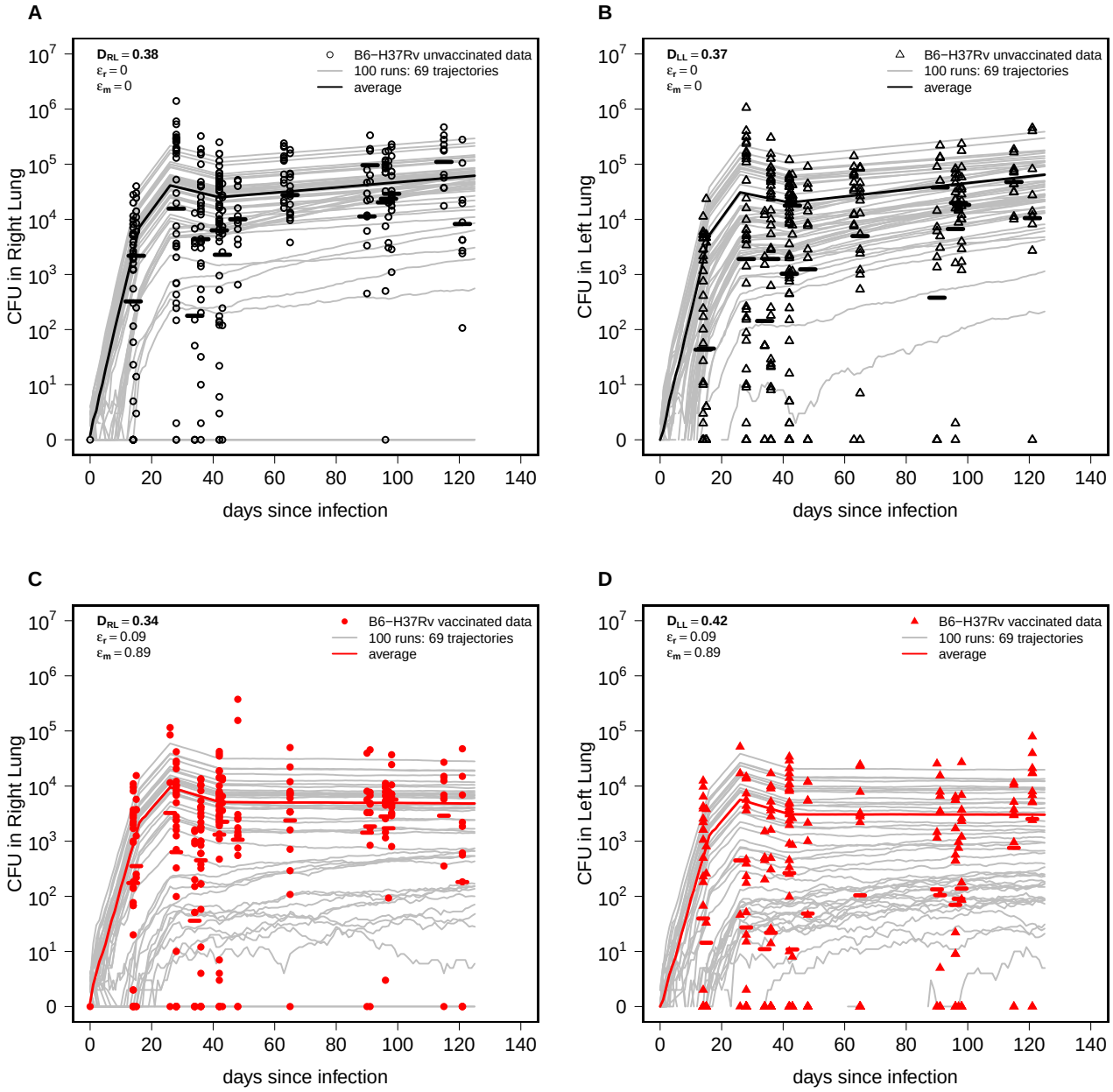


Figure 5: Stochastic simulations of Mtb infection, growth and dissemination between right and left lungs better match experimental data. Using the estimated parameters from the best fit DD model (**Figs. 2 and 3**), we ran the stochastic simulations assuming random infection of the right and left lungs and dissemination between the lungs (**Fig.** 1C-D) for unvaccinated (**A-B**) or BCG-vaccinated (**C-D**) mice. In simulations, we randomly chose the number of bacteria initiating infection from a Poisson distribution with an $\lambda = 1$ (**Suppl. Fig.** S1A), and used a probability of depositing a bacterium in the right lung of 66% (**Fig.** 1C and **Suppl. Fig.** S1B). We quantified the quality of the model match of experimental data using KS statistics D (**eqn. (9)** and Materials and methods for more detail).

Hypothetical vaccines reducing Mtb replication rate more effectively limit Mtb dissemination

The DD model predicted that BCG-induced reduction in the rate of Mtb replication results in smaller lung CFU, and, as a consequence, in reduced frequency of bilateral lung infection; however, this was insufficient to fully explain smaller CFU in the Lung 2 (**Suppl. Fig. S6** and **Suppl. Table. S2**). Having established that realistic stochastic simulations of the DD model relatively well explain variability of bacterial numbers in the lungs of unvaccinated or BCG-vaccinated mice we sought to investigate how a hypothetical vaccine that exclusively reduces the rate of Mtb replication or the rate of Mtb dissemination between the lungs, would influence the frequency of bilateral infection.

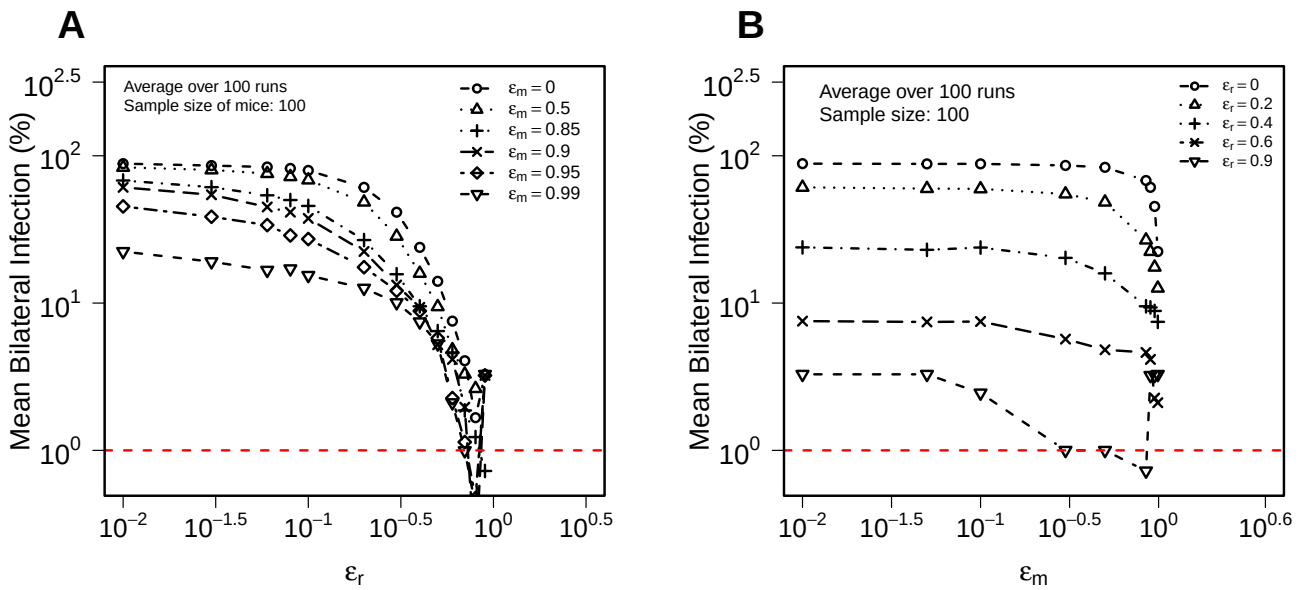


Figure 6: A vaccine reducing the rate of Mtb replication has a greater impact on probability of Mtb dissemination in the lungs as compared to a vaccine that exclusively blocks Mtb dissemination. We run stochastic simulations of Mtb replication in and dissemination between RL and LL (**Fig. 1**C-D) using parameter estimates of DD model fit to CFU data in unvaccinated mice (**Fig. 2**) for 100 mice vaccinated with a vaccine that reduces the rate of Mtb replication (ϵ_r) or a vaccine that reduces the rate of Mtb dissemination between lungs (ϵ_m); note that model predictions for $\epsilon_r = \epsilon_m = 0$ denote unvaccinated mice (value 10^{-2} on the x axis actually corresponds to $\epsilon_r = 0$ or $\epsilon_m = 0$). We calculate the percent of mice that have detectable CFU in both lungs (i.e., with bilateral infection) at day 14 post-infection with limit of detection LOD = 1 for varying levels of ϵ_r (**A**) or ϵ_m (**B**). We performed 100 runs per one set of parameters and show the average percent of bilateral infection.

As expected increasing efficacy of a vaccine blocking the rate of Mtb replication ϵ_r reduced the frequency of bilateral infection (**Fig. 6**A); interestingly, the impact of replication rate-reducing vaccine was moderate for smaller vaccine efficacy ($\epsilon_r \leq 0.3$), and was relatively independent on the efficacy of the vaccine reducing the rate of Mtb dissemination ϵ_m (**Fig. 6**A). In contrast, increasing efficacy of a vaccine that reduces the rate of Mtb dissemination had relatively minor impact on the frequency of bilateral infection unless vaccine efficacy is relatively high ($\epsilon_m > 0.9$, **Fig. 6**B). These results suggest that vaccines that reduce the rate of Mtb replication even slightly should have stronger impact on the ability of Mtb to disseminate between the lungs than vaccines that only reduce ability of Mtb to disseminate between the lungs.

Power analysis for detecting the vaccine efficacy

To detect small efficacy of BCG vaccine at reducing probability of infection of mice exposed to ULD dose, Plumlee *et al.*[41] had to combine data from 21 experiments, in total involving over 1,000 mice. It was not clear, however, whether these experiments were powered enough to actually detect a small efficacy of the vaccine ($\sim 10\%$) at preventing infections. In addition, it was not previously possible to also determine how many mice would be required to detect efficacy of a hypothetical vaccine that only reduces the rate of Mtb dissemination. We therefore used realistic stochastic simulations of the DD model to perform such power analyses, i.e., to investigate the number of mice that would be required to use in an experiment to detect efficacy of a vaccine that reduces the rate of Mtb replication ϵ_r or reduces the rate of Mtb dissemination in the lung ϵ_m; this is similar to an approach we have outlined recently[55–57].

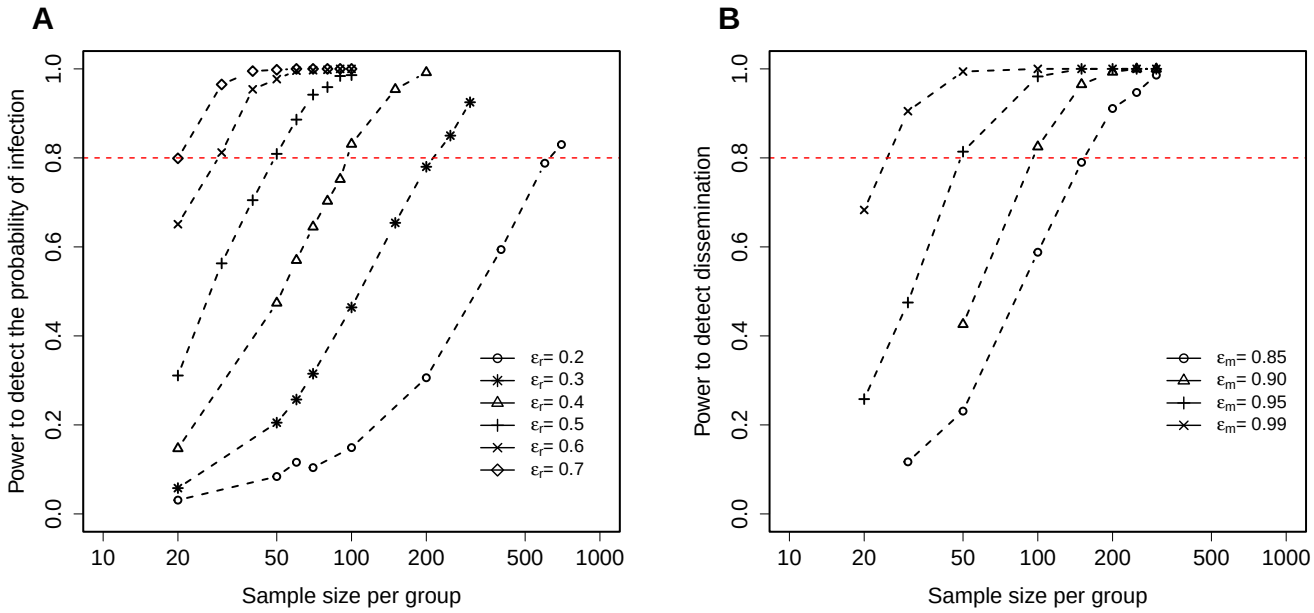


Figure 7: Realistic stochastic simulations of the DD model predict the number of animals required to detect efficacy of vaccines preventing Mtb infection or Mtb dissemination between lungs. We ran stochastic simulations of Mtb replication in and dissemination between RL and LL (**Fig. 1**C-D) in unvaccinated mice and mice vaccinated with a vaccine that either reduces the rate of Mtb replication (ϵ_r, **A**) or reduces the rate of Mtb dissemination between the lungs (ϵ_m, **B**) and calculated the proportion of mice that are uninfected (**A**) or have unilateral infection (**B**) by day 14 post infection (with LOD = 1). By varying the number of unvaccinated and vaccinated mice we calculated the probability of detecting vaccine efficacy of a particular value (ϵ_r and ϵ_m are noted on the panels, and see **Table. 3**).

We ran the model simulations by assuming 1) different numbers of mice per unvaccinated or vaccinated group, 2) different values of the vaccine efficacy ϵ_r or ϵ_m, and 3) by calculating the proportion of "experiments" (out of 1000 simulations) that result in statistically significant difference in proportion of infected mice (for ϵ_r) or difference in proportion of mice with bilateral infection (for ϵ_m, see Materials and methods for detail). As expected, increasing the number of mice per group resulted in larger power to detect a given level of vaccine efficacy and fewer mice were needed to detect larger efficacy levels (**Fig. 7** and **Table. 3**). In particular, relatively small numbers of mice ($n_{\text{mice}} = 20$) are needed to detect highly efficient replication rate-reducing vaccine ($\epsilon_r \sim 0.7$, **Fig. 7**A), and even at moderate levels of vaccine efficacy ($\epsilon_r = 0.5$), only 50 mice (per group) would be

needed to detect reduction in the frequency of infected vaccinated mice (**Table. 3**). In contrast, dissemination-blocking vaccines would require relatively large numbers of mice ($n_{\text{mice}} = 100$) to detect high vaccine efficacy levels ($\epsilon_m = 0.90$, **Fig. 7**B). This is perhaps unsurprising given that Mtb dissemination may occur continuously during the infection, and preventing dissemination to the collateral lung when CFU are sufficient large would require high vaccine efficacy. In addition, even in ULD infection setting, many mice would be infected with two or more bacteria in both lungs reducing the ability to detect efficacy of a dissemination-blocking vaccine.

Category	**Efficacy Value**	**Minimum sample size per group (from experiment[39])**	**Model predicted sample size per group**
Vaccination efficacy on replication (ϵ_r)	0.2	259	629
	0.3	112	215
	0.4	66	97
	0.5	40	50
	0.6	28	29
	0.7	21	21
	0.8	16	17
	0.9	12	-
Vaccination efficacy on dissemination (ϵ_m)	0.85	-	155
	0.90	-	97
	0.95	-	50
	0.99	-	26

Table 3: Calculation of number of mice needed to detect the efficacy of a hypothetical vaccine. We compared the sample size estimates required to achieve 80% power for different levels of vaccine efficacy, controlling replication (ϵ_r) or dissemination (ϵ_m). The power analysis is implemented by using realistic stochastic simulation to predict the sample size per group, as described in the Materials and methods section. Model-predicted sample sizes are calculated when measuring total lung CFU or bilateral infection at day 14 post-infection assuming LOD = 1; we also list the sample size predicted in Plumlee *et al.*[39]. "-" denotes not available.

Plumlee *et al.*[39] also provided power analysis to detect efficacy of a vaccine preventing infection in ULD-exposed mice. Their analysis was based on experimentally observed frequencies of infection assuming that unvaccinated mice are exposed to ~ 0.96 CFU/mouse. Interestingly, our stochastic simulation-based power calculations matched those by Plumlee *et al.*[39] at high efficacy of replication-blocking vaccine ($\epsilon_r \geq 0.5$) but diverged considerably at lower vaccine efficacy (**Table. 3**). In particular, our simulations suggested that to detect reduction of infection due to replication-blocking vaccine with efficacy $\epsilon_r = 0.2$, 629 mice per group would be required (**Table. 3** and **Fig. 7**A), and well over 1,000 mice per group would be needed to detect lower vaccine efficacy of $\epsilon_r = 0.1$ (not shown). The fact that we did detect a relatively small efficacy of replication-blocking vaccine ($\epsilon_r = 0.09$, **Fig. 3**) with about 500 mice/group suggests that time course CFU data may have more information about efficacy of BCG vaccine at blocking Mtb replication than measurements of the frequency of infected mice. This model inference remains to be tested in future experiments.

Discussion

In this study, we developed and analyzed alternative deterministic and stochastic mathematical models to investigate within-host Mtb dynamics at ULD infections of mice. By fitting direct and indirect dissemination models to experimental CFU data from Plumlee *et al.*[39], we demonstrated that

both models can successfully explain the observed Mtb kinetics, predicting rapid initial replication, transient control, and subsequent persistence (**Fig. 2**). After incorporating the effect of BCG vaccination, both models consistently indicated that BCG markedly reduces Mtb dissemination rate between lungs by approximately $\epsilon_m = 65 - -89\%$ while only modestly lowering bacterial replication rates ($\epsilon_r = 9\%$, **Fig. 3** and **Suppl. Table. S2**). Interestingly, we found similar estimates of BCG vaccine-induced reduction of Mtb replication rate by fitting a model that incorporates immune response-mediated control of Mtb growth to data from ULD or CD-infected mice (**Suppl. Fig. S8**). Realistic stochastic simulations of the DD model that capture randomness of actual number of bacteria inhaled by individual mice and possibility of bidirectional migration between right and left lungs relatively accurately reflect the variability of experimental CFU data (**Fig. 5**). Realistic stochastic simulations showed that a hypothetical vaccine that reduces the rate of Mtb replication has a larger effect at preventing bilateral infection as compared to a vaccine that only reduces Mtb dissemination rate (**Fig. 6**). Finally, our power analysis provided quantitative estimates of sample sizes needed to detect efficacy of vaccines blocking Mtb replication or dissemination (or both), offering a framework for optimizing preclinical study designs (**Fig. 7** and **Table. 3**).

Our modeling results align closely with recent animal studies and clinical evidence supporting BCG vaccine-mediated protection against TB. In particular, our results are consistent with the experimental findings of Plumlee *et al.*[39] where by infecting mice with barcoded Mtb strain the authors found reduced bilateral lung infection (by nearly 80%) in BCG-vaccinated mice. While their study quantified dissemination efficacy based on the frequency of bilateral infections, our modeling framework captures the underlying kinetics of Mtb replication and migration. Our conclusion that BCG vaccine-induced reduction in Mtb replication rate may be declining with time since infection (**Suppl. Fig. S7**) is consistent with observation in humans documenting decline in BCG vaccine efficacy over time[54]. Collectively, previous experimental studies and our results converge on a common interpretation: BCG provides partial but consistent protection that limits bacterial dissemination and disease severity rather than fully preventing the Mtb infection[58].

Clinical and epidemiological data also support this interpretation. Decades of human studies, from early randomized placebo-controlled trials in the UK and Scandinavia to more recent trials such as Chingleput in India, have revealed substantial variability in BCG efficacy[9,11]. Despite this variability, meta-analyses consistently show that BCG offers strong and durable protection against severe and disseminated forms of TB in children, including miliary and meningeal TB, while offering variable protection against adult pulmonary disease[14,15]. Though the reason behind this variable efficacy of BCG remains elusive, our analysis reveals a key mechanistic insight of a hypothetical vaccine: control of Mtb replication strongly governs dissemination outcomes, suggesting that vaccines that suppress replication are likely to provide broader and more consistent protection against pulmonary spread than those targeting dissemination alone. Thus, our work provides a predictive and quantitative framework that bridges experimental and clinical observations, reinforcing the concept that the primary impact of BCG vaccination lies in its ability to limit systemic spread of Mtb.

Our modeling framework is subject to several limitations that should be considered when interpreting the results. We reformulated the problem of Mtb dissemination in ULD-infected mice as initial infection (with 1 CFU) in Lung 1 and Mtb dissemination to Lung 2; this resulted in a simpler modeling framework (e.g., **eqns. (1)–(2)**) but required re-arranging of the data into Lung 1 (larger CFU) and Lung 2 (smaller CFU, **Fig. S3**). While this approach provided a consistent framework for fitting models, assuming different pathways of Mtb dissemination, to data, it is a simplifying assumption that may not accurately capture all infection scenarios. Indeed, stochastic simulations

of the DD ($L_1 \rightarrow L_2$) model did not well describe CFU variability observed experimentally (**Fig. 4**). It is re-assuring, though, that realistic stochastic simulations of the DD (RL $\rightleftarrows$ LL) model with parameters estimated assuming unidirectional dissemination matched better variable CFU measured experimentally (**Fig. 5**) suggesting that our parameter estimates are reasonable. The alternative approach to model Mtb dynamics at ULD doses deterministically is to consider, for example, DD (RL $\rightleftarrows$ LL) model with the ensemble of initial conditions of infection starting in the left, right, or both lungs, and weighting the model solutions proportionally to the frequency of a given initial condition. Work in this direction is ongoing and preliminary results suggest similar conclusions as those reached with a simplified approach (**Suppl. Fig. S9**).

In our analyses we assumed that Mtb starts growing exponentially from the time it lands in the lung. However, our unpublished experimental data on CD infection of mice suggest that there may be a delay of few hours to few days before the bacteria enter the phase of exponential growth. We would expect that including a delay in Mtb growth may increase the estimated rate of Mtb replication during first days after infection but because details of Mtb activation in ULD-infected mice are unknown, we opted for a model that ignores such a delay in start of exponential growth phase.

In our analysis we excluded uninfected mice; such exclusion could potentially bias parameter estimates toward higher infection probabilities and obscure vaccine effects on infection establishment. However, it is important to emphasize that the detected reduction in the rate of Mtb replication due to BCG vaccination $\epsilon_r = 0.09$ translates to 9% increase in infection clearance probability ($p_0 = \delta/(r(1-\epsilon_r)) \approx (1+\epsilon_r)\delta/r$ when $\epsilon_r \ll 1$ and for dose $= 1$) as was observed experimentally[39]. The models developed in the paper do not include explicit changes in immune components such as macrophages, neutrophils, or antigen-specific T cell response. In part, this was driven by the lack of experimental data on dynamics of these cell population in ULD-exposed mice. Extending the models to consider more mechanistic details of Mtb control in ULD infection is a key area of our current research.

We found that both direct and indirect dissemination models fit the experimental data with somewhat similar quality (**Suppl. Table. S2**); yet, the DD model was the simplest and thus was favored by statistical metrics such as AIC. The ID dissemination model predicted slightly lower CFU values in an intermediate tissue as was observed experimentally in the spleen (**Suppl. Fig. S4**), highlighting the need to further improve the ID model. Additional data on Mtb burdens in other tissues, such as lung-draining lymph nodes, spleen, and potentially blood, should be able to narrow down specific pathways of Mtb dissemination in the lung. In our preliminary analyses, we found that it was not possible to determine changes in both replication and death rates when fitting ODE-based models to experimental data, so we assumed Mtb death rates estimated previously[43,45]. Because of this constraint, we also used the time periods in which Mtb replication and death rates were estimated previously that may have influenced inference of how Mtb replication rate may be changing during the infection. Our ongoing work suggests that changing the time intervals at which Mtb replication rate changes during the ULD infection may improve the mode fit of the data.

In fitting the models to data we rescaled all CFU values by one, so that any zero CFU data become similar to the limit of detection of 1 CFU. Future work will need to explore if alternative treatments of zero CFU data, e.g., by assuming low CFU values follow a distribution[59], result in different conclusions. Finally, our stochastic simulations have shown great variability in predicted lung CFU and yet, the ODE models we fit to data primarily utilizing average CFU values. Comparing

predictions of stochastic simulations to data using KS test suggests a way to compare stochastic simulations to data but it remains to be explored if KS test could be used to rigorously fit stochastic models to data[50].

Our work opens avenues for future research. There have been limited methods of comparing predictions of stochastic mathematical models to data. Our approach, based on calculating the average KS statistics (**eqn. (9)**), could be extended to minimize D with respect to model parameters in standard optimization routines[50]. Our work also highlights that impact of a vaccine could be determined relatively early after infection (**Suppl. Fig. S8**) and that time course CFU data may be more powerful at detecting efficacy of replication-blocking vaccines as compared to measuring proportion of infected mice (**Table. 3**). Whether these inferences are repeated in testing of other vaccines in ULD-infected mice remains to be determined.

A growing number of TB vaccine candidates require small animal models for vaccine testing. One such model, the ULD infection murine model, has already demonstrated ability to infer more protection-related parameters than has been done in CD-infected mice[39,41]. Our novel mathematical modeling framework by mimicking ULD infection of mice will likely help interpret future data obtained from TB vaccine testing using the ULD murine model. For example, by using CFU burden measured at a few time points, our model should be able to predict vaccine efficacy by comparing results of realistic stochastic simulations with the data. Combining mathematical models of Mtb dynamics at ULD infection with experimental data may be a powerful novel way to access efficacy of next generation TB vaccines.

Data and code sources

The R codes and reformatted data are available on github:
https://github.com/dipanjan754/BCG_vaccination_efficacy/.

Ethics statement

No animal work has been performed.

Author Contribution

VVG and DC conceived the overall concept of the study and developed alternative models. DC performed data analysis, fitting models to the data, and stochastic simulations in the ULD framework. VVG performed simulations to fit models to the data in the conventional dose framework. DC wrote the first draft of the paper and all authors read, edited, and agreed on the final version.

Acknowledgment

We would like to thank Plumlee *et al.*[39] for making their data available with the publication and members of Urdahl lab for discussions of ULD infection and TB vaccines. We also would like to thank all members of the Ganusov lab for their feedback on previous versions of the manuscript.

Financial Disclosure Statement

This work was supported in part by the NIH/NIAID grant R01AI158963 to VVG. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. World Health Organization 2024 Global tuberculosis report. https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2024/tb-disease-burden. Accessed: [29th January, 2026].

2. Walters, E., Duvenhage, J., Draper, H. R., Hesseling, A. C., Van Wyk, S. S., Cotton, M. F. & Rabie, H. 2014 Severe manifestations of extrapulmonary tuberculosis in HIV-infected children initiating antiretroviral therapy before 2 years of age. *Archives of disease in childhood*, **99**(11), 998–1003.

3. Leeds, I. L., Magee, M. J., Kurbatova, E. V., Del Rio, C., Blumberg, H. M., Leonard, M. K. & Kraft, C. S. 2012 Site of extrapulmonary tuberculosis is associated with HIV infection. *Clinical Infectious Diseases*, **55**(1), 75–81.

4. World Health Organization 2025 *WHO Consolidated Guidelines on Tuberculosis: Module 4: Treatment and Care*. Geneva: World Health Organization. Annex 1: Summary of recommendations.

5. Harris, R. C., Sumner, T., Knight, G. M. & White, R. h. G. 2016 Systematic review of mathematical models exploring the epidemiologica l impact of future TB vaccines. *Human vac immun*, **12**, 2813–2832. doi:10.1080/21645515.2016.1205769.

6. Luca, S. & Mihaescu, T. 2013 History of BCG vaccine. *Maedica*, **8**, 53–58.

7. Calmette, A. 1931 Preventive vaccination against tuberculosis with BCG.

8. Trostdorf Monteiro Filardi, E., CRUZ CARBONELL, R. C., Pavan, F. R., Cerni, F. A. & Pucca, M. B. 2025 One hundred years of BCG: The journey of tuberculosis vaccination in Brazil. *Frontiers in Immunology*, **16**, 1655 969.

9. Hart, P. & Sutherland, I. 1950 Preventive inoculation against tuberculosis in adolescence: a further study. *British Medical Journal*, **2**(4682), 485.

10. MRC 1959 B.C.G. and vole bacillus vaccines in the prevention of tuberculosis in adolescents. *British medical journal*, **2**, 379–396.

11. ICMR 1980 Trial of BCG vaccines in south india for tuberculosis prevention: first report. *Bulletin of the World Health Organization*, **58**(5), 731–745.

12. Liao, Q., Zheng, Y., Wang, Y., Ye, L., Liu, X., Jiao, W., Liu, Y., Zhu, Y., Jia, J. *et al.* 2022 Effectiveness of Bacillus Calmette-Guérin vaccination against severe childhood tuberculosis in China: a case-based, multicenter retrospective study. *International Journal of Infectious Diseases*, **121**, 113–119.

13. Trunz, B. B., Fine, P. & Dye, C. 2006 Effect of BCG vaccination on childhood tuberculous meningitis and miliary tuberculosis worldwide: a meta-analysis and assessment of cost-effectiveness. *The Lancet*, **367**(9517), 1173–1180.

14. Mangtani, P., Abubakar, I., Ariti, C., Beynon, R., Pimpin, L., Fine, P. E., Rodrigues, L. C., Smith, P. G., Lipman, M. *et al.* 2014 Protection by BCG vaccine against tuberculosis: a systematic review of randomized controlled trials. *Clinical Infectious Diseases*, **58**(4), 470–480.

15. Andersen, P. & Doherty, T. M. 2005 The success and failure of BCG—implications for a novel tuberculosis vaccine. *Nature Reviews Microbiology*, **3**(8), 656–662.

16. Kumar, P. 2021 A perspective on the success and failure of BCG. *Frontiers in Immunology*, **12**, 778 028.

17. Moliva, J. I., Turner, J. & Torrelles, J. B. 2015 Prospects in Mycobacterium bovis Bacille Calmette et Guérin (BCG) vaccine diversity and delivery: why does BCG fail to protect against tuberculosis? *Vaccine*, **33**(39), 5035–5041.

18. Foster, M., Hill, P. C., Setiabudiawan, T. P. m., Koeken, V. A. C. M., Alisjahbana, B. & van Crevel, R. 2021 BCG-induced protection against Mycobacterium tuberculosis infecti on: Evidence, mechanisms, and implications for next-generation vaccines. *Immunological reviews*, **301**, 122–144. doi:10.1111/imr.12965.

19. Moliva, J. I., Turner, J. & Torrelles, J. B. 2017 Immune responses to Bacillus Calmette-Guérin vaccination: Why do they fail to protect against Mycobacterium tuberculosis ? *Frontiers in immunology*, **8**, 407. doi:10.3389/fimmu.2017.00407.

20. Gong, W., Liang, Y. & Wu, X. 2020 Animal models of tuberculosis vaccine research: an important component in the fight against tuberculosis. *BioMed Research International*, **2020**(1), 4263 079.

21. McMurray, D. N. 1994 Guinea pig model of tuberculosis. *Tuberculosis: pathogenesis, protection, and control*, pp. 135–147.

22. Tsenova, L., Harbacheuski, R., Sung, N., Ellison, E., Fallows, D. & Kaplan, G. 2007 BCG vaccination confers poor protection against M. tuberculosis HN878-induced central nervous system disease. *Vaccine*, **25**(28), 5126–5132.

23. Singhal, A., Mathys, V., Kiass, M., Creusy, C., Delaire, B., Aliouat, E. M., Dartois, V., Kaplan, G. & Bifani, P. 2011 BCG induces protection against Mycobacterium tuberculosis infection in the Wistar rat model. *PLoS One*, **6**(12), e28 082.

24. Zhang, L., Ru, H.-w., Chen, F.-z., Jin, C.-y., Sun, R.-f., Fan, X.-y., Guo, M., Mai, J.-t., Xu, W.-x. *et al.* 2016 Variable virulence and efficacy of BCG vaccine strains in mice and correlation with genome polymorphisms. *Molecular Therapy*, **24**(2), 398–405.

25. Henao-Tamayo, M., Obregón-Henao, A., Creissen, E., Shanley, C., Orme, I. & Ordway, D. J. 2015 Differential mycobacterium bovis BCG vaccine-derived efficacy in C3Heb/FeJ and C3H/HeOuJ mice exposed to a clinical strain of Mycobacterium tuberculosis. *Clinical and Vaccine Immunology*, **22**(1), 91–98.

26. Chen, H., Liu, X., Ma, X., Wang, Q., Yang, G., Niu, H., Li, S., He, B., He, S. *et al.* 2017 A new rabbit-skin model to evaluate protective efficacy of tuberculosis vaccines. *Frontiers in Microbiology*, **8**, 842.

27. Dijkman, K., Sombroek, C. C., Vervenne, R. A., Hofman, S. O., Boot, C., Remarque, E. J., Kocken, C. H., Ottenhoff, T. H., Kondova, I. *et al.* 2019 Prevention of tuberculosis infection and disease by local BCG in repeatedly exposed rhesus macaques. *Nature medicine*, **25**(2), 255–262.

28. Wei, J., Li, J., Li, X., Du, W., Su, C., Sheng, X., Huang, Y., Wang, J., Niu, Q. *et al.* 2025 The impact of animal models and strain standardization on the evaluation of tuberculosis vaccine efficacy. *Vaccines*, **13**(7), 669.

29. Clark, S., Hall, Y. & Williams, A. 2015 Animal models of tuberculosis: Guinea pigs. *Cold Spring Harbor perspectives in medicine*, **5**(5), a018 572.

30. del Pozo-Ramos, L. & Kupz, A. 2025 A review of the efficacy of clinical tuberculosis vaccine candidates in mouse models. *Frontiers in Immunology*, **16**, 1609 136.

31. Ouh, I.-O., Kim, M. J., Kim, K., Lim, H., Yang, Y. J., Heo, J. W., Choi, H. N., Kim, H. H., Lee, H.-J. *et al.* 2025 Bacillus Calmette–Guérin vaccination promotes efficient and comprehensive immune modulation in guinea pig models. *Vaccines*, **13**(3), 305.

32. Darrah, P. A., Zeppa, J. J., Maiello, P., Hackney, J. A., Wadsworth, M. H., Hughes, T. K., Pokkali, S., Swanson, P. A., Grant, N. L. *et al.* 2020 Prevention of tuberculosis in macaques after intravenous BCG immunization. *Nature*, **577**(7788), 95–102.

33. Simonson, A. W., Zeppa, J. J., Bucsan, A. N., C hao, M. C., Pokkali, S., Hopkins, F., Chase, M. R., Vicker s, A. J., Sutton, M. S. *et al.* 2025 Intravenous BCG-mediated protection against tuberculosis requires CD4+ T cells and CD8a+ lymphocytes. *The Journal of experimental medicine*, **222**. doi:10.1084/jem.20241571.

34. Kaushal, D., Mehra, S., Didier, P. & Lackner, A. 2012 The non-human primate model of tuberculosis. *Journal of medical primatology*, **41**(3), 191–201.

35. Reed, S. G., Coler, R. N., Dalemans, W., Tan, E. V., DeLa Cruz, E. C., Basaraba, R. J., Orme, I. M., Skeiky, Y. A., Alderson, M. R. *et al.* 2009 Defined tuberculosis vaccine, Mtb72F/AS02A, evidence of protection in cynomolgus monkeys. *Proceedings of the National Academy of Sciences*, **106**(7), 2301–2306.

36. Peña, J. C. & Ho, W.-Z. 2017 Non-human primate models of tuberculosis. *Tuberculosis and the Tubercle Bacillus*, pp. 163–176.

37. Singh, S., Saavedra-Avila, N. A., Tiwari, S. & Porcelli, S. A. 2022 A century of BCG vaccination: Immune mechanisms, animal models, non-traditional routes and implications for COVID-19. *Frontiers in immunology*, **13**, 959 656. doi:10.3389/fimmu.2022.959656.

38. Irvine, E. B., Darrah, P. A., Wang, S., Wang, C., McNamara, R. P., Roederer, M., Seder, R. A., Lauffenburger, D. A., Flynn, J. L. *et al.* 2024 Humoral correlates of protection against Mycobacterium tuberculosis following intravenous BCG vaccination in rhesus macaques. *iScience*, **27**, 111 128. doi:10.1016/j.isci.2024.111128.

39. Plumlee, C. R., Barrett, H. W., Shao, D. E., Lien, K. A., Cross, L. M., Cohen, S. B., Edlefsen, P. T. & Urdahl, K. B. 2023 Assessing vaccine-mediated protection in an ultra-low dose Mycobacterium tuberculosis murine model. *Plos Pathogens*, **19**(11), e1011 825.

40. Lindenstrøm, T., Knudsen, N. P. H., Agger, E. M. & Andersen, P. 2013 Control of chronic Mycobacterium tuberculosis infection by CD4 KLRG1- IL-2–secreting central memory cells. *The Journal of Immunology*, **190**(12), 6311–6319.

41. Plumlee, C. R., Duffy, F. J., Gern, B. H., Delahaye, J. L., Cohen, S. B., Stoltzfus, C. R., Rustad, T. R., Hansen, S. G., Axthelm, M. K. *et al.* 2021 Ultra-low dose aerosol infection of mice with Mycobacterium tuberculosis more closely models human tuberculosis. *Cell host & microbe*, **29**(1), 68–82.

42. Ganusov, V. V. 2016 Strong Inference in mathematical modeling: A method for robust science in the twenty-first century. *Front Microbiol*, **7**, 1131. doi:10.3389/fmicb.2016.01131.

43. McDaniel, M. M., Krishna, N., Handagama, W. G., Eda, S. & Ganusov, V. V. 2016 Quantifying limits on replication, death, and quiescence of Mycobacterium tuberculosis in mice. *Frontiers in Microbiology*, **7**, 862.

44. Friesen, A. & Ganusov, V. V. 2025 Immortal genome assumption significantly underestimates replication and death rates of Mycobacterium tuberculosis in mice and monkeys. *bioRxiv*, pp. 2025–08.

45. Gill, W. P., Harik, N. S., Whiddon, M. R., Liao, R. P., Mittler, J. E. & Sherman, D. R. 2009 A replication clock for mycobacterium tuberculosis. *Nature medicine*, **15**(2), 211–214.

46. Miller, J., Burch-Smith, T. M. & Ganusov, V. V. 2022 Mathematical modeling suggests cooperation of plant-infecting viruses. *Viruses*, **14**(4). doi:10.3390/v14040741.

47. Burnham, K. P. & Anderson, D. R. 2002 *Model selection and multimodel inference: a practical information-theoret ic approach.* Springer-Verlag, New York.

48. Cannoodt, R. & Saelens, W. 2023 *GillespieSSA2: Gillespie's Stochastic Simulation Algorithm for Impatient People*. doi:10.32614/CRAN.package.GillespieSSA2. R package.

49. Gillespie, D. T. 1977 Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, **81**(25), 2340–2361.

50. Chakraborty, D., Batabyal, S. & Ganusov, V. V. 2024 A brief overview of mathematical modeling of the within-host dynamics of Mycobacterium tuberculosis. *Frontiers in applied mathematics and statistics*, **10**, 1355 373.

51. Chackerian, A. A., Alt, J. M., Perera, T. V., Dascher, C. C. & Behar, S. M. 2002 Dissemination of Mycobacterium tuberculosis is influenced by host factors and precedes the initiation of T-cell immunity. *Infect Immunity*, **70**, 4501–4509. doi:10.1128/iai.70.8.4501-4509.2002.

52. Krishnan, N., Robertson, B. D. & Thwaites, G. 2010 The mechanisms and consequences of the extra-pulmonary dissemination of Mycobacterium tuberculosis. *Tuberculosis*, **90**, 361–366. doi:10.1016/j.tube.2010.08.005.

53. Moule, M. G. & Cirillo, J. D. 2020 Mycobacterium tuberculosis dissemination plays a critical role in pathogenesis. *Front Cell Infec Microbiol*, **10**, 65. doi:10.3389/fcimb.2020.00065.

54. Trauer, J. M., Kawai, A., Coussens, A. K., Datta, M., Williams, B. M., McBryde, E. S. & Ragonnet, R. 2021 Timing of Mycobacterium tuberculosis exposure explains variation in BCG effectiveness: a systematic review and meta-analysis. *Thorax*, **76**, 1131–1141. doi:10.1136/thoraxjnl-2020-216794.

55. Rajakaruna, H. & Ganusov, V. V. 2022 Mathematical modeling to guide experimental design: T cell clustering as a case study. *Bulletin of mathematical biology*, **84**, 103–115. doi:10.1007/s11538-022-01063-x.

56. Majumder, B., Budhu, S. & Ganusov, V. V. 2023 Cytotoxic T lymphocytes control growth of B16 tumor cells in collagin-fibrin gels by cytolytic and non-lytic mechanisms. *Viruses*, **15**, 1454. doi:10.3390/v15071454.

57. Friesen, A. D. & Ganusov, V. V. 2025 Immortal genome assumption significantly underestimates replication and death rates of Mycobacterium tuberculosis in mice and monkeys. *bioRxiv*. doi:10.1101/2025.08.15.670581. Microbe Spectrum: in review.

58. Fine, P. 1988 BCG vaccination against tuberculosis and leprosy. *British medical bulletin*, **44**(3), 691–703.

59. Thiebaut, R., Guedj, J., Jacqmin-Gadda, H., Chene, G., Trimoulet, P., Neau, D. & Commenges, D. 2006 Estimation of dynamical model parameters taking into account undetectable marker values. *BMC Med Res Methodol*, **6**, 38.

60. Shamskhou, E. A., Duffy, F. J., Cross, L. M., Plumlee, C., Gern, B. H., Diercks, A. H., Cohen, S. B., Aitchison, J. D., Ganusov, V. V. *et al.* 2025 Monocytic niches enable mycobacterium tuberculosis persistence in lymph nodes. *Available at SSRN 5588229*.

61. Le, D., Miller, J. D. & Ganusov, V. V. 2015 Mathematical modeling provides kinetic details of the human immune response to vaccination. *Front Cell Infect Microbiol*, **4**, 177. doi:10.3389/fcimb.2014.00177.

62. Riley, R. L. 1957 The j. burns amberson lecture: aerial dissemination of pulmonary tuberculosis. *American Review of Tuberculosis and Pulmonary Diseases*, **76**(6), 931–941.

63. Delahaye, J. L., Gern, B. H., Cohen, S. B., Plumlee, C. R., Shafiani, S., Gerner, M. Y. & Urdahl, K. B. 2019 Cutting edge: Bacillus Calmette-Guérin-induced T cells shape Mycobacterium tuberculosis infection before reducing the bacterial burden. *J Immunol*, **203**, 807–812. doi:10.4049/jimmunol.1900108.

BCG vaccination reduces the rate of Mycobacterium tuberculosis dissemination between murine lungs

Dipanjan Chakraborty and Vitaly V. Ganusov

Supplemental Information

ID model assuming intermediate tissue is blood

In the ID model, if blood is treated as an intermediate tissue, Mtb neither replicates nor dies in this compartment. Instead, blood will serve solely as a transport medium, carrying Mtb from Lung 1 to Lung 2. Thus, the ID model equations will be modified as mentioned below:

$$\frac{dL_1(t)}{dt} = (1-\epsilon_r)r_L(t)L_1(t) - \frac{\delta(t)}{1-\epsilon_\delta}L_1(t) - (1-\epsilon_m)m_1L_1(t), \tag{S.1}$$

$$\frac{dT(t)}{dt} = (1-\epsilon_{m_1})m_1L_1(t) - m_2(1-\epsilon_m)T(t), \tag{S.2}$$

$$\frac{dL_2(t)}{dt} = \theta(L_2-1)\Big[(1-\epsilon_r)r_L(t)L_2(t) - \frac{\delta(t)}{1-\epsilon_\delta}L_2(t)\Big] + m_2(1-\epsilon_m)T(t), \tag{S.3}$$

Modeling Mtb dynamics in the whole lung

To evaluate how BCG vaccination impacts early Mtb dynamics in mice infected with an ULD or conventional dose (**CD**) of Mtb, we modified our previously proposed mathematical model[60] to include dynamics of immune response E and suppression of Mtb replication by immunity[61]:

$$\frac{\mathrm{d}B}{\mathrm{d}t} = \frac{(r+\delta)B}{1+h^nE(t)^n} - \delta B, \tag{S.4}$$

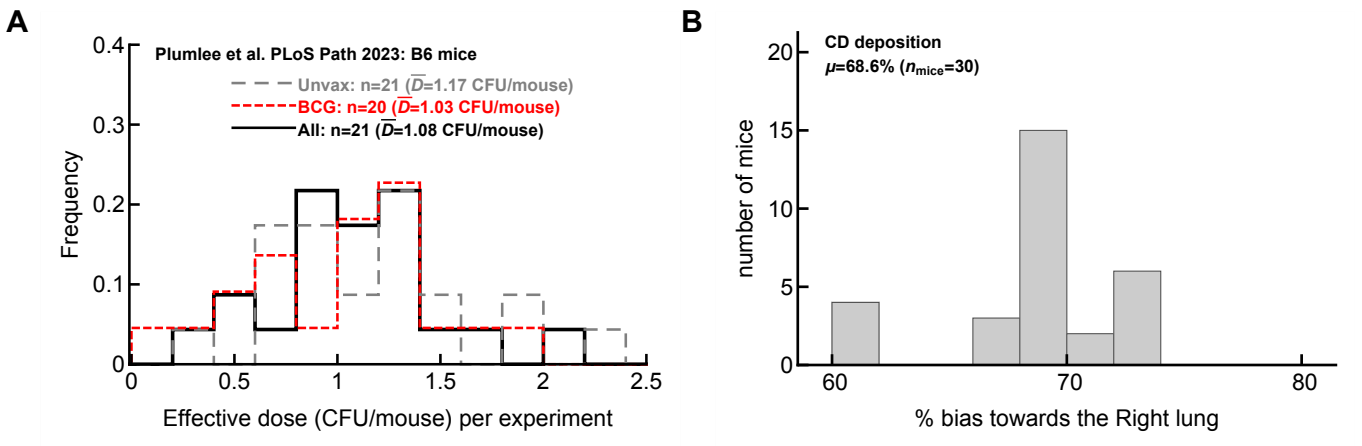
where B is the number of bacteria in the lung, r and δ are the rates of Mtb replication and death, respectively, h is the level of immunity at which immune-mediated suppression is half of its maximum, and $E(t) = \mathrm{e}^{\rho t}$ where ρ is the expansion of the immune response. We characterize kinetics of the immune response by time $T_E = -\ln(h)/\rho$ that is the time when immunity reaches 50% of its suppressive capacity. We estimated $\rho \approx 0.4-0.6$/day for Mtb-specific CD4 T cell response (Batabyal et al. (in prep)) but specific value of ρ is typically not very critical for the model behavior as long as it is sufficiently large[61]. In the model, infection starts with B_0 bacteria.

Calculation of effective dose in ULD-infected mice

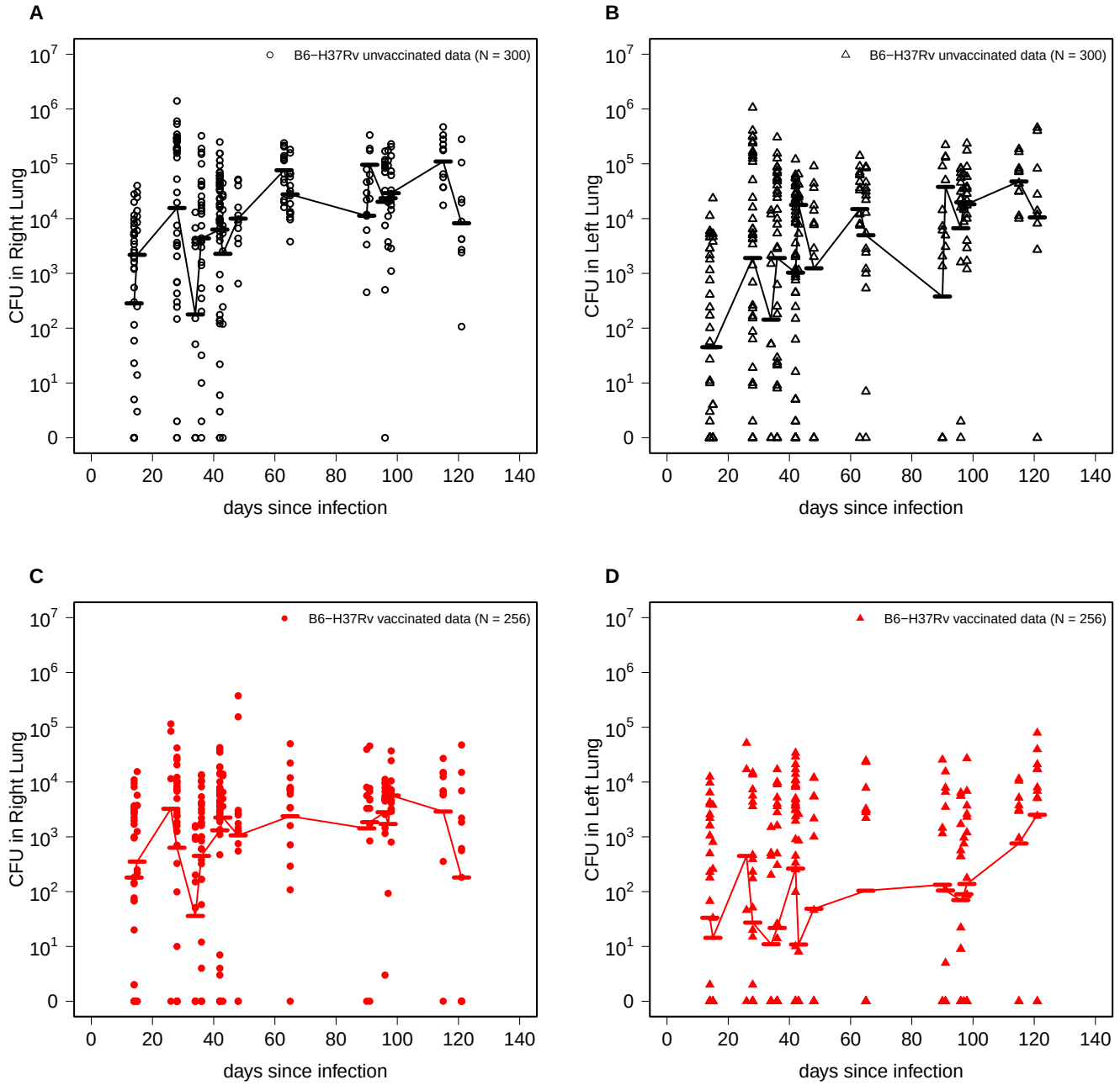
Let f_0 be the fraction of mice that remain uninfected, and D be the average dose of Mtb infection. Since Mtb infection is random, the number of bacteria inhaled by individual mice will follow a Poisson distribution[41,62]. Then, the probability that a mouse receives k bacteria is,

$$P(k) = \frac{D^k e(-D)}{k!}, \tag{S.5}$$

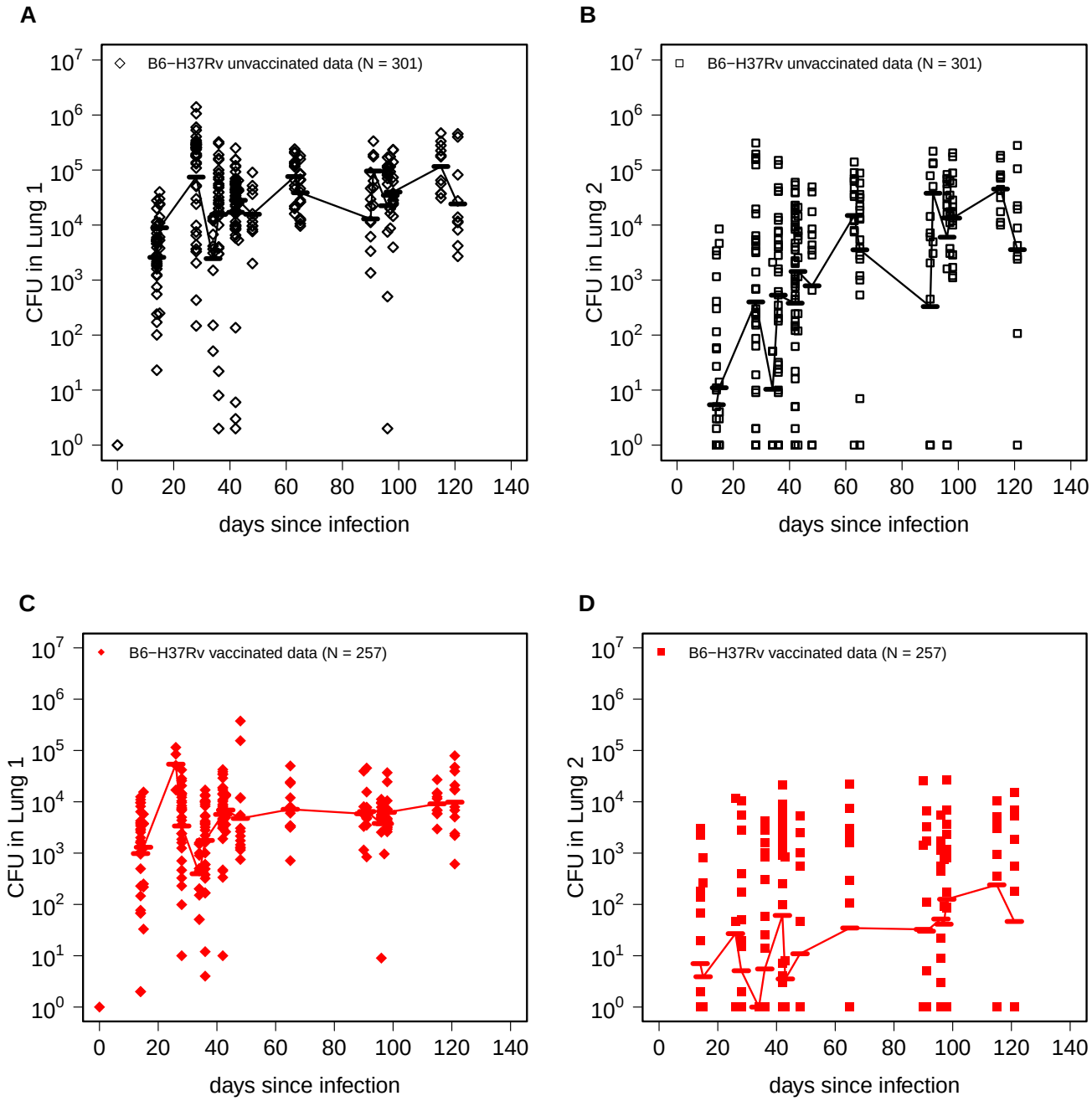
where the probability of no infection is: $P(0) = e^{-D} = f_0$. Thus, the effective dose $D = -\ln(f_0)$. Note that the actual dose that mice are exposed to may be higher because of additional death rate of Mtb during the infection (**eqn. (3)**).



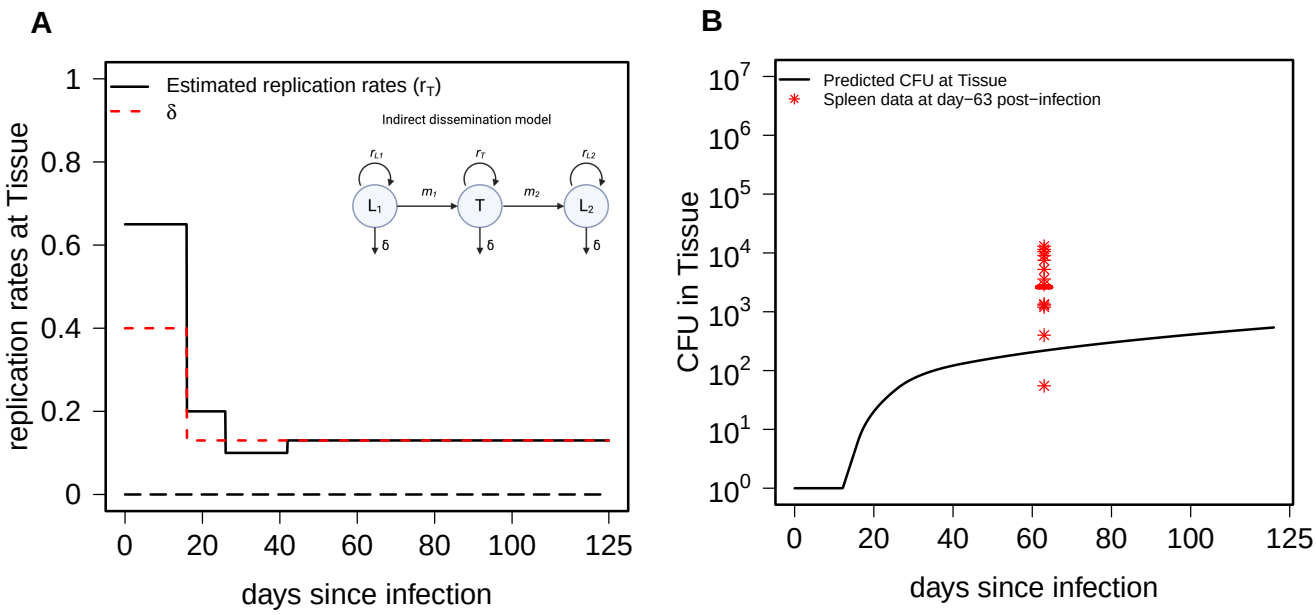
Supplemental Figure S1: Distribution of estimated dose in mice exposed to ultra-low dose of Mtb and bias of initial infection towards the right lung of mice. **A**: For every experiment of $n = 21$ experiments of Plumlee *et al.*[39] we calculated the proportion of uninfected mice for unvaccinated group, BCG-vaccinated group, or all mice in the experiment (f_0) and then calculated the average effective deposition dose using a Poisson distribution $D = -\ln(f_0)$ (ref[41]). In cases when all mice were infected, the dose could not be calculated, and such an experiment was excluded from the calculations. (Note a different number of experiments n used for calculation of the dose.) **B**: We calculated the bias in the CFUs recovered from right ($\mathrm{CFU_{RL}}$) or left ($\mathrm{CFU_{LL}}$) lung 1 day after Mtb infection of B6 mice with a conventional dose[41]; the bias was calculated for each mouse as $b = \mathrm{CFU_{RL}}/(\mathrm{CFU_{RL}} + \mathrm{CFU_{LL}})$.



Supplemental Figure S2: Bacterial burden at right lung and left lung of ultra-low dose infected mice Plumlee *et al.*[39]. Mice were exposed to an ultra-low dose of Mtb H37Rv and bacterial burden at right lung (RL) and left lung (LL) was measured at different days post-infection after sacrificing the mice. Plumlee *et al.*[39] conducted multiple experiments, categorizing the mice into two cohorts: unvaccinated and BCG-vaccinated, comprising a total of (300 + 256) mice (excluding entries marked as 'NA' and the mice with no detectable infection i.e., mice with zero CFUs in both lungs or mice in which CFU was only measured in the whole lung without subdivision into right and left lungs). (A) and (B) present the CFU counts in the RL and LL of unvaccinated mice across all experiments, while (C) and (D) show the corresponding CFU data for BCG-vaccinated mice.



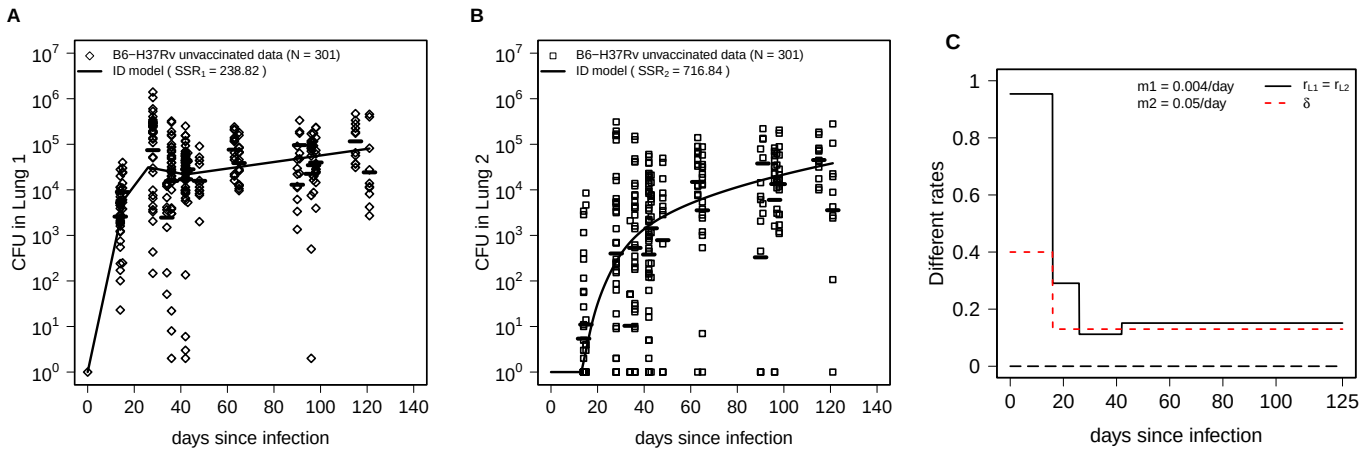
Supplemental Figure S3: Reclassification of right lung and left Lung to Lung 1 and Lung 2. We have redefined the CFU data of right lung and left lung to Lung 1 and Lung 2 such that Lung 1 always has higher CFU than Lung 2. This reclassification reflects our assumption that the initial infection starts at Lung 1 (Lung 1 = 1 CFU at day 0), and then it could go to Lung 2 or other tissues as prescribed in the mathematical model section. Thus, due to the addition of an initial CFU data point, the number of mice in the dataset used for our analyses increases to 301 for unvaccinated mice (A, B) and to 257 for vaccinated mice (C, D).



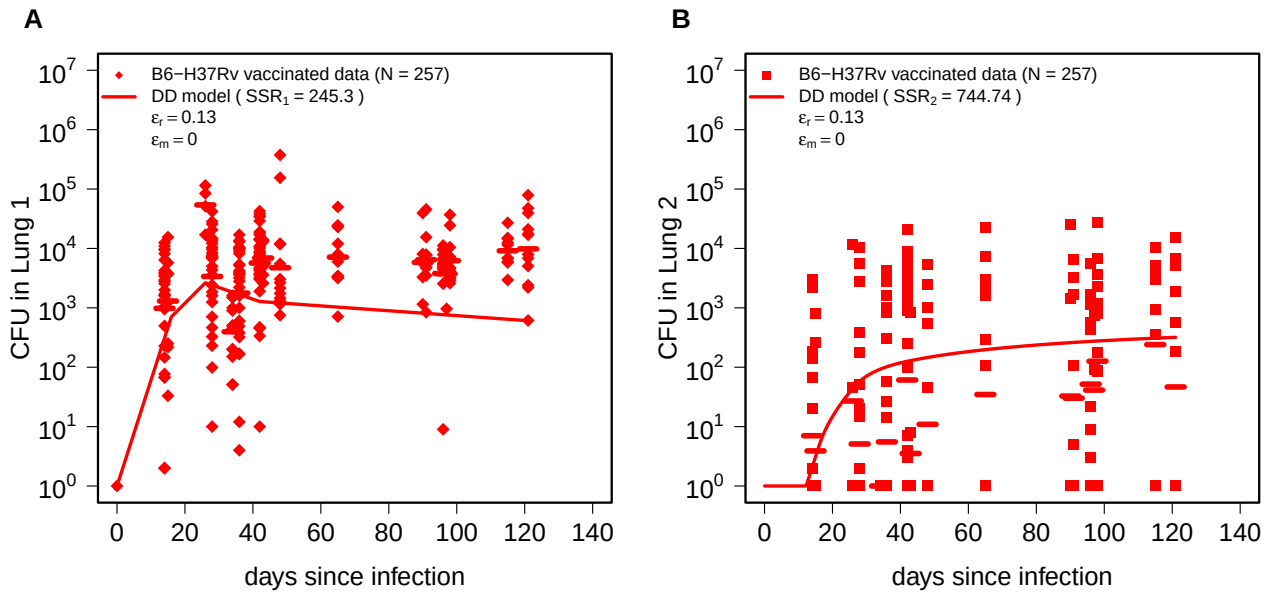
Supplemental Figure S4: Estimated replication rates and predicted bacterial burden at an intermediate tissue predicted by the best fit of the ID model. **A**: We plot time-varying replication rates of Mtb in the intermediate tissue as estimated from fitting the indirect dissemination model to the data. **B**: We plot prediction of the model on the bacterial numbers in an intermediate tissue (line) and experimentally measured CFU in the spleen of ULD-infected mice (markers).

Model	r_1	r_2	r_3	r_4	r_{T_1}	r_{T_2}	r_{T_3}	r_{T_4}	m	m_1	m_2
DD model	0.94 (0.89–0.98)	0.32 (0.28–0.33)	0.10 (0.09–0.11)	0.14 (0.13–0.16)	-	-	-	-	0.0027	-	-
ID model	0.97	0.23	0.11	0.15	0.65	0.20	0.10	0.13	-	0.004	0.042

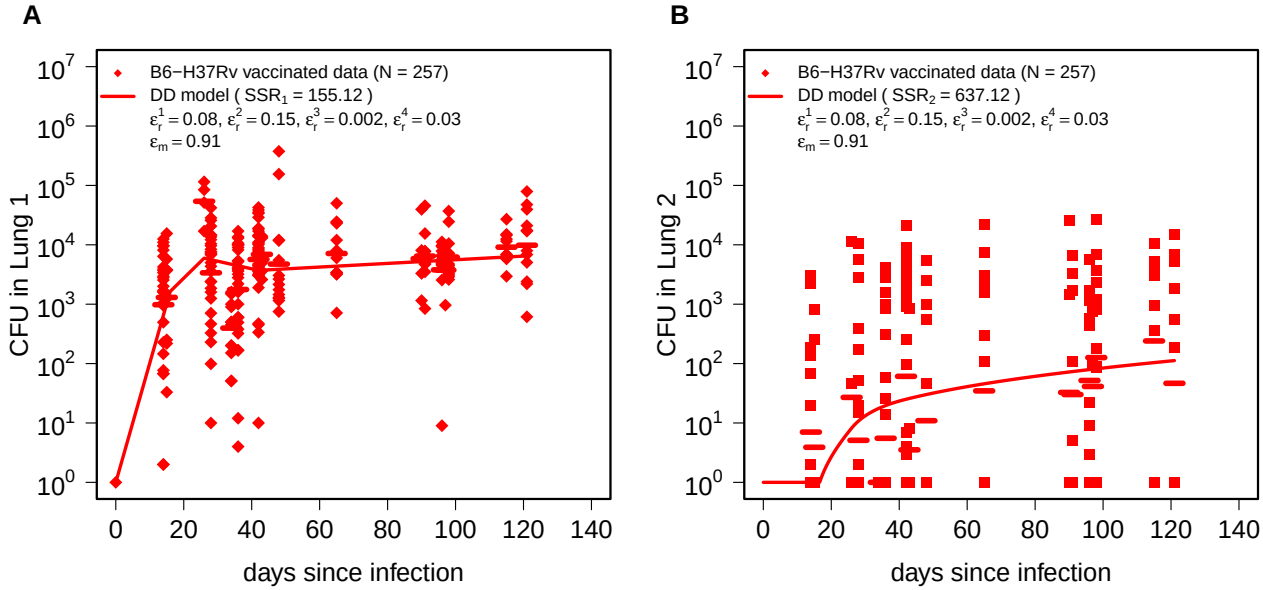
Supplemental Table S1: Estimates of replication rates of Mtb in the lung (r_i) or lung (r_i) and intermediate tissue (r_{T_i}), providing the best fit of the DD or ID model, respectively (model fits are shown in **Fig. 2**). The rates are time-dependent as defined in **eqn. (4)**. All rates are given in per day units, and − denotes not applicable. The 95% confidence intervals (CIs) for the parameter estimates were generated using `modFit`. We only report CIs for the DD model because `modFit` failed to provide CI estimates for the ID model due to the larger number of parameters.



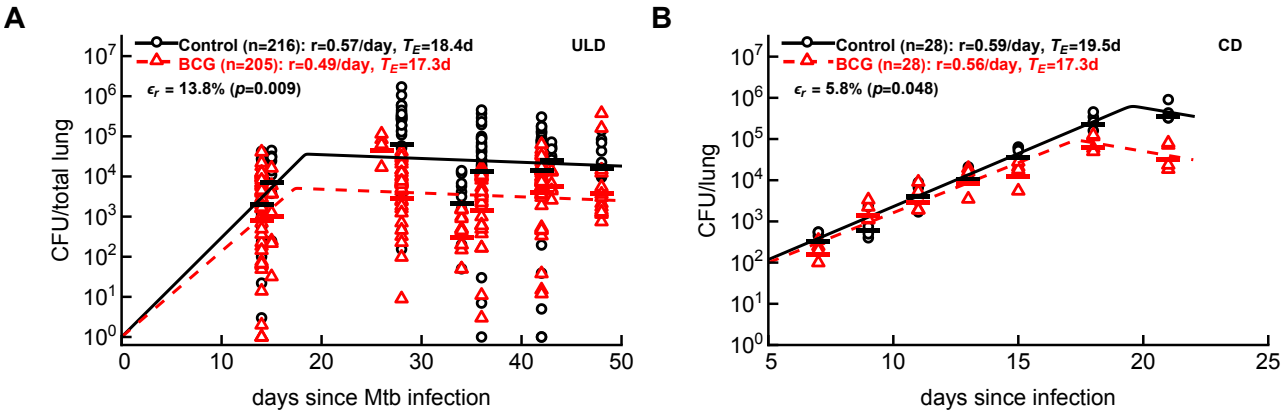
Supplemental Figure S5: Indirect dissemination model explains the data well even if the intermediate tissue is assumed to be blood. We assume that in the ID model the intermediate tissue is blood, and there is no replication and death in the this intermediate tissue. Then, we fit the model to the data and the best-fitted model predicts the CFU burden at Lung 1 and Lung 2 (A, B). The best-fitted model estimate of the dissemination rate from blood to Lung 2 was $m = 0.05$/day, and the time-dependent Mtb replication rate in the lung behaves similarly to that in **Fig. 2**C.



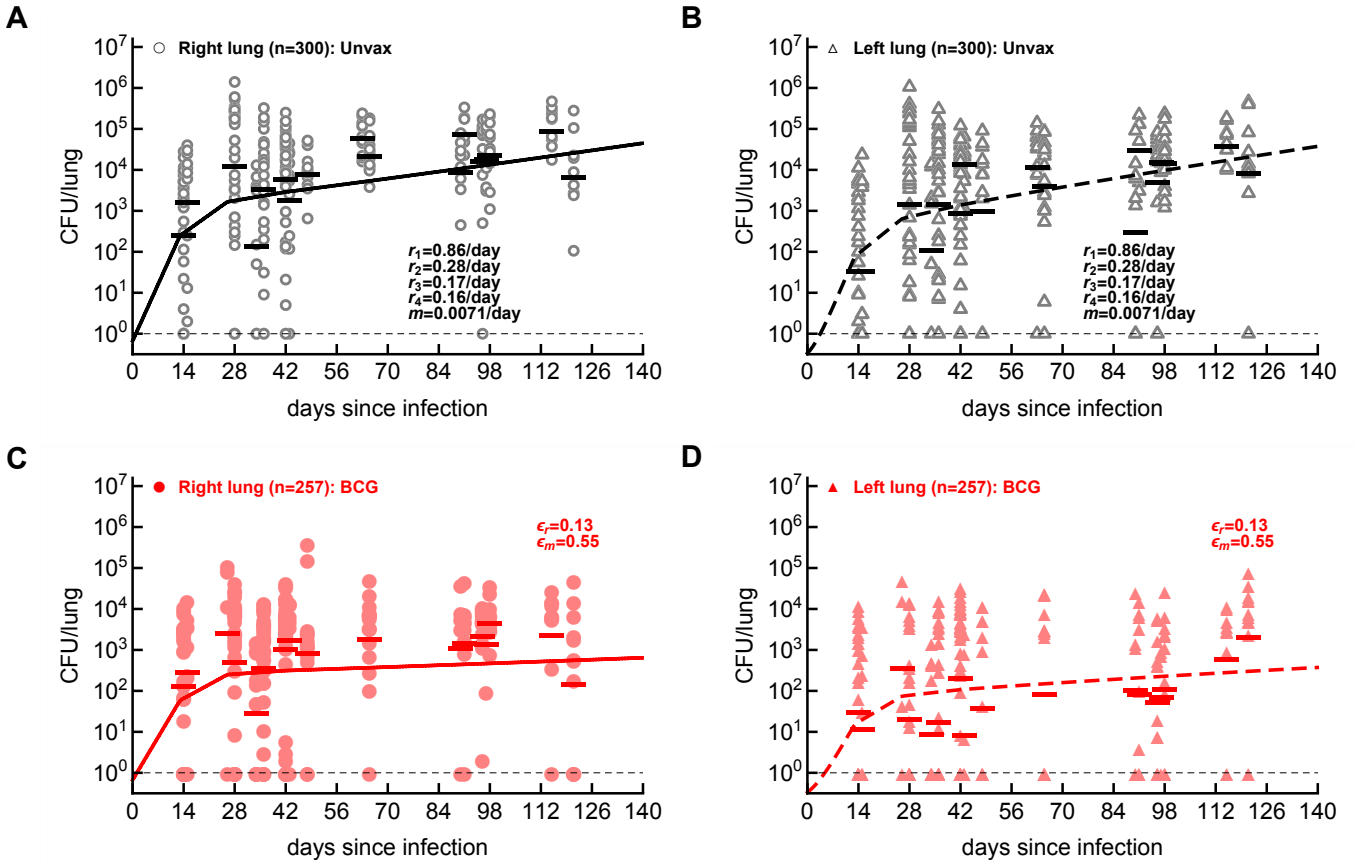
Supplemental Figure S6: The DD model assuming that BCG only reduces the rate of Mtb replication does not accurately fit the data. We fit the DD model (**eqns. (1)–(2)**) to the BCG-vaccinated data by assuming that BCG vaccination only impacts the Mtb replication rate (ϵ_r) but not the Mtb death rate and dissemination rate ($\epsilon_\delta = \epsilon_m = 0$); estimated $\epsilon_r = 0.13$. Other notations are similar to those in **Fig. 3**A-B. SSR values of the fit are shown on individual panels and AIC = 338.95 (see also **Suppl. Table. S2**).



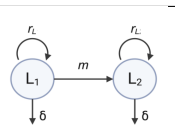
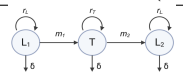
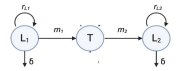
Supplemental Figure S7: Efficacy of BCG vaccine declines with time since infection. We fit the DD model (**eqns. (1)–(2)**) to the data from BCG-vaccinated mice by assuming that BCG vaccine efficacy in reducing the rate of Mtb replication ϵ_r depends on time since infection in time intervals defined in **eqn. (3)**). The fit had much lower SSR (see **Suppl. Table. S2**) and slightly lower AIC = 232.5.



Supplemental Figure S8: BCG vaccination reduces net Mtb replication rate early during ULD and CD infection of B6 mice. We fit our novel model that predicts Mtb dynamics in the presence of exponentially expanding immune response (**eqn. (S.4)**) to the data on Mtb numbers in lungs of B6 mice infected with an ultra-low dose (**A**, **Suppl. Fig. S2**) or conventional dose (**B**, ref[63]) of H37Rv strain of mice. In panel A we only take the data the first 50 days of the total lung CFU of ULD-infected mice (**Suppl. Fig. S2**). We fit the model to data from both unvaccinated or BCG-vaccinated mice simultaneously only allowing for the parameters r and h (or $T_E = -\ln(h)/\rho$) to vary between the datasets; fixing the rate of Mtb replication between the datasets resulted in fits of poorer quality (A: $F_{1,416} = 6.84$, $p = 0.009$; B: $F_{1,49} = 4.1$, $p = 0.048$). Estimated efficacy of BCG vaccination was calculated as $\epsilon_r = (r_{unvax} - r_{BCG})/r_{unvax}$. Other estimated parameters are as follows. Panel A: $B_0 = 1$ (assumed), $\delta = 0.22$/day, $h_{unvax} = 1.62 \times 10^{-5}$, $h_{BCG} = 3.02 \times 10^{-5}$, $r_{unvax} = 0.57$/day, $r_{BCG} = 0.49$/day; panel B: $B_0 = 6.12$, $\delta = 0.24$/day, $h_{unvax} = 8.22 \times 10^{-6}$, $h_{BCG} = 3.1 \times 10^{-5}$, $r_{unvax} = 0.59$/day, $r_{BCG} = 0.56$/day. Parameters $\rho = 0.6$/day and $n = 20$ (**eqn. (S.4)**) were fixed in model fits to data.



Supplemental Figure S9: Deterministic version of realistic simulations of the DD model also predicts that BCG vaccination reduces Mtb replication and dissemination rates. We fit the DD model in which we consider random deposition of inhaled Mtb bacilli into right or left lungs, growth in the lungs, and dissemination between the lungs to the data from unvaccinated (panels **A-B**) or BCG-vaccinated (**C-D**) mice together assuming that BCG vaccine reduces the rate of Mtb replication (by ϵ_r) and the rate of Mtb dissemination (by ϵ_m). We plot experimental data (by markers) and model predictions (by lines); horizontal dashed line denotes the limit of detection of 1 CFU. Best fit parameters are shown on the panels and are $r_1 = 0.86$/day, $r_2 = 0.28$/day, $r_3 = 0.17$/day, $r_4 = 0.16$/day, $m = 0.0071$/day, $\epsilon_r = 0.13$, and $\epsilon_m = 0.55$; the death rates are assumed as in **eqn. (3)**. The best model fit resulted in SSR = 3494.14. To generate model predictions and fit the model to data, we solved the DD model (**eqns. (1)–(2)**) for a collection of initial conditions with an infection initiated by n bacteria ($n = 1, 2 \ldots n_{\max}$) that distributed between right and left lung (e.g., $(1,0)$, $(0,1)$, $(2,0)$, $(1,1)$, $(0,2) \ldots (i,j)$ for (RL, LL), respectively). For each of these initial conditions we calculated the probability $p_{(i,j)}$ of such initial condition (i,j) happening using 0-truncated Poisson distribution and 2:1 bias in infection of the right:left lung. After solving the model for all initial conditions (e.g., assuming $n_{\max} = 3$) we generated the average model predictions by weighting the solutions with a given initial conditions by their respective probabilities. To calculate the SSR, we use paired model predictions of the CFU number in the right and left lung and those observed experimentally; the residuals for a model prediction with initial condition (i,j) were squared and weighed by the probability $p_{(i,j)}$ and summed over to generate SSR.

Fitting Models to:	Type of model	SSR	AIC	Akaike weights
Unvaccinated	 Direct dissemination (DD) model	939.1	**277.79**	**0.91**
	 Indirect dissemination (ID) model	930.52	282.53	0.08
	 ID model (blood as a carrier)	955	290.35	0.00
BCG vaccinated	DD model fitted with ϵ_r	990.04	338.95	0.0
	DD model with ϵ_δ	1081.22	384.23	0.0
	DD model with ϵ_r, ϵ_m	804.66	**234.4**	**0.82**
	DD model with $\epsilon_\delta, \epsilon_m$	835.91	244.58	0.005
	DD model with $\epsilon_r, \epsilon_m, \epsilon_\delta$	805.25	237.41	0.18
	DD model with varying ϵ_r,	792.24	232.5	Not counting

Supplemental Table S2: Comparison of model fits (SSR and AIC) for unvaccinated and BCG-vaccinated data. We show statistics of fits of alternative direct and indirect dissemination models, fitted to data from unvaccinated or BCG-vaccinated mice infected with ULD of H37Rv.

Category	**Efficacy Value**	**Model predicted sample size per group**
Vaccination efficacy on replication (ϵ_r)	0.2	500
	0.3	177
	0.4	69
	0.5	36
	0.6	<20
	0.7	<20
	0.8	<20
	0.9	<20
Vaccination efficacy on dissemination (ϵ_m)	0.85	165
	0.90	95
	0.95	48
	0.99	22

Supplemental Table S3: Calculation of number of mice needed to detect the efficacy of a hypothetical vaccine if the initial dose is $\lambda = 2$. We compared the sample size estimates required to achieve 80% power for different levels of vaccine efficacy, controlling replication (ϵ_r) or dissemination (ϵ_m). The power analysis is implemented by using realistic stochastic simulation to predict the sample size per group, as described in the Materials and methods section (and see **Table. 3**).