

A spatial model of the efficiency of T cell search in the influenza-infected lung (Response to Reviewers)

Drew Levin, Stephanie Forrest, Soumya Banerjee, Candice Clay, Judy Cannon, Melanie Moses, Frederick Koster

We thank the reviewers for their constructive and insightful comments. We have done our best to address each point raised and feel the incorporation of the reviewers' comments has greatly improved the paper. We include detailed responses to each comment below.

Reviewer 1

1. To me, the most interesting result would be estimating parameters for the dynamics of the virus and production of chemokines but the authors provide little to no details of how this was done. This is not sufficient as the model has so many parameters I would like to see how you estimated those. What are the model fits? Which parameters were fixed in fitting and why? What if Mitchell et al. study has parameters which are not compatible with your experiments? you need to explore this much deeper than the current ms. does.

We added one new paragraph and modified another in the Models for Parameter Estimation subsection to address the reviewer's three main questions, as well as others the reviewer did not explicitly ask for. In summary:

- The model was fit using a genetic algorithm to minimize the sum of the squared log residuals between the model and the data. In an earlier analysis the GA fits were superior to non-linear regression methods in both the Berkeley Madonna and Matlab software suites.
- All parameters were held constant during the fitting procedure aside from the chemokine secretion rate.
- The model was fit to experimental data measured during the experiments reported in Mitchell 2011. However, the chemokine data (Fig. 3, Table S1) have not been published previously.
- 1,000 bootstrapping runs were performed to generate confidence intervals over each estimated parameter value.

The new text to address the comment is:

To provide estimates of chemokine concentrations and secretion rates in lung tissue, chemokine levels were measured at 4-6h intervals during the first 48 h of infection in wells containing approximately one million human bronchial epithelial cells (Fig. 3, Table S1). The dynamic viral loads at these intervals have been reported previously by us for cultures infected with seasonal H1N1 virus, pandemic H1N1 virus, and avian H5N1 virus (Mitchell et al., 2011). IP-10 concentration increases were observed by 8h post-infection (p.i.), and RANTES by 16h p.i..

We estimate chemokine production rates, r , by adapting the delay differential equation model of influenza infection described in (Mitchell et al., 2011) Eq. 1 by adding one new equation ($\dot{C} = rI_{1\tau_3} - dC$) to model chemokine production. Strain-specific values for r were found by fitting the equations to the experimental data in Table S1 using a genetic algorithm to minimize the log squared error between the model and the data while holding the rest of the parameter values constant. Next, 1,000 bootstrapping runs were performed on each modeled strain to generate confidence intervals over r . The results of these fits are shown in Table 1 (viral secretion rates are from previous fits in Mitchell et al., 2011). IP-10 and RANTES secretion rates are

aggregated in the spatial model (S2.2).

... Eq. 1 ...

τ subscript variables denote delay terms, signifying the value is the population quantity in existence at time $t - \tau$. Table S2 summarizes population and parameter values and descriptions. ~~Strain-specific values for r were found by fitting the equations to the experimental data in Table S1. The results of these fits are shown in Table 1. IP-10 and RANTES secretion rates are aggregated in the spatial model (S2.2).~~

2. **I am not completely clear on many of the parameter choices in your model. Yes, you cite several studies for the parameter choice but I would like to see a more thorough review of which parameters you chose and why. The issue is that your model is not ODE model used in many of the studies you cite, so it is not straightforward to go from one model to another assuming that parameters must be the same (or similar). Specifically, how do you obtain 1257 cell/hour emergence from LNs? T cells decay with average lifespan of 3 days? IgM-mediated viral clearance increase of 10 fold? Constant T cell production after day 5?**

We agree with the reviewer that there is no easy answer for which parameters to include or exclude from the model. We added a new paragraph to the beginning of the Model Parameters subsection of the Modeling and Methods section to better explain our reasoning:

We include only parameters that directly address the role of T cells and T cell migration because our study focuses on the role of T cells and the chemokine effects of T cells in influenza virus infection. Thus, we included parameters that affected virus, T cells, and chemokines. With regards to the influenza, we chose parameters involving viral replication, infectivity in epithelial cells, and virus decay and diffusion rates based on our earlier study. The chemokine decay rate, diffusion rate, and secretion rate parameters are relevant to dynamic chemokine gradients required for T cell chemotaxis. Because we wanted to use our model to test the role of T cells and T cell migration in influenza clearance, we included multiple T cell parameters, such as T cell production rate, T cell death rate, T cell kill time, and T cell migration rate. We added IgM because many studies have argued for the importance of antibody mediated virus clearance. However, because our study is not addressing the role of antibody in clearance of influenza, we did not test the full range of B cell responses directly by including them in our parameters.

The reviewer is correct in observing that agent-based modeling is different from ODE models, we think that using parameters from the ODE models is a good first estimate. We included parameters that ODE models have determined (cited in Table 2) and in cases where appropriate (such as the viral infection rate), we adapted values from an ODE model to mirror the expected behavior in a spatial environment.

The specific parameters identified by Reviewer 1, each one is listed with a citation or justification in Table 2. For the case of the T cell production rate, there is also a subsection in the paper that describes our methodology (Models for Parameter Estimation). For the case of the T cell decay rate, three days was an error as the correct value was four days. This has been fixed in the manuscript. These values are first referenced in the Model Implementation subsection of the Materials and Methods section, before the justification for the values are referenced. To address this inconsistency, we moved the Model Parameters subsection to appear before the Model Implementation subsection, and we added more direct references to Table 2 and the Models for Parameter Estimation subsection for clarity.

The main finding of our study through sensitivity analysis shows that few parameters actually affect influenza clearance. Our findings confirm that despite the difficulty selecting parameters to populate the model, selection of most parameters is not likely to be a crucial factor in determining our results.

3. **The number of simulations performed seems small (50). Will statistics be better if you run more reasonable number of simulations as recommended for bootstrap analyses (e.g., 500-1000).**

The spatial nature of the model made it difficult to perform a full 1,000 sample bootstrapping run. Further, bootstrapping works by shuffling residuals of a model fit to generate new data sets for the purpose of generating confidence intervals over fit parameters. In this case, we were simply interested in sampling the variance of the stochastic model itself. We chose 50 runs of each model parameter set as that number was at the larger end of what was computationally tractable.

To satisfy the reviewer's request, we performed an additional 50 samples of each modeled strain of influenza, doubling our sample size. Avian influenza's maximum standard deviation changed from 37.2 to 35 (5.9%), seasonal influenza changed from 242 to 212 (12.4%), and pandemic influenza changed from 230 to 231 (0.4%). The text and Figure 4 have been updated to include the new model runs.

We think the small change between the old and new values demonstrates that 100 runs is sufficient, but we can perform more samples if the reviewer deems it necessary.

4. **How important is the assumption of diffusion rate of chemokines being slower than rate of viral spread? What if chemokine spreads faster? Will control be better?**

This is a good question. It is important to note that this is a different question than that regarding the direct comparison between the viral diffusion rate and the chemokine diffusion rate as the chemokine's effect is only seen in the gradient of the concentration and the location of the local maxima, while the viral effect is seen in the spread and concentration of the virions themselves.

We added an explanatory text to the Spatial Effects subsection of the Results section to better address this question:

Finally, the spatial nature of our model reveals that chemokines can diffuse more slowly than the rate infected cells and virus expand, thus misdirecting the T cells. It takes time for infected cells to begin producing chemokine, while the preexisting areas of high chemokine density are slow to decay. Thus, T cells whose movements respond to the spatial layout of the chemokine gradient can become trapped, failing to locate new regions of infected cells in the growing plaque. The differences in the infection outcomes between the three strains illustrate the effect of the infection spreading faster than the chemokine maxima can move in the case of pH1N1. The velocity of the chemokine maxima positions depends on a combination of many factors, including the chemokine diffusion, decay, delay, and secretion rates (Table 2). ~~This effect is most pronounced in the pH1N1 strain, where the plaque expands more quickly than the chemokine can diffuse.~~

5. **Your references are not formatted properly. Please correct.**

We use the Mendeley reference manager to manage our citations. We used the official JTB style formatter for Mendeley, along with the specific bibliography style file supplied by JTB, `elsarticle-num-names.bst`, to format our references with latex and bibtex. We are happy to continue working with the editor to make our reference style compatible with the expectations of the journal but require further guidance on this point.

Reviewer 2

- **(Introduction) P5 L4: "Pathogenicity is a function of the virus"....too vague. Function of what in the virus?**

We added a new sentence to clarify our goals. The question, "a function of what in the virus?" is exactly the question this paper tries to answer.

Pathogenicity is a function of both the virus and its interaction with the immune response. Our model explores how various features of the virus and the host immune system interact to produce observed differences between strains.

- **P8: “T cells emerge from the lymph node at five days post infection (p.i.) at a rate of 1,257 cells per hour. T cell travel time from the lymph node to a random location on the lung’s surface is six seconds. If chemokine is not encountered, the circulating T cell returns to a new location in the lung after another six seconds. If a circulating cell encounters chemokine, it changes state to chemotaxing and follows the chemotactic gradient to the FOI. Circulating T cells decay exponentially with an average lifespan of three days.” Add some references for all these data.**

This statement is related to Reviewer 1, point 2 (see above for how we addressed it). In summary, we added a new section describing our choice of the parameter set. We moved the Model Parameters subsection up to appear before the Model Implementation subsection. We also added references to Table 2, which contains citations and justifications for each of our parameter values.

- **P8: “IgM is modeled by...”. IgM never defined.**

We have added a new sentence to the Model Definition section.

After four simulated days, Immunoglobulin M (IgM) is introduced into the model by increasing the viral decay rate.

- **P11: “This ODE was fit to data taken from (Miao et al., 2010) using Matlab’s nlinfit function in order to obtain a value for σ . The final value was found to be 1,257 per hour.” Any confidence intervals over σ ?**

Although we don’t report confidence intervals, we explain how we obtained the value in the section “Models for Parameter Estimation”. σ is part of our sensitivity analysis (T cell secretion rate) and shows its value does not have a strong effect on the model results within a biologically plausible range.

- **P12: “Because the CyCells model is stochastic, we conducted 50 runs of the model for each of the three influenza strains, each run initialized with a unique random seed (Fig. 4). ” So 50 runs with the same parameter set? Not clear**

We have expanded our sample from 50 to 100 runs as per Reviewer 1, point 3. We have clarified that all 100 runs use the same parameter set.

Because the CyCells model is stochastic, we conducted ~~100~~50 runs of the model with the same parameter set for each of the three influenza strains, each run initialized with a unique random seed (Fig. 4).

- **P12: “We estimated variance across runs by calculating the standard deviation at each time point.” Standard deviation of what? Viral load? Please specify.**

We added a sentence to the beginning paragraph of that (Model Results) section:

Because the CyCells model is stochastic, we conducted 100 runs of the model with the same parameter set for each of the three influenza strains, each run initialized with a unique random seed (Fig. 4). We count the number of infected cells in the simulated plaque at every time point. For the avian and seasonal strains, all model runs cleared the infection by day 10. Conversely, each simulation of the 2009 pandemic influenza led to uncontrolled infection.

We estimated variance across runs by calculating the standard deviation of the number of infected cells at each time point. ...

- **P13: Not clear how the Sensitivity Analysis has been performed. Were the parameters varied simultaneously? It appears that they have been varied independently (“Each of the remaining sixteen parameters was varied independently”). See PMCID: PMC2570191 for a comprehensive review of global sensitivity analysis methods, especially in the context of an ABM/stochastic model.**

We thank the reviewer for this suggestion and reference. We performed a new sensitivity analysis using Latin hypercube sampling (LHS) and partial rank correlation coefficient (PRCC) analysis following the suggested citation to supplement our original analysis. The new sensitivity analysis was performed only for the sH1N1 data due to computational constraints. 333 sample points involving 17 parameters (those of Table 3 plus the viral secretion rate) were generated using LHS and each point was evaluated three times, for a total of 999 model runs. One sample point caused undefined model behavior and was removed, leaving 996 samples.

The results of the PRCC analysis confirm and quantitate the original sensitivity analysis. As we found previously, all sensitive parameters remain sensitive under PRCC ($P < 0.01$ with a high absolute Spearman's ρ). Infected Cell expression time and T Cell Expected Lifespan at the FOI were confirmed to be significant with lower values of ρ . These results have been added as a new paragraph to the Sensitivity Analysis subsection of the Results section in the main paper (text), Table 3 in the main paper (below), and expanded upon in the Supplement with text (S2.4) and three new figures (Figs. S6-S8).

Partial rank correlation coefficient (PRCC) analysis was performed on the sH1N1 strain using the same 16 parameters plus the viral secretion rate to determine the strength and significance of the parameters' partial correlation with the model output (S2.4). The five viral parameters are confirmed by the PRCC analysis to be significantly correlated with model output (Table 3). This indicates our less computationally intensive OFAT analysis is accurate.

Category	Parameter	Avian H5N1	Seasonal H1N1	Pandemic H1N1
Chemokine	Chemokine Decay Rate	bounded stable	bounded stable	stable
	Chemokine Diffusion Rate	bounded stable	stable	stable
	Chemokine Secretion Rate	bounded stable	stable	stable
T Cell	Circulation Time	bounded stable	bounded stable	bounded stable
	T Cell Kill Rate	stable	bounded stable	bounded stable
	T Cell Speed	stable	bounded stable	bounded stable
	T Cell Age in Blood	stable	bounded stable	bounded stable
	T Cell Age at FOI	stable	bounded stable [†]	bounded stable
	T Cell Production Rate	bounded stable	bounded stable	bounded stable
Delay	Apoptosis Time	stable	stable	stable
	Expression Time	peak change	peak change [†]	peak change
	Incubation Time	peak change	peak change	peak change
Virus	Viral Response to IgM	sensitive	sensitive [†]	sensitive
	Infectivity	sensitive	sensitive *	sensitive
	Viral Decay Rate	sensitive	sensitive *	sensitive
	Viral Diffusion Rate	sensitive	sensitive *	sensitive

Table 3. Sensitivity Results: The above parameters were varied over predetermined ranges in isolation, resulting in new model runs for every new value tested (Figs. S3-S5). The results of the sensitivity analysis were then qualitatively evaluated for each individual parameter. A model run's behavior was determined by examining the height of the peak of the infection at day 5 post-infection and the number of infected cells at day 10 post-infection. Each combination of influenza strain and free parameter was classified as belonging to one of four categories. Parameters were classified as **stable** if all runs follow the same behavior, **bounded stable** if intermediate parameter adjustments did not affect the model's behavior, even if the more extreme adjustments did, **peak change** if the peak of the infection differs but the result at day 10 is the same, and **sensitive** if any level of change in the parameter affects the resulting model behavior. PRCC analysis was also performed for the sH1N1 strain over these parameters (Figs. S6-S8). Bold text in the seasonal column denotes significant Spearman rank correlation ($P < 0.01$) over the time period where the parameter was active. † indicates a maximum *absolute* Spearman's ρ of less than 0.5. * indicates a maximum *absolute* Spearman's ρ of over 0.5.