MODELING NASAL TRACT INFECTIONS OF SARS-CoV-2 VARIANTS

Jason J. Pearson

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Mathematics.

Chapel Hill
2022

Approved by:

M. Gregory Forest

Boyce Griffth

Jingfang Huang

David Adalsteinsson

Alexander Chen

ABSTRACT

Jason J. Pearson: Modeling Nasal Tract Infections of SARS-CoV-2 Variants
(Under the direction of M. Gregory Forest)

The SARS-CoV-2 coronavirus responsible for COVID-19 has undergone waves of documented mutations and dominant variants. Sequencing of the SARS-CoV-2 spike, membrane, envelope and nucleocapsid structural proteins has revealed mutations specific to each variant. The cell infection and viral replication mechanisms induced by these mutations, however, remain to be quantified or identified. Clinical nasal test data has revealed a 1000-fold rise in infectious nasal titers of delta over alpha, and another 70-fold rise of omicron over delta. Here we seek to understand what infection-replication mechanisms can, and cannot, explain these clinical data. We employ a mechanistic model of SARS-CoV-2 nasal tract infection [2] that estimates infectious viral load and infected cells in the 1-3 days following an initial nasal tract infection. We explore sensitivity in nasal viral load to mechanisms impacted by mutations: (i) viral spike binding affinity to cell receptors, modeled as cell infectability; (ii) latency time: spanning virion-cell receptor binding to cellular uptake, assembly, replication, and onset of virion shedding; and (iii) the number of virions shed per day by infected cells. We find that latency time has the dominant effect on viral load: reducing the 12-hour latency for the alpha variant to 6, respectively 2, hours is sufficient to reproduce the rises in nasal test data of delta and omicron, respectively. Increase replication rate of infectious virions by infected cells has a significant influence, one order of magnitude higher rate can produce two orders of magnitude higher viral titers. We further show viral load and cell infection are weakly sensitive to cell infectability, even though stronger delta and omicron spike-receptor binding affinities are the most publicized mutation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ASL	Airway surface liquid
MCC	Mucociliary clearance
PCL	Periciliary cell liquid
RT	Respiratory tract
URT	Upper respiratory tract

CHAPTER 1: INTRODUCTION

1.1: BACKGROUND

SARS-CoV-2 has caused a global pandemic affecting almost every aspect of life in since early 2020. As of April 2022, at least 500 million cases of SARS-CoV-2 have been reported, and at least 6 million deaths globally have been attributed to it [1]. In excess of this vast human toll, the loss in economic output since the beginning of the pandemic until now is estimated to be at least $3 trillion USD [45]. Clearly, the world was not prepared at any level.

To gain further insights into what happens to a person infected with SARS-CoV-2, the Forest Group [7] developed a model to simulate novel viral infectious in the human respiratory tract (RT). The model incorporates relevant (and not previously modeled) aspects of RT infections, including the anatomy of and physiology of the RT (Table 1.2), characterization of the movement of SARS-CoV-2, and parameters characterizing the infection properties of the virus (Table 1.1). Given its generality, the model can be used to simulate any type of respiratory virus, including variants of the same virus if one knows the properties altered by mutations. The initial model focus was to gain insights into two observed clinical outcomes: nasal/upper respiratory tract infections occuring within 1-3 days of very mild exposures; and 3% to 5% [22, 24, 48] of infections in the nasal passage progress to alveolar pneumonia [8].

The model differs from other within-host models of RT infections [47, 27, 32, 35, 37, 10, 34, 39, 46, 20, 19, 28, 14] in that it incorporates mucociliary clearance (MCC). In the absence of previous exposure to a virus, MCC is the body's first immune response [25, 30, 15, 4, 41, 42, 38]. MCC refers to the advection of the mucus layer, and is the result of the the coordination of beating cilia in the periciliary liqiud lining every part of the lung (except the alveolar space). MCC functions as an immune response by moving inhaled virions to the esophagus where they are digested.

1.2: MODEL DESCRIPTION

The relevant model presented in Chen et al. 2022 to this thesis is the infection kernel. The geometry nasal passage and the different generations of the LRT are treated as cylinders (See Figure 1.1 (A)). The branching structure shown mimics the fractal nature of the lung. Parts of the URT are not included, like the nasopharynx or oropharynx, because of their complex geometry. This would be a desirable extension of this model.

Due to the intensive nature of the computation involved in the infection kernal, the generations are not connected in any sense, that is, once a virion exits generation 0 (trachea), it is not allowed to continue to move into one of the regions labeled generation 1 and would instead be removed from the model. In reality, the lung naturally is connect, and so another extension of this model could simulate connected generations in parallel, however, SARS-CoV-2 mostly affects the nasal passage and deep lung, and so this approach does not benefit the current work.

In Figure 1.1 (B), a close up of a sample cylinder modeling a RT generation is shown. Assume the radial direction is the x-coordinate, the axial direction is the y coordinate, and the azimuthal direction is the z coordinate. The computational model is concerned with three layers. The first, outermost layer is the cell wall, where epithelial cells can become infected. The second layer is the PCL, populated by cilia protruding from the epithelial cells, where diffusion is in effect. The third layer is the mucus layer, where both advection and diffusion are in effect. Inside of the third layer is the air space of the RT, but this is not incorporated into the model.

Along the outermost purple shell of the cylinder, along $x = 0$ is the cell-wall PCL interface. It has an absorbing boundary condition with probability p_{infect} in the case of a successful virion infection encounter of the cell, and a reflecting boundary condition with probability $1 - p_{infect}$ for a failed virion infection encounter. The innermost green shell of the cylinder, along $x = PCL_{gen} + mucus_{gen}$ (Note $mucus_{gen} = h_{mucus}$ elsewhere in this work. See Section 2.3.2.1). has a reflecting boundary condition. This ensure virions stay trapped inside the mucus layer and cannot escape back into the ambient air space of the RT.

Each generation begins at $y = 0$. Since MCC biased infection and viral load towards the esophagus, this boundary condition is not encountered in practice. From the biological perspective, infections do not spread via retrograde propagation, that is, mucus rarely flows backwards. Each generation ends at $y = length_{gen}$, which has an absorbing boundary condition removing virions from the computation.

The azimuthal coordinate has a periodic boundary condition, as is suggested by the cylindrical geometry. A virion, in principle, could traverse this entire dimension, however, in practice, this boundary condition is not encountered.

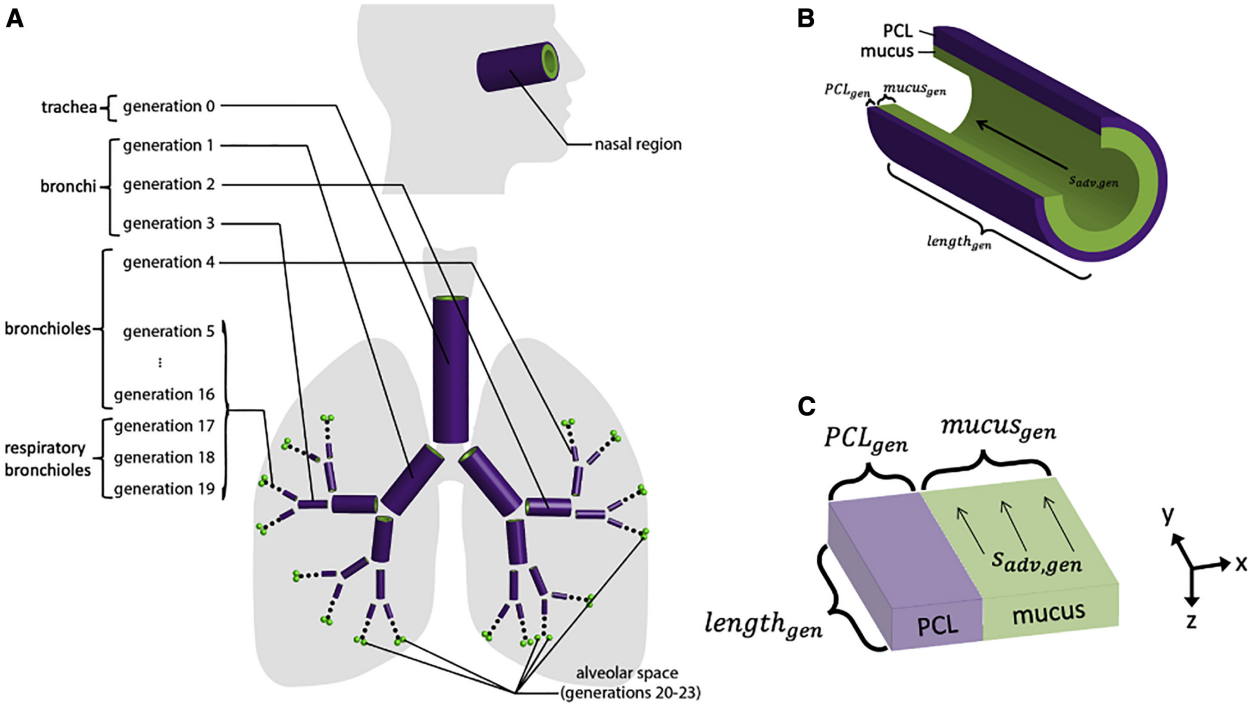

Figure 1.1: Model of the human respiratory tract. Chen et al. 2022 [7].

Figure 1.1 (C) shows an unrolled cylindrical domain. It is this rectangular prism which serves as the computational domain. To obtain this domain we can imagine making a cut on the cylinder in Figure 1.1 and flattening it. This approximation is possible because of the relatively large size of the azimuthal dimension.

The cells comprising the cell wall shown in Figure 1.2 (A) are modeled as squares filling out the rectangular (y, z) plane located along $x = 0$. Their side length is $4\mu m$. Each cell is in one of four states: uninfectable (white), uninfected (blue), infected (yellow) and shedding (red). The initially infected cell used in simulations (just transitioned from blue to yellow at $t = 0$) is, by convention, located at the center of the upstream of the generation. All the remaining cell states are randomly set to uninfectable or uninfected based on the percentage of infectable cells, a model parameter (See 2.3.2.3). Uninfectable cells (white) maintain their state throughout the entire simulation. As their name indicates, they are not susceptible to virion infection. Infectable cells (blue) when having a successful virion encounter transition to infected cells (yellow). A timer is started at the time of infection, t_i. Once the period of latency time $t_{latency}$ (See Section

2.3.1.2) has passed the cell finally transitions to a shedding cell (red). At time $t_i + t_{latency}$, the cell begins shedding virions into the ASL. The initial location of the virions is at the center of the infected cell they are released from, right at the PCL-cell interface. No virions are present in the ASL at the beginning of the simulation.

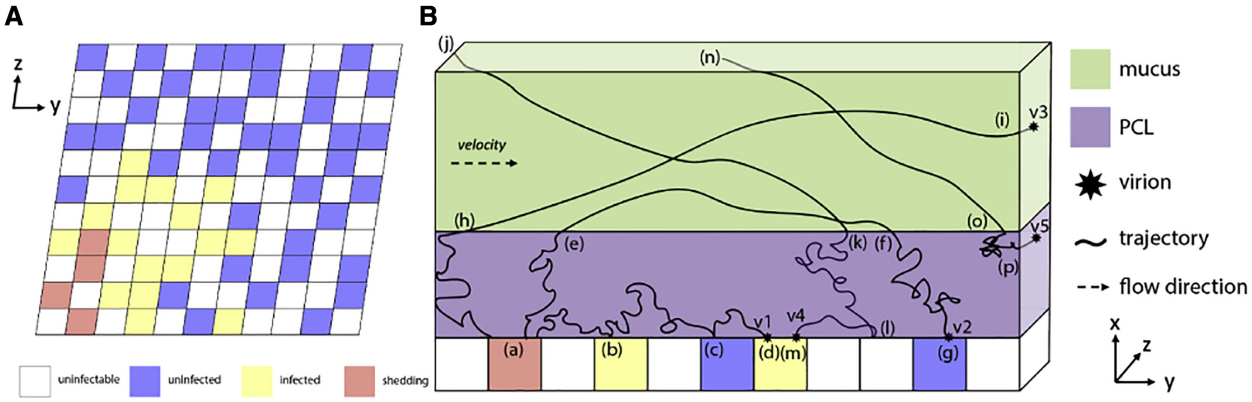

Figure 1.2: (A) Cell wall with possible states and (B) virion trajectories. Chen et al. 2022 [7].

The virions are modeled as inertialess point particles, undergoing diffusion and advection depending on their location as shown in. Once shed, the ordinary differential governing their evolution is

$$dx = \sqrt{2D_v}dW_1 \tag{1.1}$$

$$dy = \sqrt{2D_v}dW_2 + s_{adv,gen}\mathbb{1}_{x>PCL_{gen}}dt \tag{1.2}$$

$$dz = \sqrt{2D_v}dW_3. \tag{1.3}$$

such that D_v is the diffusivity of SARS-Cov-2 inside mucus, dW_i is 1-D Brownian motion, $s_{adv,gen}$ is the advection velocity of the mucus, $\mathbb{1}_{x>PCL_{gen}}$ is an indicator function, and dt is a time step. The time step of the infection in practice is $dt = 1s$, and the virions trajectories are obtained using Forward Euler. The Brownian motion is generated by the choice of randomly oriented three-dimensional unit vectors.

Diffusion is the same in each coordinate dimension, dictated by the fact the virions are diffusing in the same fluid, mucus, in any selected generation. Only the y-coordinate direction features advection, which has a strength dependent on the generation being simulated. Advection is greatest in the nasal passage and decreases when traveling deeper into the lung. The advection of the mucus layer is caused by the coordinated

beating of cilia protruding from the cell wall into the PCL layer. These cilia are assumed to not cause any advection in the PCL layer itself.

Figure 1.2 (B) depicts a single veritcal column of the cell wall in Figure 1.2 (A) with the PCL and mucus layer (advecting to the right) laid over top and multiple virion trajectories. Shed virions have two ultimate fates: they either infect an epithelial cell or are cleared from the generation. The virions labeled $v3$ and $v5$ are examples of virions which end up clearing the generation, and those labeled $v1$, $v2$, and $v4$ are examples which infect epithelial cells.

The infection kernel keeps track of various aspects of the infection. Efficient data structures store the time of infection for each infected cell t_i, its location (y, z) on the cell wall, how many times the cell is infected by a virion during the simulation (Infected and shedding cells can continue to uptake virions. See Figure 1.2 (B) virions $v1$, and $v4$), the number of shed virions, the number of free virions, the number of cleared virions (flux). This information can be synthesized to generate a heatmap of the cell wall, showing the time lapse of the infeciton over time, or the number of free virions over time.

1.3: MODEL PARAMETERS

The parameters used in the model described in Section 1.2 are shown in Table 1.1, 1.2. Both tables are populated from clinical or experimental data.

Hydrodynamic radius of SARS-CoV-2 (nm)	100
Diffusivity in airway surface liquids $(\mu m^2/s)$	1.27
Alveolar space: probability of infection of infectible cell by infectious SARS-CoV-2 virion (% per encounter per second)	0.03
Non-alveolar space: probability of infection of infectible cell by infectious SARS-CoV-2 virion (% per encounter per second)	0.3
Duration of eclipse or latency phase $(hours)$	12
Duration of virus replication phase $(days)$	3
Replication rate of virions inside infected cell (counts per day)	50,000
Percent of replicated virions that are infectious $(\%)$	4 (2000/day)

Table 1.1: Assumed alpha variant coronavirus properties. Chen et al. 2022 [7].

Table 1.1 lists the best available measurements or parameter estimates for SARS-CoV-2. The table includes the hydrodynamic radius of a SARS-CoV-2 virion, which can be taken to be the radius of the bounding sphere around the virus. The diffusivity in the airway surface liquids is next, describing the extent thermal motion can move the virion in the ASL. Two differing values exist for the probability of infection of

a infectible cell by a infectious SARS-CoV-2 virion, corresponding to the alveolar space in the deep lung, and the non-alveolar space outside that region. This is due to the differing proportion of the types of cells in those regions. The non-alveolar space has a greater density of epithelial cells, and therefore has a higher susceptibility to infection. The duration of the eclipse phase is the time it takes a cell from a successful infection event to the moment it begins shedding. The duration of the virus replication phase is how long the infected cell is thought to shed virion prior to the immune system mounting a response to replace it with an uninfected cell. This places an effective time scale limit on how long the model can be run, as by assumption it does not immune responses. Lastly, the replication rate of virions inside an infected cell refers to the raw number of virions the cell produces. Most of these produced virions are incapable of causing infections, and so only the small fraction capable of causing an infection are considered, i.e., the percent of replicated virions that are infectious.

Table 1.2 gives measurements or parameter estimates for the human RT. First it has the length of each generation. Deeper locations in the RT tend to have smaller lengths. Next is the assumed mucus layer thickness (See the green layer in Figure 1.1). This is added to the $7\mu m$ thick PCL to create the entire ASL. Next is the assumed advection, largest in the nasal passage, and monotonically decreasing until generation 19. In the deepest part of the lung (not shown in the table), no advection is present. Next the time for the mucus layer to advect the full length is computed from the assumed length and assumed advection. The final three columns are a result from [7], describing the probability a generation will become infected with a virion starting at the air-ASL interface. Virions started this way will either infect or advect out of the generation. Many such realizations were performed to determine these probabilities. Particularly noteworthy is the relative accuracy of a virion started at the midpoint of capturing the probability to infect for a virion started in the middle two quartiles of a generation.

Note that with the exception of the the hydrodynamic radius and the diffusivity in ASL, that all the parameters are estimates for the alpha variant of SARS-CoV-2. The remaining parameters are later (Section 2.3) the subject of of a sensitivity analysis of this model.

1.4: MODEL OUTPUTS

The infection model outputs of greatest interest are the total number of shed virions and total number of the infected cells. However, as alluded to previously, the model does keep track of other information, like

Generation	Assumed length (cm)	Assumed mucus layer thickness (μm)	Assumed advection (mm/min)	Time for mucus layer to advect the full length (hrs)	Probability of infection (starting at upstream end)	Probability of infection (starting at midpoint)	Probability of infection (starting in middle two quartiles)
Nasal	1.3×10^{1}	17	8.8	0.246	0.9093	0.6863	0.6721
0	1×10^{1}	50	5.5	0.303	0.4177	0.1733	0.1727
1	4.36×10^{0}	44.4	3.91	0.186	0.3	0.0989	0.1056
2	1.78×10^{0}	38.8	2.49	0.119	0.2229	0.0644	0.07
3	9.65×10^{-1}	33.1	1.54	0.104	0.2619	0.0868	0.088
4	9.95×10^{-1}	27.5	0.889	0.187	0.6024	0.3164	0.312
5	1.01×10^{-1}	21.9	0.496	0.339	0.9046	0.6857	0.6625
6	8.9×10^{-1}	19.6	0.296	0.501	0.9857	0.8693	0.8463
7	9.62×10^{-1}	17.3	0.167	0.96	0.9999	0.9901	0.9799
8	8.67×10^{-1}	15	0.101	1.43	1	1	0.9983
9	6.67×10^{-1}	12.7	0.0616	1.8	0.9988	1	0.9999
10	5.56×10^{-1}	10.4	0.0396	2.34	0.9901	1	1
11	4.46×10^{-1}	9.32	0.0252	2.95	0.9575	1	1
12	3.59×10^{-1}	8.29	0.0165	3.63	0.8973	1	1
13	2.75×10^{-1}	7.26	0.0113	4.06	0.8029	1	1
14	2.12×10^{-1}	6.23	0.008	4.44	0.7757	1	1
15	1.68×10^{-1}	5.2	0.0057	4.93	0.7188	1	1
16	1.34×10^{-1}	4.7	0.0041	5.47	0.6548	1	1
17	1.2×10^{-1}	4.21	0.0028	7.07	0.6509	1	1
18	9.2×10^{-2}	3.72	0.0014	11	0.5678	1	1
19	8×10^{-2}	3.22	0.001	13.9	0.5957	1	1

Table 1.2: RT physical dimensions, advection speeds and probability of infection within generation. Chen et al. 2022 [7]

the infection times, the number of virions which infect each cell, and the amount of virions (flux) ultimately leaving a generation.

The results, showing some of the model outputs, from Chen et al. 2022 [7] are displayed in Figures 1.3, and 1.4.

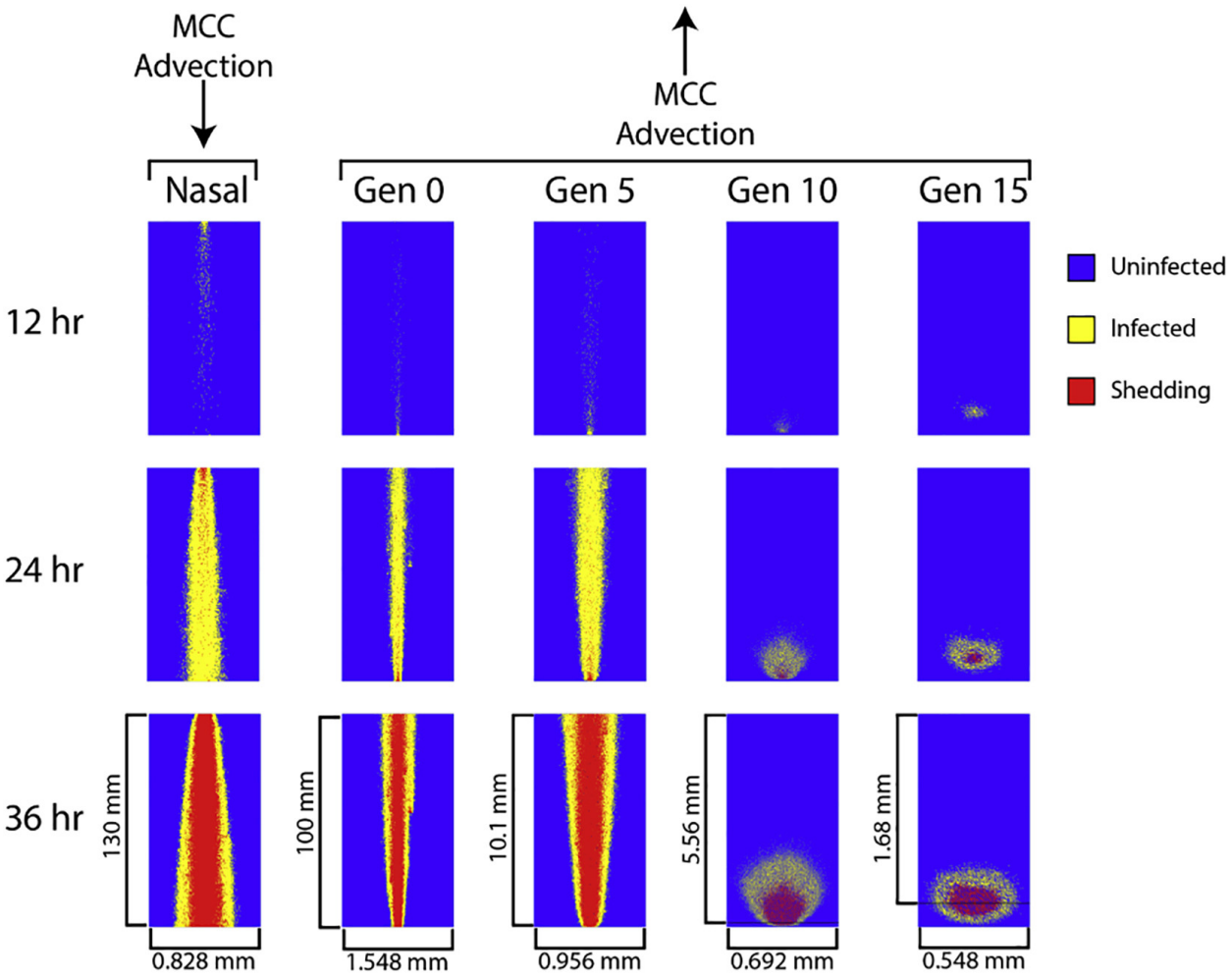

Figure 1.3: Model outputs: cell heatmap time lapse. Chen et al. 2022 [7].

Figure 1.3 shows a time lapse of how infections can evolve inside of the various parts of the RT. Each pane in the figure shows the cell wall at a specified time shown off to the right. The initial condition, not shown, is a single infected cell which has only just begun shedding in the upsteam direction In each generation depicted. Essentially, the cell walls start almost entirely uninfected, and so the initial condition of the heatmaps is a state of being completely blue (uninfected) with the exception of a single cell. The arrows on top of each column indicate the direction of advection. The nasal passage is the only region simulated with downward mucus advection (towards the esophagus). Generations 0, 5, 10, 15 all feature upward mucus advection.

Diffusion is the same in each case, but the affect from advection is relatively stronger than that of diffusion in the generations greater than 10, as can be seen from the streaks of infection gradually seen to emerge as the simulation time reaches 36 hours. In the LRT, i.e., generations 10 and 15, diffusion dominates

over advection, and the spread of the across the entire length is not realized. The final spread of the infection also assumes a more radially symmetric character, as would be expected in the case of a purely diffusive phenomena.

One major caveat to making comparisons between the columns is their scaling. The horizontal dimension for the plots of the nasal passsage, generation 0 and 5 is zoomed in to focus on the region where the infection occurs, and does not show the entire width of the cell wall. As already mentioned, the advection in these three locations dominates the diffusion, and as a consequence, the region of infected cells is never able to wrap around the width of the generation, that is, the periodic boundary condition in the azimuthal direction is never in pratice encountered. The vertical dimension of these plots does include the entirety of each generation.

For generations above 10, the general pattern shown is at 12 hours post initial cell shedding, a thin streak of cells becomes infected with daughter virions. A second generation of cells have yet to begun shedding their own daughter virions. Next, at 24 hours post initial cell shedding, a second generation of cells have begun shedding, as can be seen by the small red regions near the initial infection site. Naturally, these cells close to the initially infected cell are very likely to become infected earlier, and transition to shedding cells themselves earlier. Last, at 36 hours post initial cell shedding, a large streak of cells have entered a shedding phase. The heatmap further suggests this large streak of infected cells extends beyond the nasal passage, and into the adjacent nasopharynx or oropharynx.

For generations 10 and 15, the general pattern of infected cell propagation is similar, but the combination of weakening advection and thinner mucus layers contribute to diffusion becoming the dominant mechanism of spread of viral load and infection.

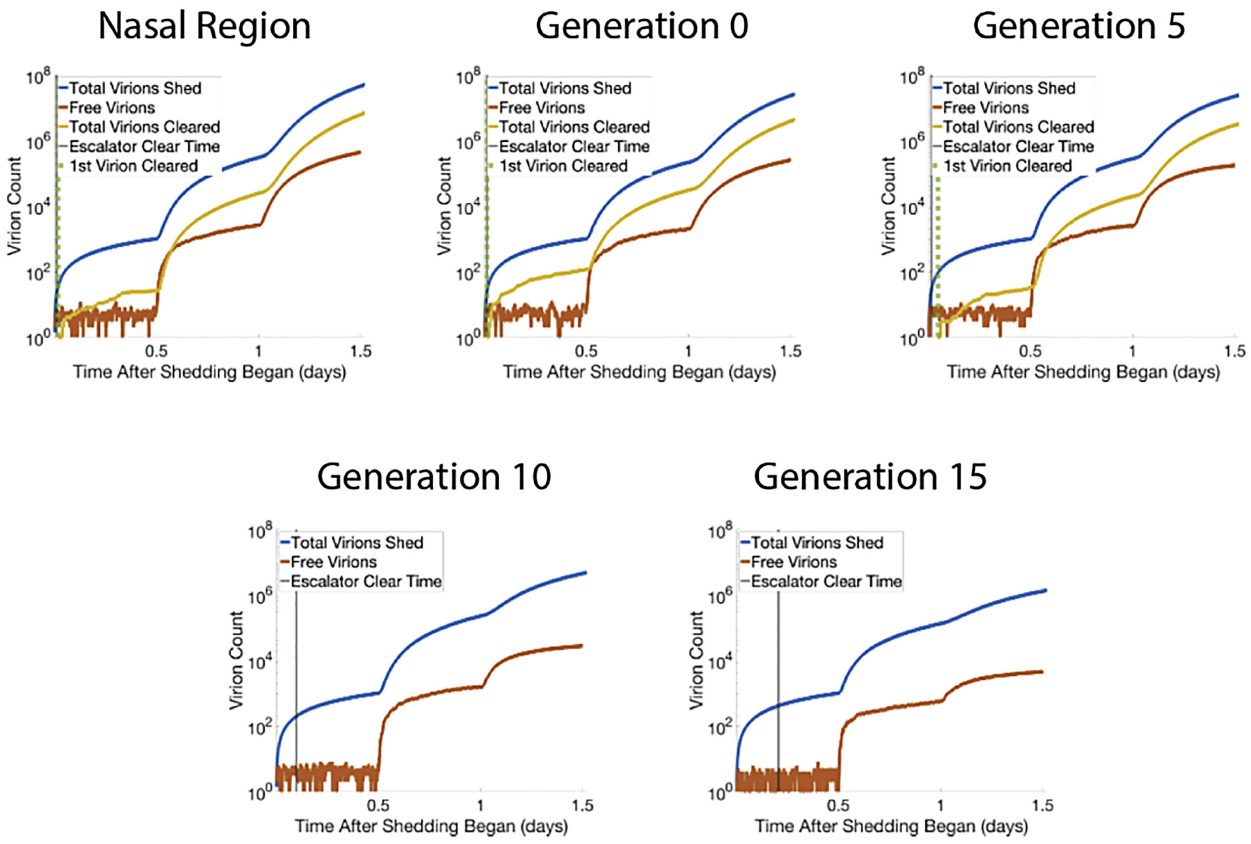

Figure 1.4: Model outputs: infections vs time. Chen et al. 2022 [7].

Figure 1.3 shows the change of the cell infection over time, whereas 1.4 shows the change of the viral infection of the same simulations over time.

The blue curves show the total virions shed over time. The yellow curves show the total number of virions cleared, also referred to as virion flux. The plots include a dotted green line indicating the time it took for the first virion to leave the specified generation, and a black line, referred to as escalator clear time, which specifies the time it takes for a mucus tracer particle starting upstream in the generation to exit. Especially in the generations above 10, advection is strong enough that it takes only around 1 hour at most for the first virion to exit. This time is at least the escalator clear time, as a virion will almost certainly not diffuse only the direction of advection. The red curves show the free virions over time. Free virions are those which are inside the ASL, continuing to diffuse and advect and have yet to find a cell to infect or clear. The initial 0.5 days of each plot show the number of free virions to be small. The reason for this is the combination of a large number of infectible cells which have yet to become infected combined with a small number of virions

being released. This implies most of the virions released from the initially infected cell find a new cell to infect. The blue curves are a sum of the yellow and red curves.

The shape of the total virions shed (blue) curves exhibit an oscillation between periods of exponential and linearly dominated growth. Periods of exponential growth are begun when a new generation of cells transitions from infected to shedding. Once a generations transition has mostly completed, a period of linear growth sets in, which appears like a decrease in slope in the semi-log plots.

Generation 10 and 15 lack the 1st virion cleared time, and the reason again can be traced to the fact advection is relatively weak compared to diffusion in this region. No virions are able to escape these generations as a result.

1.5: CONCLUSION

In this introductory chapter we have summarized the baseline pre-immunity model I inherited from my collaborators, published in Chen et al. 2022 [7], with best known parameter values for the 2020 alpha variant of SARS-CoV-2. Some key information generated by their work is:

- There is a strong likelihood of nasal infection following exposure.
- Nasal infection is the precursor and source of alveolar infection (deep lung infections only occur from direct deposition of virions into deep lung, not by retrograde progression from the lower RT).
- This fact compels a dedicated focus on the sensitivities associated with nasal passage infection and the deep lung.
- Virion parameters in Chen et al. 2022 [7] were for the alpha variant.

From here, my contributions begin, focusing on the nasal passage and the striking exposure outcomes that have been documented from the alpha variant to the delta variant and then omicron variant.

CHAPTER 2: SENSITIVITY ANALYSIS

2.1: INTRODUCTION

We would like to answer various questions about SARS-CoV-2 respiratory infections. What mechanisms are modified by the waves of viral mutations and variants? Which mechanistic parameters dominate or weakly affect infection outcomes? How stochastic are infection outcomes in the nasal passage where we have the majority of clinical data? Can the model reproduce various types of clinical data?

To answer these and other questions, we need to perform sensitivity analysis on the model in Chapter 1. Informally, sensitivity analysis is a collection of techniques to determine which input parameters have the greatest affect in the outputs. Formally, sensitivity analysis considers a physical model $\mathbb{M}$ of form

$$\mathcal{Y} = \mathcal{M}(\mathcal{X}, P) \tag{2.1}$$

such that $\mathcal{X}$ is a random vector in $\mathbb{R}^n$, capturing the variability of the input parameters to the model, $\mathbb{Y}$ is a vector of model responses (also random), and P is a vector of deterministic parameters in the model (if present). The components of the vector $\mathcal{X}$ are considered to be independent, although methods do exist in the case where the components are dependent [18].

Many different sensitivity analysis methods have been developed which can be used to determine the most influential inputs for a model [12]. The essence of these methods is to generate measures of sensitivity based on how much of the variances, in the output is captured by the variation of the input.

We employ a type of linearization method. The essential idea here is that the variation in the model output can be captured by observing partial derivatives around certain values. Larger magnitude gradients correspond to greater output sensitivity.

2.2: SENSITIVITY ANALYSIS PLATFORM OVERVIEW

To perform a sensitivity analysis, we develop a platform capable of exploring a specified parameter space R, simulating many infections in that parameter space to obtain statistics, and while making efficient use of computational resources.

The steps taken to acheive this in order are:

1. Select a parameter sweep region $R \subset P$. See Figure 2.1.
2. For each parameter combination $z \in R$, simulate a number of realizations $n_{realization}$ sufficient to obtain reliable statistics, that is, characterize the distribution of outcomes. See Figure 2.2.
3. Compute statistics on the distributions obtained in the previous step, for example, compute the mean. See Figure 2.3.
4. Interpolate the values in obtained the last step to create a function. See Figure 2.4.
5. Observe gradients in the generated function to determine sensitivity.

Once n_p parameters of interest have been identified, a parameter sweep region R can be defined,

$$R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_{n_p}, b_{n_p}] \tag{2.2}$$

The minimum and maximum range of these parameters is generally informed by clinical or experimental data. Within these bounds, the discretization can be arbitrarily chosen, that is, it need not be uniform as shown in Figure 2.1, which shows a sample discretization (black lines) of latency times $t_{latency} = \{4, 6, 8, 10, 12\}$ and infectious virion shedding rates of infected cells post latency phase, $r_{shed} = \{500, 1000, 1500, 2000\}$.

The Cartesian product of the discretizations of R forms the set of all parameter combinations for a parameter sweep. Each parameter combination is shown as a black point in Figure 2.1. The selection of multiple parameters allows the investigation of interaction between parameters on outcomes, if any exist.

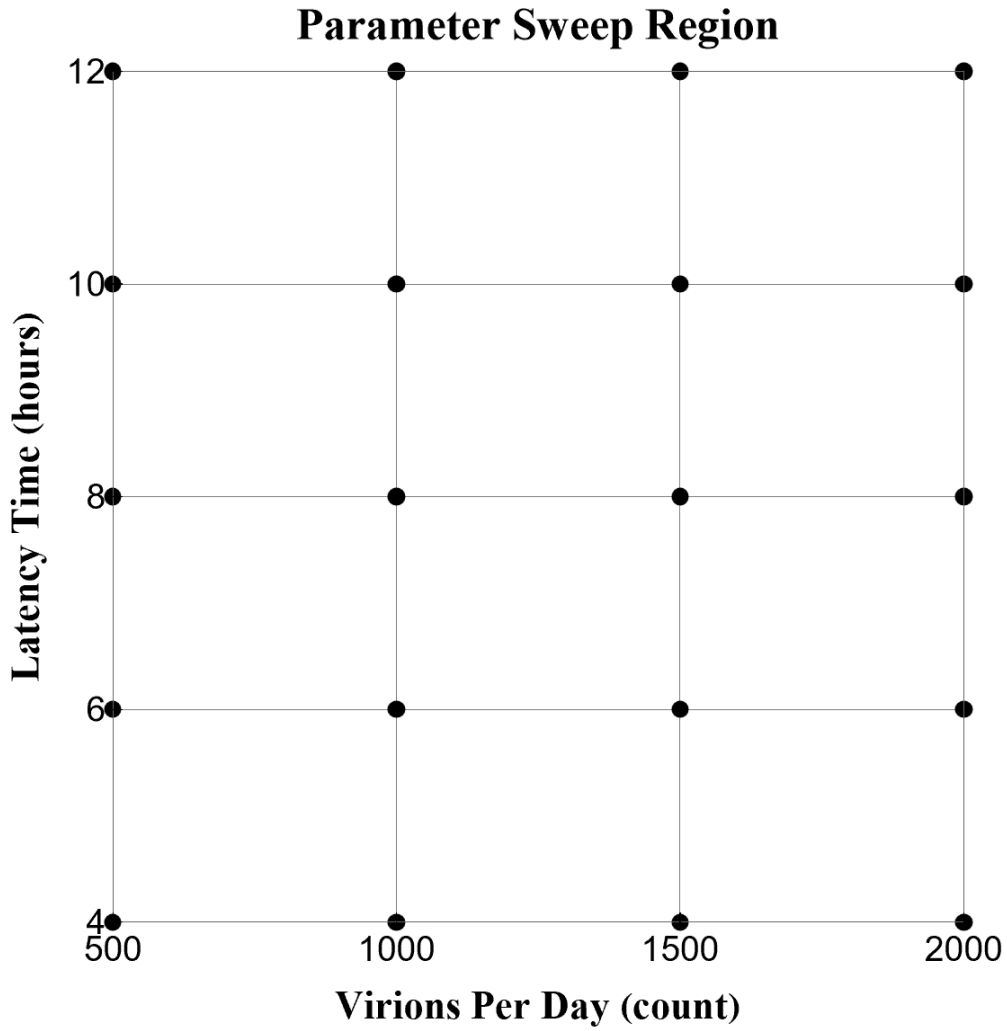

Figure 2.1: User defined parameter sweep

Each parameter combination requires some number of realizations $n_{realization}$ to obtain a distribution of outcomes. In Figure 2.2, $n_{realization} = 100$, and a sample histogram is shown to the upper-right of each parameter combination (black point). The superimposed blue curve is a Gaussian distribution fit to the data. This is intended to be suggestive, as a few of the model outcomes, namely the number of infected cells and shed virions, are both approximately Gaussian.

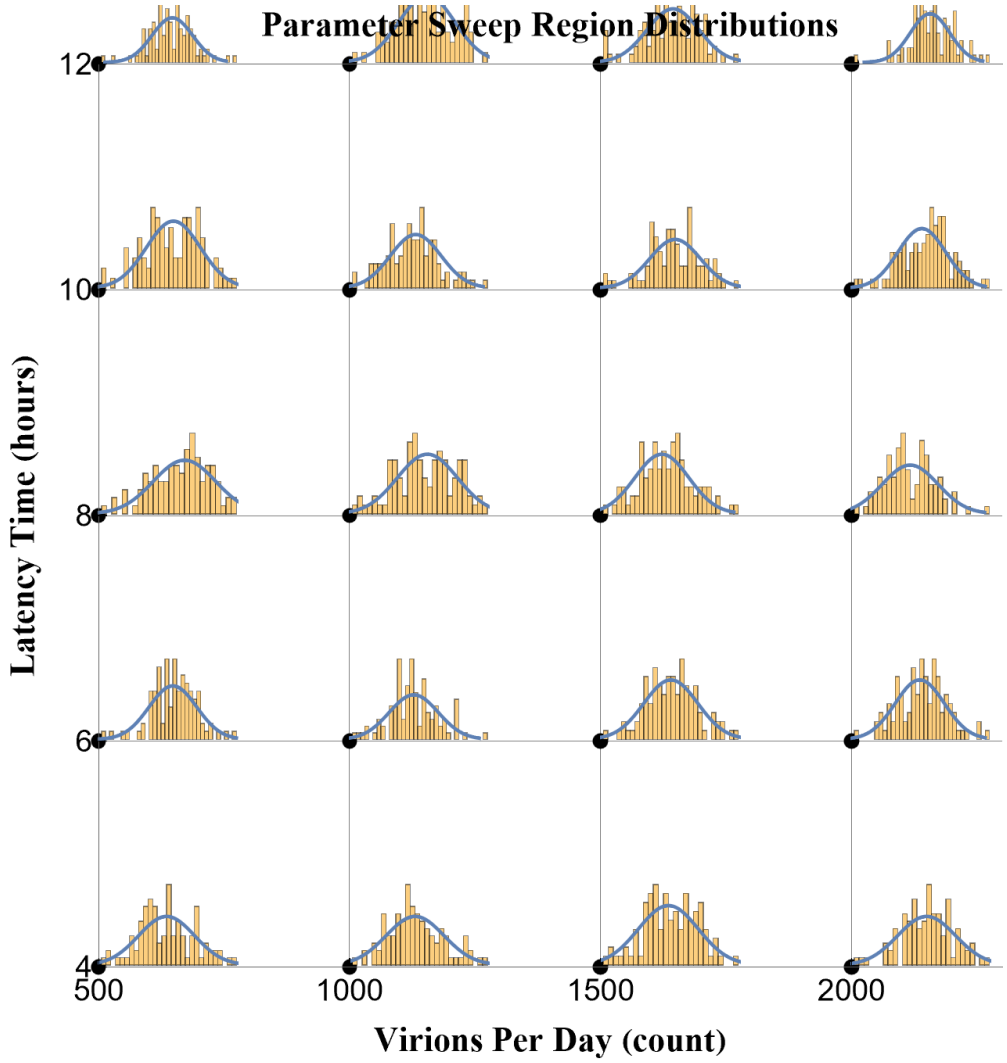

Figure 2.2: Outcome distributions on parameter sweep

Once a distribution of model outcomes is obtained, statistics are performed at each parameter combination. In the simplest case, an mean is computed of the outcome distribution,

$$\mu = \sum_{i=1}^{n} Y_i \tag{2.3}$$

however, the variance, or additional moments could be computed. The mean is shown in red below each histogram in Figure 2.3.

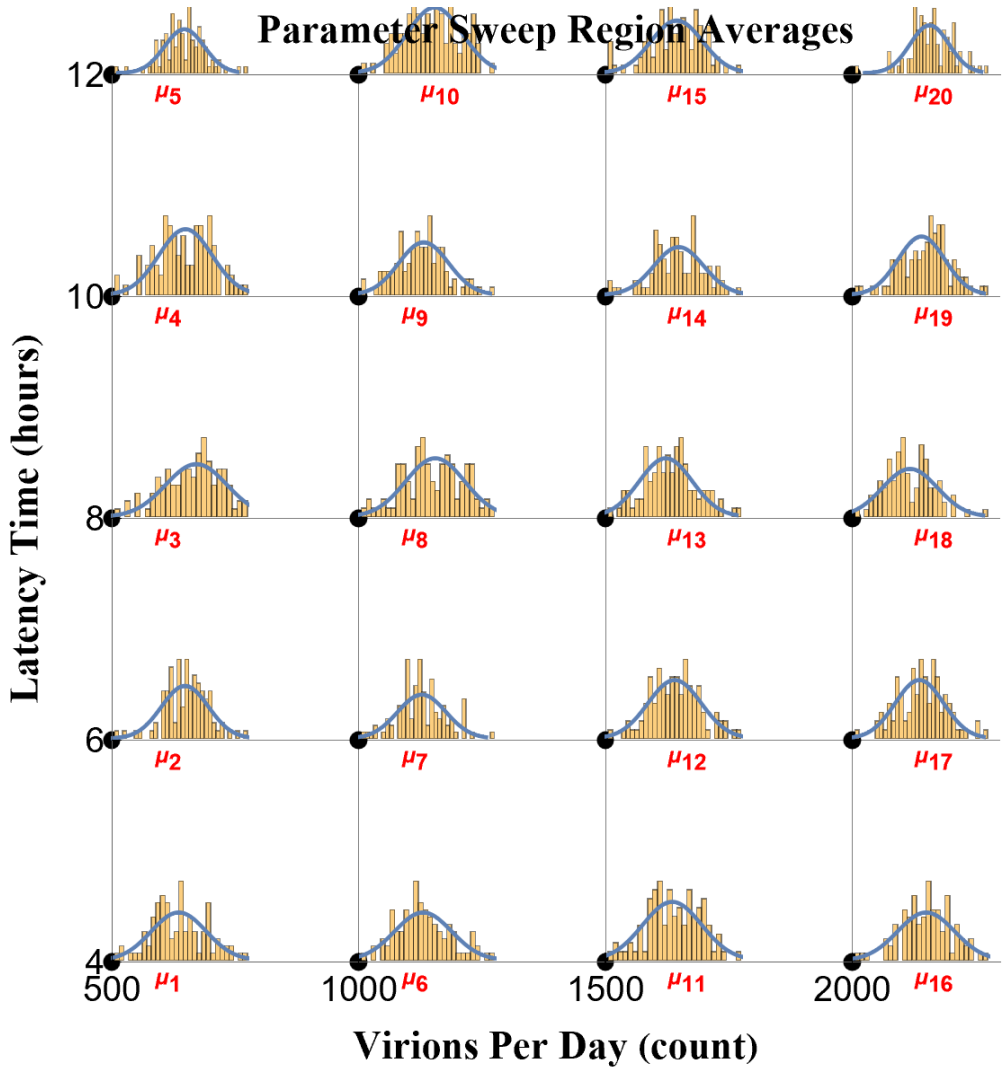

Figure 2.3: Statistics on parameter space

Lastly, to generate an interpretable function f spanning all of R, the computed statistics are interpolated. This is achieved by using so-called match functions over each grid element. In the simplest case in 2D (two parameters of the n_p are chosen), a bilinear interpolant is employed. If the corner values $f_{00}, f_{10}, f_{01}, f_{11} = f(x_0, y_0), f(x_1, y_0), f(x_0, y_1), f(x_1, y_1)$ of a grid element are known, then the bilinear interpolant on that grid element is

$$f(x, y) = a_{00} + a_{10}x + a_{01}y + a_{11}xy, \tag{2.4}$$

with coefficients determined by [36]

$$\begin{bmatrix} a_{00} \\ a_{10} \\ a_{01} \\ a_{11} \end{bmatrix} = \frac{1}{(x_1 - x_0)(y_1 - y_0)} \begin{bmatrix} x_1 y_1 & -x_1 y_0 & -x_0 y_1 & x_0 y_0 \\ y_1 & y_0 & y_1 & y_0 \\ x_1 & x_1 & x_0 & -x_0 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} f_{00} \\ f_{10} \\ f_{01} \\ f_{11} \end{bmatrix}. \tag{2.5}$$

Once an interpolant is generated for each grid element, a function over the entire region R is completed. A sample interpolant (over random data) is shown in Figure 2.4 To fill in the entire parameter space region R.

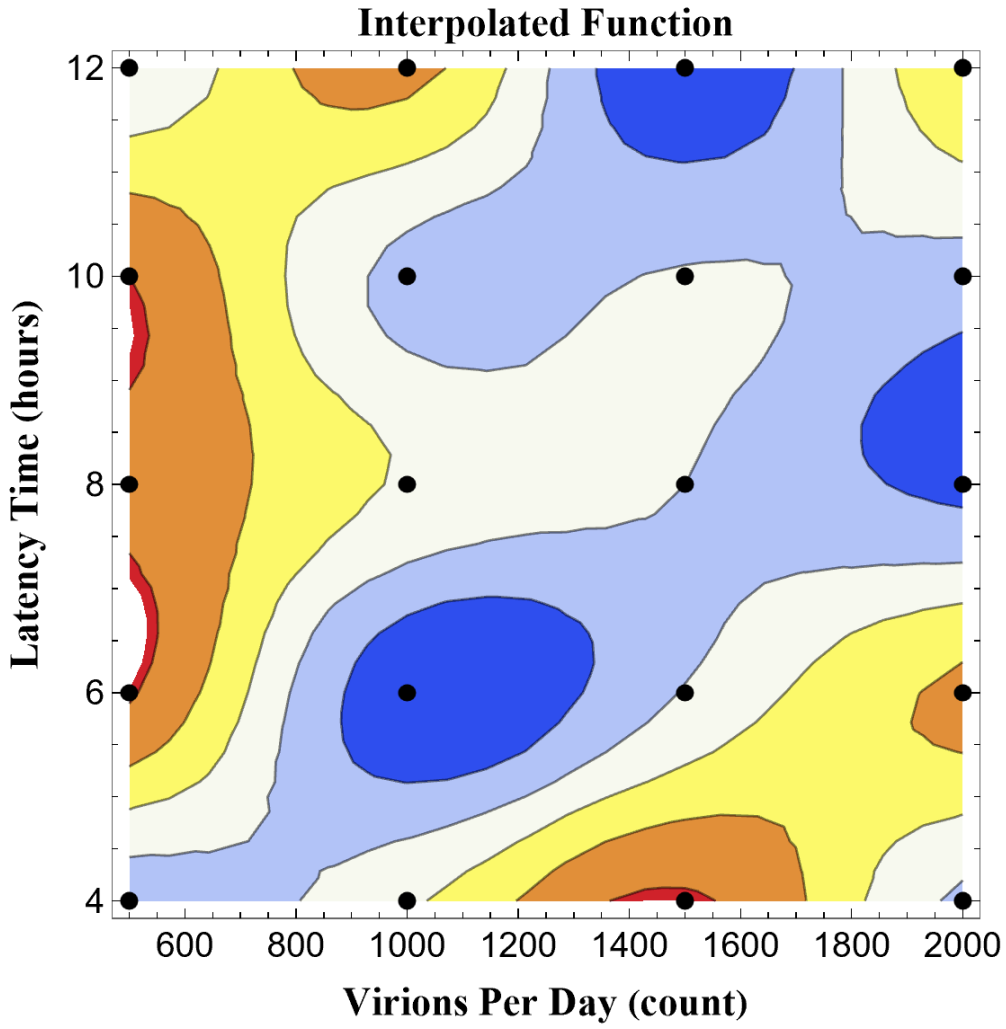

Figure 2.4: Function interpolation on parameter sweep

2.3: PARAMETER SPACE

Excluding the simulation time, the parameter space of inputs P of the model is a collection of six parameters, i.e. $P \subset \mathbb{R}^6$. Three of are viral properties: latency time, $t_{latency}$, probability of infection per encounter-second, p_{infect}, and number of infectious virions shed per day, r_{shed}. Differing respiratory track infections can be modeled by altering these. The remaining variable RT properties are: mucus thickness, h_{mucus}, the advection velocity of the mucus of the periciliary layer, v_{adv}, and percentage of infectible cells i_{perc}. These account for the variations among individual respiratory systems. A definition of each of these parameters is included for each of these parameters to provide some context.

2.3.1 Viral-specific model parameters

The virus-specific model parameters are latency time, probability of infection per encounter-second and the number of infectious virions produced per day, i.e., $t_{latency}, p_{infect}, r_{shed}$. These parameters are all depicted in Figure 2.5. In each picture, the pink is the interior of an infectible cell, the darker pink is the cell membrane, the blue is a cell receptor to which an invading virion is capable of binding and the white region is the ASL. The small ball with red spikes is a virion, black arrows show the motion of the virions.

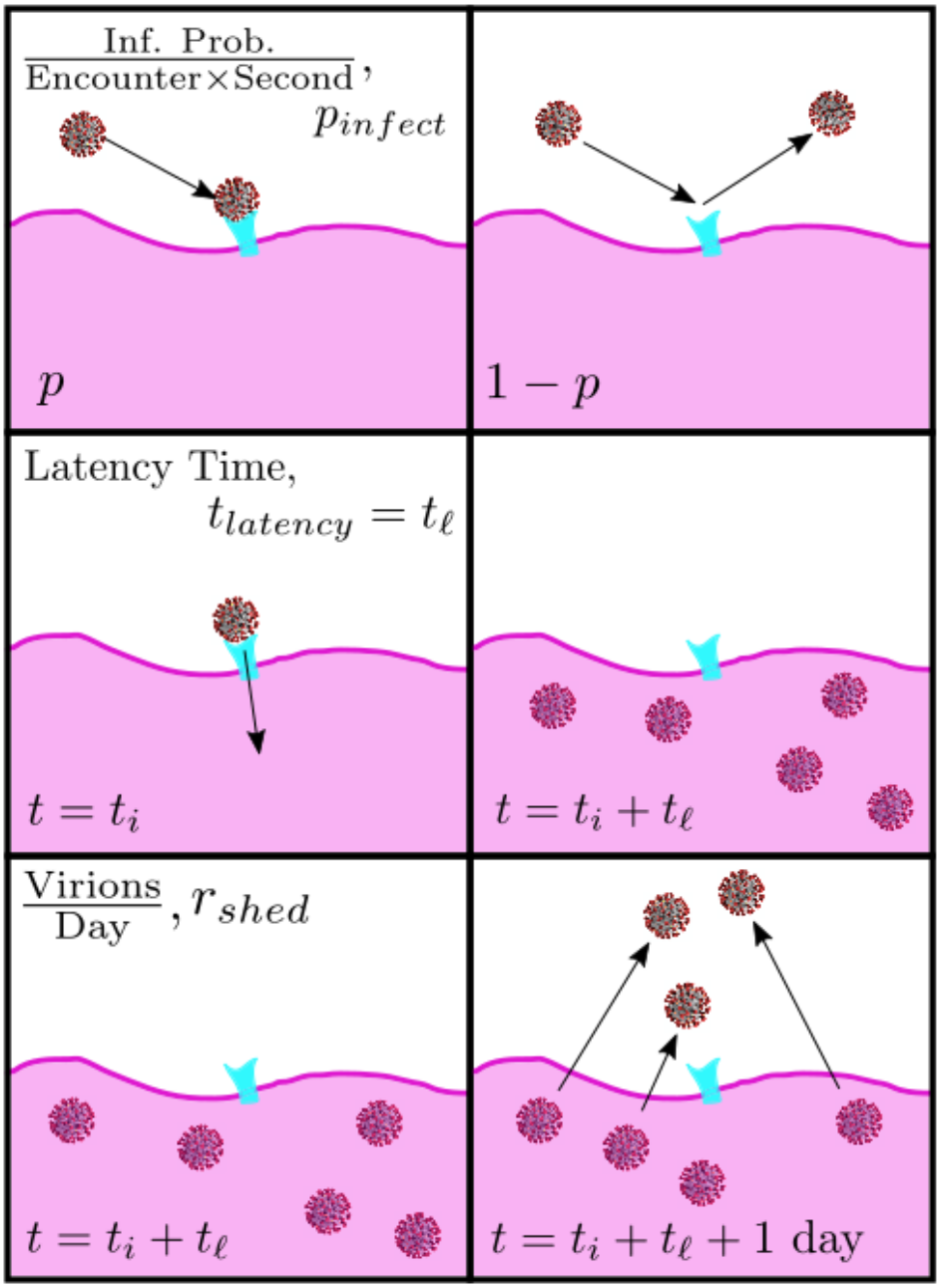

Figure 2.5: Viral parameters

2.3.1.1 Probability of infection per encounter-second

The probability of infection per encounter-second p_{infect} refers to the probability a virion within attachment distance of an infectible cell will successfully attach over the time span of one second. Outside of

the alveolar region (generations 20-23), the estimated value used for the alpha variant in [7], shown in Table 1.1 was $p_{infect} = 0.3$.

This parameter characterizes the cell wall boundary condition of the model. Virions encountering infectible cells have p_{infect} probability to successfully infect (absorbing boundary), or $1 - p_{infect}$ to fail to infect (reflecting boundary). A depiction of p_{infect} as it relates to the biology is shown in the top two panels of Figure 2.5. The left panel shows a successful infection event by a virion with probability p, whereas the right panel shows a failed infection event of probability $1 - p$, reflecting the virion back into the ASL.

2.3.1.2 Latency time

The latency time $t_{latency}$ is defined as the time taken from the moment an infectious virion binds to an infectable cell that leads to a successful infection until the infected cell starts shedding replicated virions into the airway surface liquid. (We will show that shorter latency times have a dominant, exponential effect on shed viral load within the first 24-48 hours after a cell is first infected and emerges from the eclipse phase.)

The estimate used for the alpha variant in [7], shown in Table 1.1, was 12 hours. Subsequent mutations of the alpha variant may have reduced the latency time, causing accelerated infection.

The middle two panels of Figure 2.5 show latency time. The left panel shows a successful infection event at $t = t_i$, and starts the clock for latency time. The right panel shows the cell interior after a full latency time has elasped. The interior of the cell is full of daughter virions ready to be released back into the ASL.

2.3.1.3 Number of infectious virions produced per day

The number of infections virions produced per day r_{shed} is the number of virions viable to go on to infect another cell produced in a day. A large number of the virions produced are not viable, and these are not included. A higher rate of virions ejected from a cell causes more severe infection outcomes. In [7], shown in 1.1 the baseline value for this shedding rate used was 2000 virions per day estimted for the alpha variant. Subsequent information from colleagues worldwide suggest a lower shedding rate is more realistic for the delta variant.

The bottom two panels of Figure 2.5 show the effect off the number of infectious virions per day. The left panel is the instant the cell changes its state to shedding, i.e., the moment before the first virion is released into the ASL. The right panel shows the virions have exited the cell to begin their journey to infect other cells or become cleared.

As a last note, when virions are released, they are placed in the center of the cell, at the PCL-cell interface. However, the placement of these virions when they exit infected cells may be another parameter worth considering for future research.

2.3.2 RT parameters

Three parameters which characterize the RT in individuals are the mucus thickness, the advection of the mucus layer above the PCL, and the percentage of cells which are infectible. Figure 2.6 depicts two of these parameters.

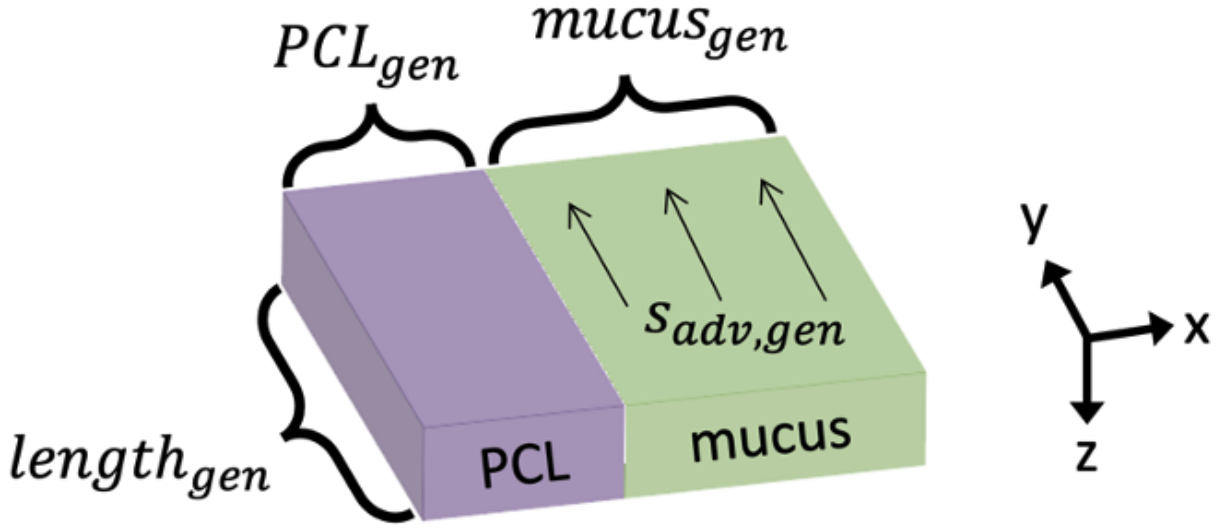

Figure 2.6: RT parameters

2.3.2.1 Mucus thickness

The mucus thickness h_{mucus} is the thickness of the mucus layer between the PCL layer and air core. The PCL is populated by cilia which when fully extended in the power stroke reach just inside the mucus layer and propel it, causing an escalator-like advection always toward the esophagus for mucus to be swallowed into the acidic environment of the stomach. Changing this parameter value is the same as expanding or shrinking the mucus layer, and can be visualized in Figure 2.6 as increasing or decreasing respectively the dimension of the green rectangular region.

Estimates of the mucus thickness in each generation are given in Table 1.2. The nasal passage has a thickness of $17\mu m$. For the remaining parts of the RT, generation 0 has the thickest mucus, around $50\mu m$, and monotonically decreases as we move deeper into the RT, and reaches a minimum in generation 19, with mucus thickness $3.22\mu m$. Mucus thickness naturally varies between individuals, due to genetic or lifestyle factors. Thicker mucus layers allow a larger space for virions to get advected to other parts of the RT, thereby causing infection elsewhere or possibly clearing virions more quickly.

2.3.2.2 Advection velocity of the mucus layer

The advection of the mucus layer v_{adv} is the speed at which the mucus layer travels resulting from the aggregate beating of the cilia populating the PCL and protruding from the cell layer. This advection dominates over diffusion in the nasal passage and upper generations of the RT. Altering advection velocity strength can be thought of as changing the length of the vectors embedded inside the green rectangular region in Figure 2.6. Note that in the deepest part of the RT, i.e., generations 20-23, no advection occurs.

Similar to mucus thickness, estimates for the values of advection in each generation are given in Table 1.2. It starts off greatest in the nasal passage at $8.8\frac{mm}{min}$ and monotonically decreases to $0.001\frac{mm}{min}$ in generation 19. Mucus advection also may vary between individuals. Additionally, as an infection takes hold of cells, it may impair their ability to beat with other nearby cilia, causing a reduction in MCC.

2.3.2.3 Percentage of infectible cells

The percentage of infectible cells i_{perc} is the percentage of cells which can potentially become infected by a virion. The chance for infection varies markedly amongst different cell types, due to the presence or absence of certain types of cell receptors, and the density of those receptors.

In [7], a value of 50% was used. Based on the composition of the RT tissue a person has, this percentage may differ. A larger percentage of infectible cells may cause infections to take hold more easily and spread further, whereas a smaller percentage of infectible cells may offer greater protection from infection.

Changes to this parameter can be visualized as increasing or decreasing the number of cells colored blue (infectible) in Figure 1.2 A.

2.4: OUTPUT SPACE

The model originally developed in [7] determines the possible outputs which can be analyzed. See 1.4 for additional discussion. For the tracking of the progression of infection, the number of shed virions and number of infected cells are collected. The number of virions which exit a generation (also called the flux) is recorded.

2.5: ORIGINAL MODEL MODIFICATIONS

The original code in [7] developed had to undergo a few modifications to be deployed on the Longleaf platform. Some of these were simple to implement, like avoiding duplication of the source code in many directories by using a common server directory and changing some of the packages to be server compatible. Some of the most involved modifications are discussed in this section.

2.5.1 State data extraction

In the original version of the code [7], the entire state of the simulation would be saved to hard disk. The advantage of this was all the data generated by the simulation would persist to be utilized later. In the event additional information needed to be derived from the simulations, this would be possible. The disadvantage was these files tended to be quite large, consuming at minimum GBs of space.

When the number of simulations was sufficiently small, we wrote a collection of scripts to efficiently extract any desired information from the simulations themselves. This was useful in the cases where we wished to keep track of the flux of virions out of a generation prior to it being incorporated into the base model output of [7].

2.5.2 Memory constraints

As already mentioned, a major drawback of saving the state of the simulations was the vast amount of memory space consumed. A single run of the sensitivity analysis platform with saving state enabled can easily overwhelm the 50 GB partition on Longleaf. Exceeding this limit prohibits users from saving additional files, and so further attempts to save simulation output are lost entirely.

To address this issue, the source code in [7] was modified to save only the outputs of interest, typically the number of virions sheds and number of infected cells. Naturally, the drawback of this is if additional

information is desired about a simulation, then it requires a complete rerun. Rerunning a simulation can cost days of computation time, however, this approach allowed many more computations to beat deployed.

2.5.3 New functions

Removing the state saving functionality in the base code created the need to write new functions to extract and format the output prior to the end of execution. These routines include functions to calculate the number of shed virions, the number of infected cells and the virion flux over time. These function sample at an equal number of locations specified by the user in the input file. Another function implemented was one to calculate the maximum number of virions which infect a cell over the lifetime of the simulation. This function's output was used to determine if the number of reinfections over the length of the simulation was biologically feasible.

2.6: MANY-INFECTION STATISTICS

Every parameter combination chosen requires many realizations for good statistics, i.e., enough so that the distribution of outcomes is sufficiently characterized. For all the parameter sweeps performed to date, around ≈ 100 realizations, i.e., $n_{realization} = 100$ are required to obtain good statistics. Additionally, in the case of number of shed virions, and number of infected cells, both appear to be approximately Gaussian. This may make these distributions amenable to the metamodeling technique of Kriging.

One test to determine the amount of realizations necessary is shown in 2.7 and 2.8. Both these distributions are generated from the parameter combination $T = 24\ hrs, p_{infect} = 0.3, t_{latency} = 6\ hrs, r_{shed} = 4000, v_{adv} = 146.7\frac{\mu m}{s}, h_{mucus} = 10\mu m$. Both Figures 2.7 and 2.8 compare the distribution by varying the number of realizations from 100 to 1000. The means and variances for both of these are approximately equal, as indicated by the small relative errors. A z-test was performed as well, suggesting both of these distributions from the same underlying random variable.

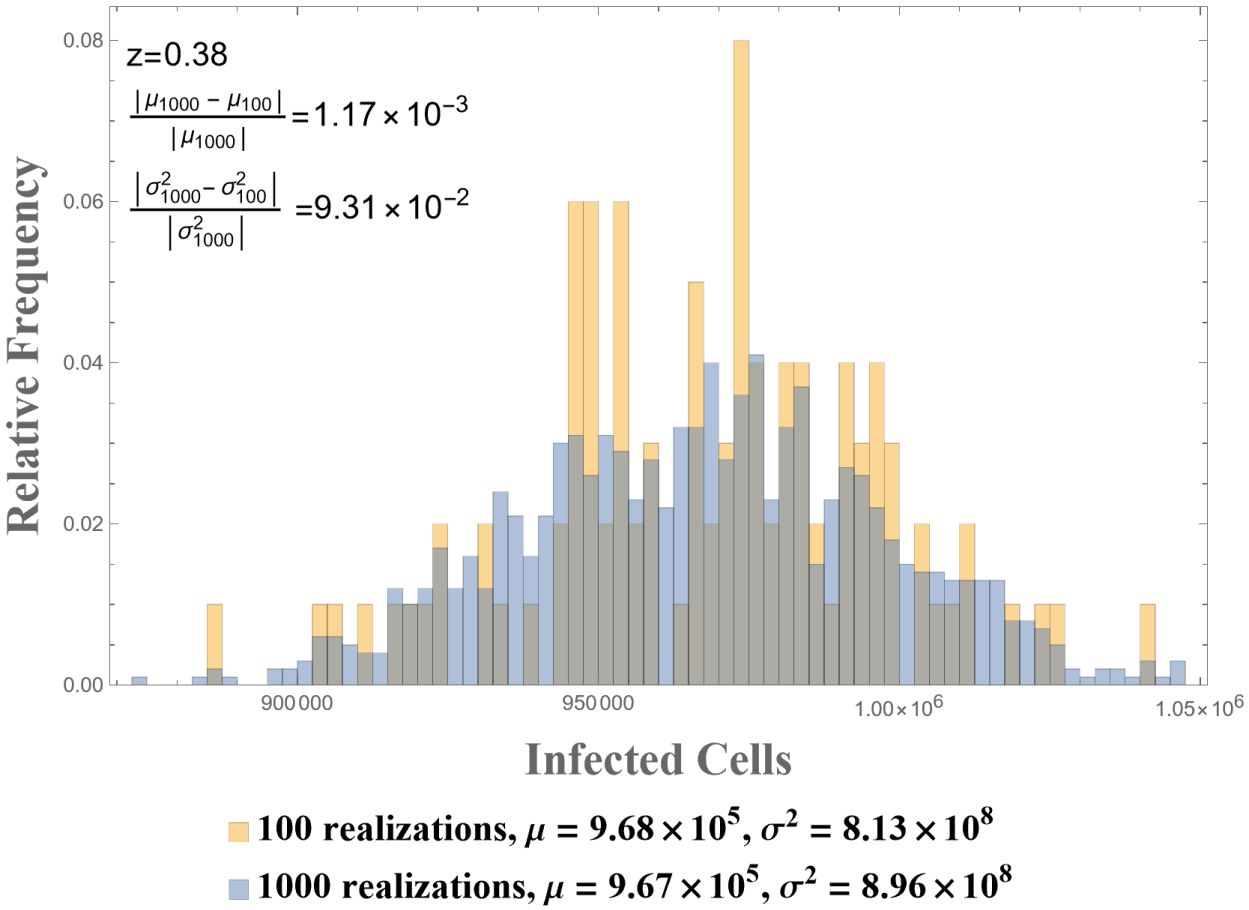

Figure 2.7: Sample distribution of infected cells

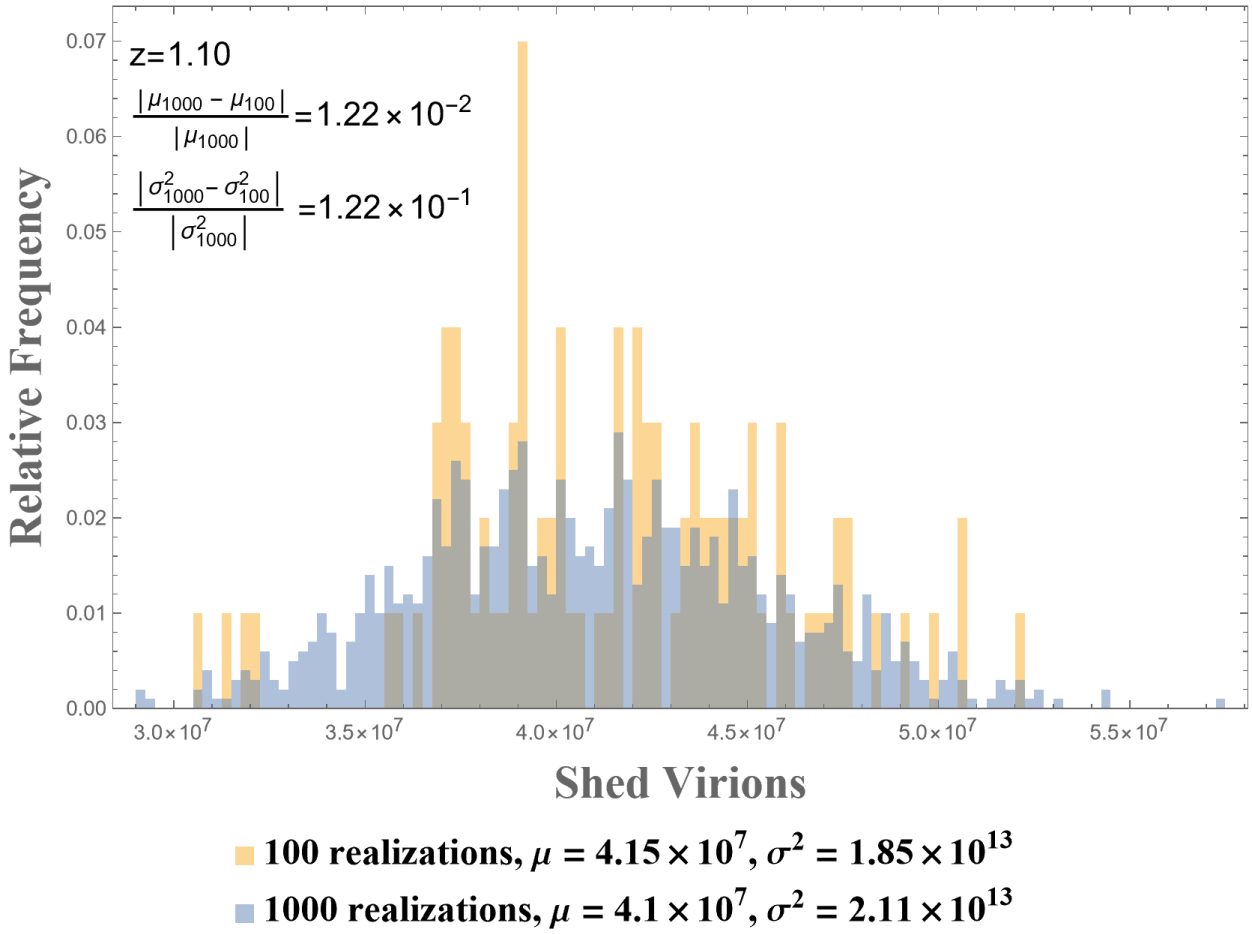

Figure 2.8: Sample distribution of shed virions

When considering the efficency of sampling, it is worthwhile to consider the question of what is the minimum number of realizations sufficient to characterize an outcome? If a distribution can be characterized with fewer samples, less computational resources are required, and results can be obtained faster. To explore this question, random subsamples can be taken from a specified parameter combination, and the relative error incurred from these subsamples can be computed.

One such subsampling process is shown in 2.9. The various curves correspond to different proportions of the original sample used as a subsample. For example, 0.9 indicates 90% of the original sample was used, that is, 90 randomly selected realizations of the original 100. The main point of this plot is that 80 realizations maintains a relative error of less than 1% from the mean off the distribution.

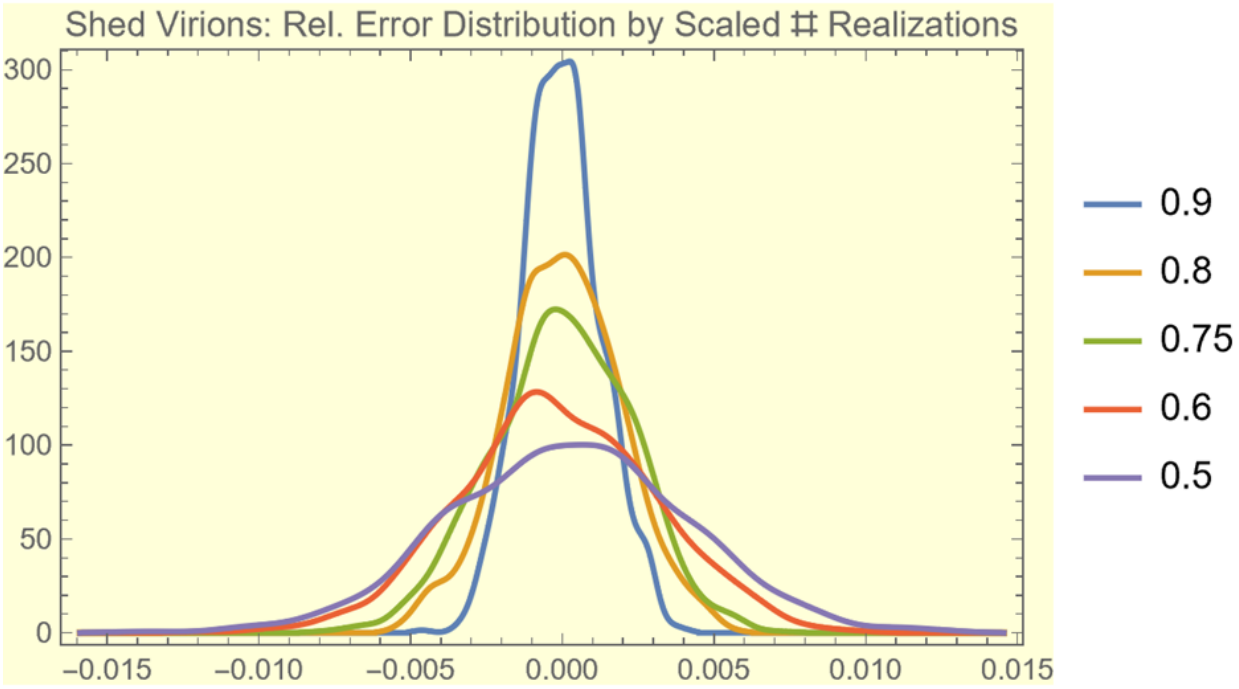

Figure 2.9: Subsampling to obtain minimum required number of realizations

An important caveat to bear in mind is the number of realizations required to obtain good statistics may depend upon the parameter values themselves. All parameter space regions explored to date have conformed to $n_{realization} = 100$, however, it is certainly possible more would be necessary for other yet to be investigated parameter regimes.

2.7: EFFICIENT COMPUTATION

A single run of the model for certain values of the input parameters for $t_{latency} = 2$ and $r_{shed} = 2000$ requires around 60 GB of memory and 4 days of CPU time. Computing many simulations in serial would take a prohibitively long amount of time. We use UNC's Longleaf cluster to simultaneously run many simulations.

The UNC Longleaf computing cluster is a shared resource amongst many users. It has a scheduler called SLURM [50] from which users request resources to be applied to submitted jobs. On Longleaf, good CPU time and memory requests must be made for a job to complete successfully. When too few resources are requested, simulations will not finish. When too many resources are requested, job throughput is reduced, slowing the completion of simulations, and also diminishing the user's associated fair share score. The fair

share score determines what priority is given to submitted jobs. Jobs with low priority scores stay in the scheduler queue for longer and drive up the time to complete jobs.

2.7.1 Resource requirements

One challenge of making appropriate resource requests for jobs is the simulations dependence on the parameter values themselves. The effect of a given parameter on the model is not known a priori, otherwise it would not be necessary to conduct a sensitivity analysis in the first place.

Another challenge of making appropriate resource requests for jobs is the stochastic nature of the model. The virions undergo diffusion, a fundamentally random process (See Sec. 1.2). This randomness ensures the simulations themselves are random, which in turn means the CPU time and memory required are random.

A demonstration of these two challenges is shown in Figure 2.10 and Figure 2.11. Both of these figures are for the same parameter combination of $T = 48hrs$, $p_{infect} = 0.2$, $t_{latency} = 2hrs$, $v_{adv} = 16\% \times 146\frac{\mu m}{s}$, $h_{mucus} = 48hrs$, $i_{perc} = 50\%$,

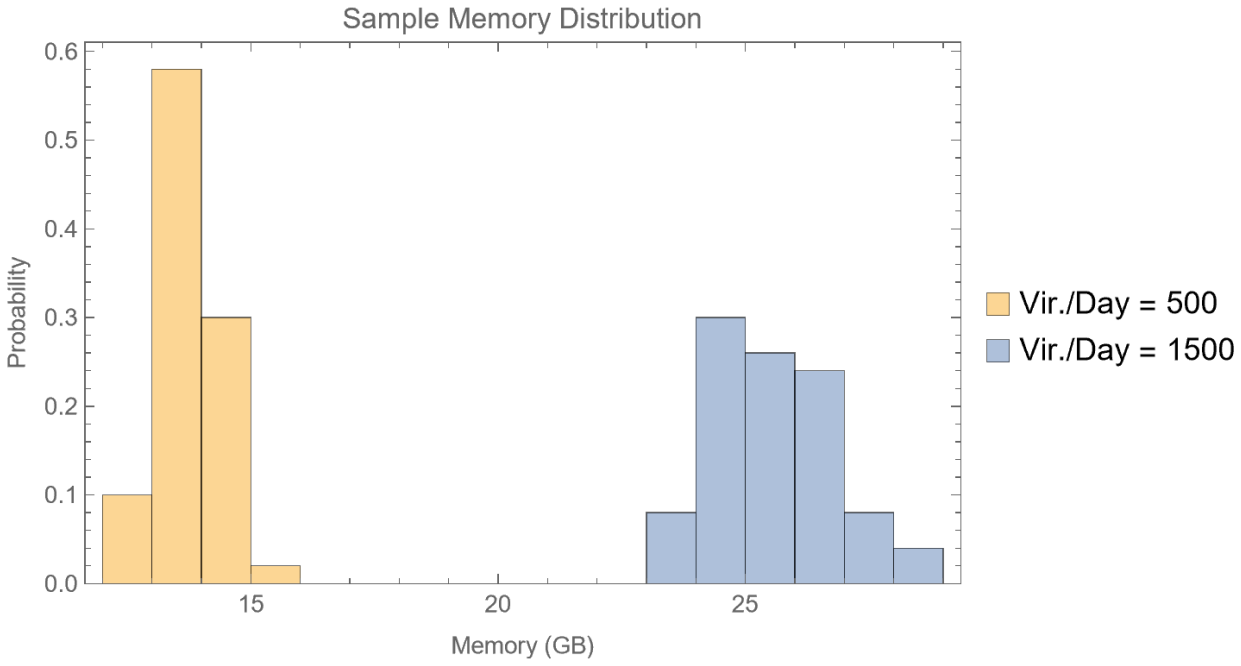

Figure 2.10: Required memory distribution for a specific parameter combination.

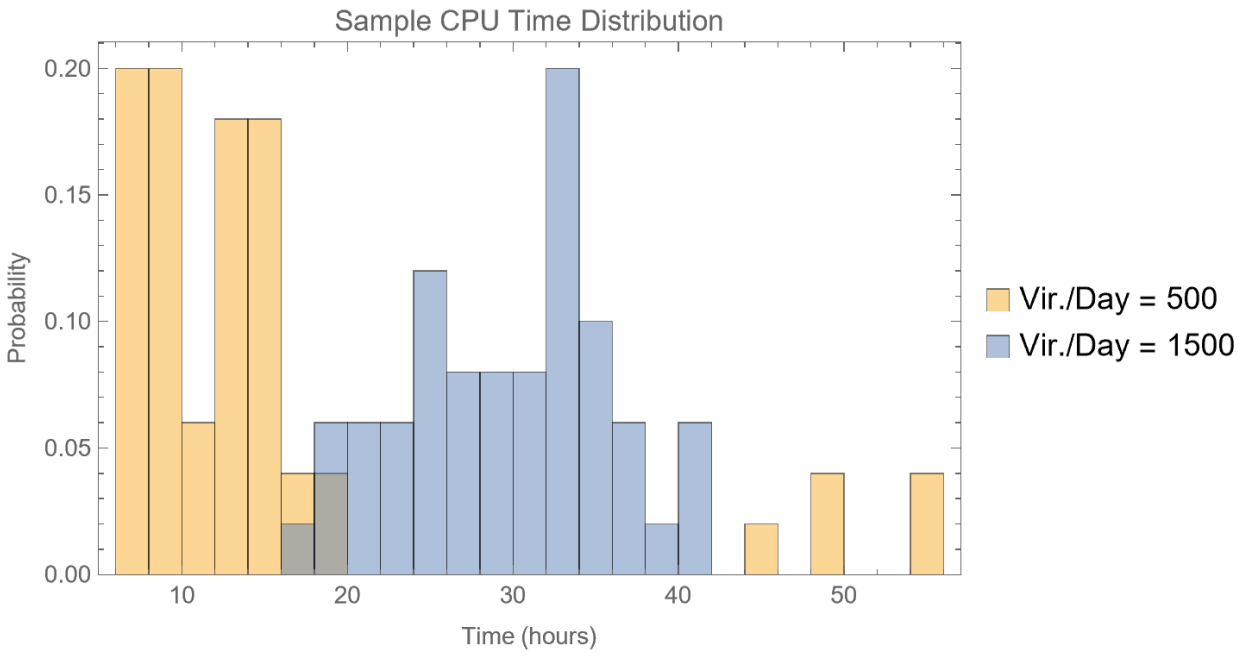

Figure 2.11: Required CPU time distribution for a specific parameter combination.

Both the memory and CPU time show a clear dependence on the parameters, while forming an approximately Gaussian distribution of their own. In the case of the memory usage, for $r_{shed} = 500$, the average memory value is around 14 GB, whereas for $r_{shed} = 1500$, the average memory value is around 26 GB. The variance of in the latter distribution is greater than the former. In the case of CPU time, the time required for r_{shed} has an average of 15 hours of CPU time, and r_{shed} has an average of 30 hours of CPU time. The variance again is greater in the latter distribution.

An interesting note is the presence of several outliers in the CPU time required. A few of the $r_{shed} = 500$ simulations required even more CPU time than those associated with $r_{shed} = 1500$. This suggests resource values must be chosen carefully to ensure all the simulations are able to be completed.

2.7.2 Parameter space exploration costs

When discretization of the parameter sweep subspace R in Eq. 2.2 is sufficiently refined, the computational resources to complete the parameter sweep may be prohibitively large. Assuming a fixed simulation time T, and $n_{realization}$ realizations for each parameter combination, a general parameter sweep subspace R in n_p dimensions discretized at $\{m_1, ..., m_{n_p}\}$ points along each respective dimension needs

$$Memory = m_1 \times ... \times m_{n_p} \times n_{realization} \times \bar{m} \tag{2.6}$$

$$CPU\ Time = m_1 \times ... \times m_{n_p} \times n_{realization} \times \bar{t}_{CPU} \tag{2.7}$$

such that $\bar{t}_{CPU}, \bar{m}$ are the average CPU time and memory it takes per simulation.

Supposing the estimated average CPU time per simulation is $\bar{t}_{CPU} = 1hrs$, and the average memory usage is $5GB$, and each of the $n_p = 6$ dimensions are sampled at $m_i = 3$ locations, the resources required are

$$Memory = 3^6 \times 100 \times 5\ GB = 364500\ GB \approx 365\ TB \tag{2.8}$$

$$CPU\ Time = 3^6 \times 100 \times 1\ hrs = 72900\ hrs \approx 8.2\ yrs \tag{2.9}$$

To work around these costs, smaller dimensional subspaces can be chosen. Discretizing along two selected dimensions allows for greater refinement and better resource scaling. The effects of all the parameters can be compared in a binary fashion, and a hierarchy can be formulated ordering the input parameters from greatest to least impact on model outcomes.

2.8: CONCLUSION

This chapter covered some of the basics of sensitivity analysis and the specific implementation of the sensitivity analysis platform. An overview of how the platform functions, its inputs and outputs were described. The essenctial nature of performing many simulations to obtain reliable statistics with computational efficiency was also outlined. With this in place, the questions asked at the beginning of this chapter can be answered. To this end, this platform will be used in the next chapter to capture clinical and experimental outcomes for SARS-CoV-2.

CHAPTER 3: CLINICAL DATA VALIDATION

We now turn to answering the question: which mechanisms of infection and replication might explain exponential rises in nasal titers from the waves of dominant variants?

3.1: VARIANT INDUCED NASAL TITER CHANGES

In [7], a mechanistic, within-host, respiratory tract model was developed and applied to predict immediate (hours to days) outcomes of infection and viral load from inhaled SARS-CoV-2 exposures in the respiratory tract. The model assumes all adaptive immune system protection is absent on these timescales, consistent with exposure to a novel viral exposure, and results were presented based on the best-available information for the 2020 alpha variant. Here we use the model to explore three likely mechanistic effects of the numerous SARS-CoV-2 mutations documented for the delta and omicron variants [16, 29, 43, 3, 11]: (1) cell infectivity (probability to bind and successfully infect per virion-cell encounter per second); (2) latency time (from cell binding and uptake of a virion to cellular assembly, replication, and shedding of infectious daughter virions); and (3) efficiency of virion replication (number of infectious virions shed per day) [28]. These three mechanistic effects represent segments of the stages of viral infection (see Figure 3.1), and each is represented in the model by a specific parameter: (1) cell infectivity is modeled by p_{infect}, (2) latency time (often called the eclipse phase) by $t_{latency}$, and (3) efficiency of virion replication (or rate of replicated infectious virions being shed) by r_{shed}. Our aim is to explore whether, and if so which, variations in the best-known parameter values of the alpha variant can account for the dramatic rises in nasal infectious viral load of the delta variant ($\approx$ 3 orders of magnitude) [16, 29, 43, 3] and omicron variant (another $\approx$ 1-2 orders of magnitude [11]).

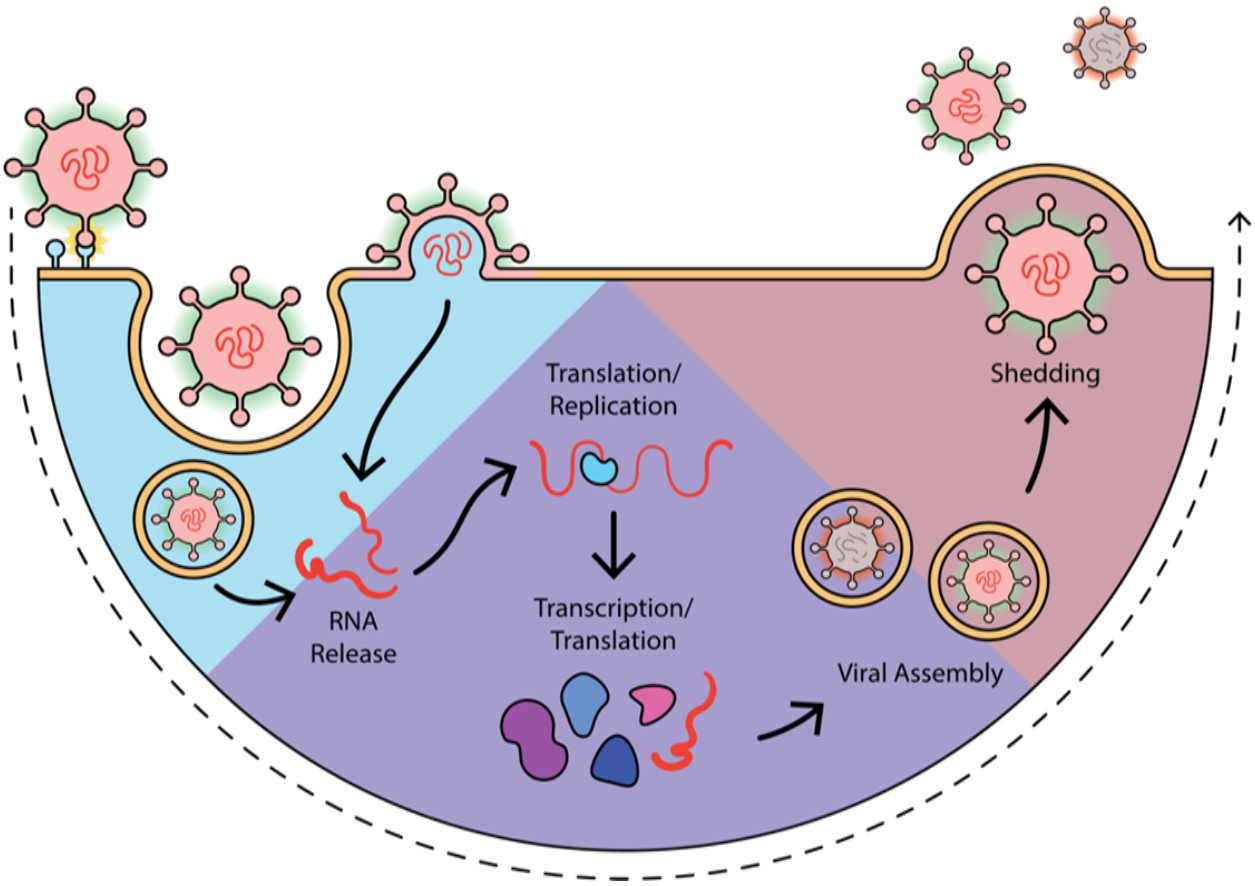

Figure 3.1: SARS-CoV-2 infection schematic. Stage 1 (upper left, blue triangle): a successful encounter-binding-infection event by an infectious (green corona) virion with an infectable cell, parametrized in the model by p_{infect} (probability of a successful infection event per encounter per second). Stage 2 (lower middle, purple triangle): the latency or eclipse phase including cellular processes of RNA translation, transcription, and assembly of infectious and non-infectious (orange corona) RNA copies, parametrized in the model by $t_{latency}$ (the time between the first successful infection event and onset of virion shedding); and Stage 3 (upper right, COLOR triangle): the shedding phase, parametrized in the model by r_{shed}, the number of infectious virions shed per day per infected cell post eclipse phase

We focus this study on the nasal passage only (where there is strong testing data of the viral loads for alpha, delta, and omicron variants). The model outputs the viral load within the nasal passage, the flux into the nasopharynx [25, 15, 4, 2], and the number of infected nasal cells, in the 24-72 hours immediately following infection of a single nasal epithelial cell at the entrance of the nasal passage. Simulations are conducted at the entrance of the nasal passage (i.e., furthest from the nasopharynx). Chen et al. 2022 [7]

show the viral load and infected cell outcomes initiated at the entrance are statistically similar to results from a cell initially infected within the middle two quartiles of the nasal passage. We note that very few inhaled infectious aerosols suffice to generate a high-titer nasal infection, [7, 22, 24, 48]. Further, as shown in [7], each infected nasal cell produces a thin streak of infected cells and viruses in the mucus layer during the first 1-3 days because of strong nasal mucus advection toward the oropharynx and subsequently esophagus, in contrast to other within-host models of viral infection [14, 13, 44, 5]. As such, we assume no overlap in the small number of infectious streaks sufficient for a high titer nasal infection, not valid for larger inhaled exposures.

We model outcomes in viral load and infected cells across feasible ranges in the kinetic parameters for these 3 mechanistic effects, anchored by results for the alpha variant in Chen et al. 2022 [7]. We refer the reader to [7] and [2, 31, 4, 6, 9] for details of: (i) the physiological details of the nasal passage, including length and circumferential dimensions, thickness and clearance velocity of the mucosal layer, and thickness of the periciliary liquid (PCL) layer; (ii) the percentage of SARS-CoV-2 infectable epithelial cells (predominantly ciliated cells) in the nasal passage; and (iii) the computational model for tracking the spread of viral load and infected cells.

Simulating a single infection is now expanded to simulating many infections with differing parameters. This enables the exploration of various factors determining infection outcomes. In practice, the parameters characterizing a viral infection in the respiratory track may be difficult to measure precisely in experimental settings, due to technical or ethical reasons. The human respiratory track also shows considerable variation from individual to individual. Additionally, viruses often undergo mutations profoundly affecting their ability to spread and infect. Lastly, the infections themselves have a stochastic component, and the ability to simulation many infections to determine the statistics of many infections is desirable.

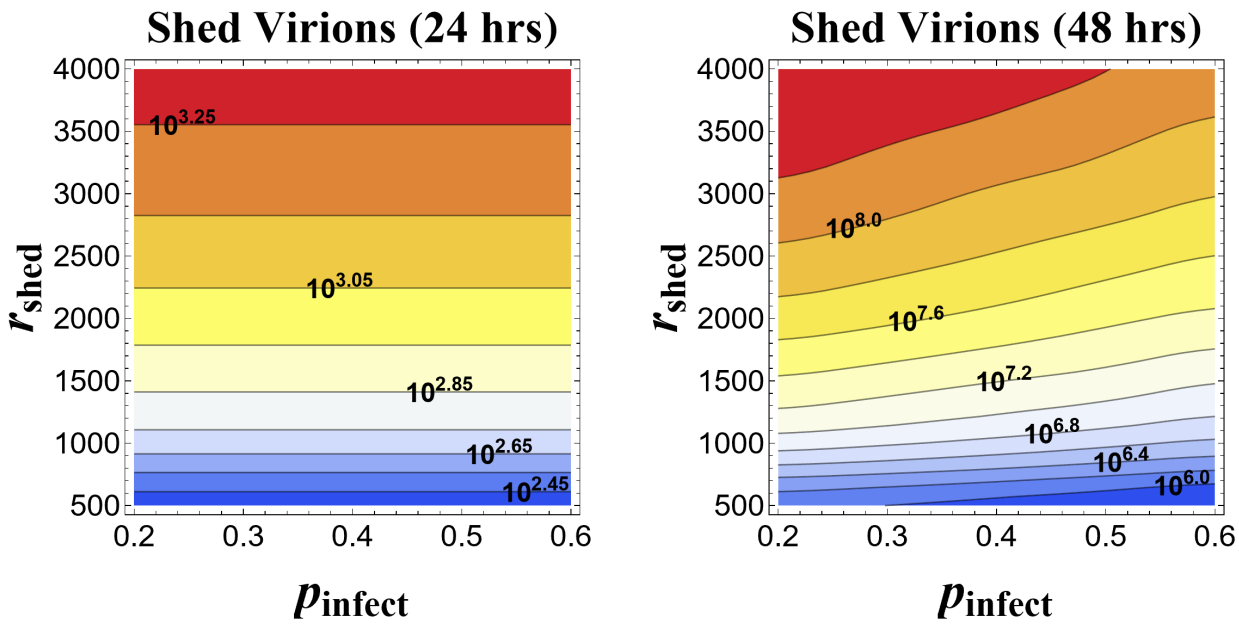

Figure 3.2: State diagrams depicting the level sets of total shed virions from one infected cell at the entry of the nasal passage at 24 hours (left) and 48 hours (right) post cell infection, over ranges of model parameters (p_{infect}, r_{shed}), for fixed latency time, $t_{latency} = 12$ hours, of the alpha variant. p_{infect} is cell infection probability per virion encounter per second, r_{shed} is virions shed per day by infected cells once the latency time expires. The underlying data are mean shed virions over 100 realizations at each fixed (p_{infect}, r_{shed}). Solid curves pass through level sets, with some orders-of-magnitude depicted; all results follow from a uniform ratio between successive level sets, $10^{0.10}$ at 24 hours and $10^{0.20}$ at 48 hours.

Figure 3.2 shows the total infectious viral load at 24 and 48 hours post infection from a single infected cell in the upper nasal passage due to variations in (i) cell infection probability per encounter per second (p_{infect}) and (ii) shedding rate (#/day) of infectious virions from a single infected nasal cell (r_{shed}), for a fixed 12-hour latency time ($t_latency$ =12 hours) of the alpha variant. The ranges p_{infect} = 0.2 to 0.6 and r_{shed} = 500 to 4000 are chosen to extend the baseline values p_{infect} = 0.3 and r_{shed} = 2000 in Chen et al. [7]. Results for much lower p_{infect} are given in Sec. 3.2. Figures 3.2, 3.2 reveal robustness of viral load to cell infectability, p_{infect}, the model proxy for spike-receptor binding affinity even for much lower p_{infect}. Higher p_{infect} has a negligible effect at 24 hours post infection, and a slight reduction in nasal viral load at 48 hours. The widely reported spike mutations in delta and omicron leading to stronger ACE2 and TMPRSS2 receptor binding affinities [26, 17, 33, 23] are therefore not the source of dramatic rises in infectious nasal

titers! Figure 3.2 further reveals that increased shedding rate, e.g., from mutations that improve efficiency of infectious RNA assembly and replication, significantly raises viral load. One order-of-magnitude increase in infectious virions per day produces 2 orders-of-magnitude gain in shed virions after 2 days. Figures 3.2, 3.5 show: mutation-induced changes in cell infection probability p_{infect} and shedding rate r_{shed} cannot produce the 3-5 orders-of-magnitude rises in viral loads over the alpha variant.

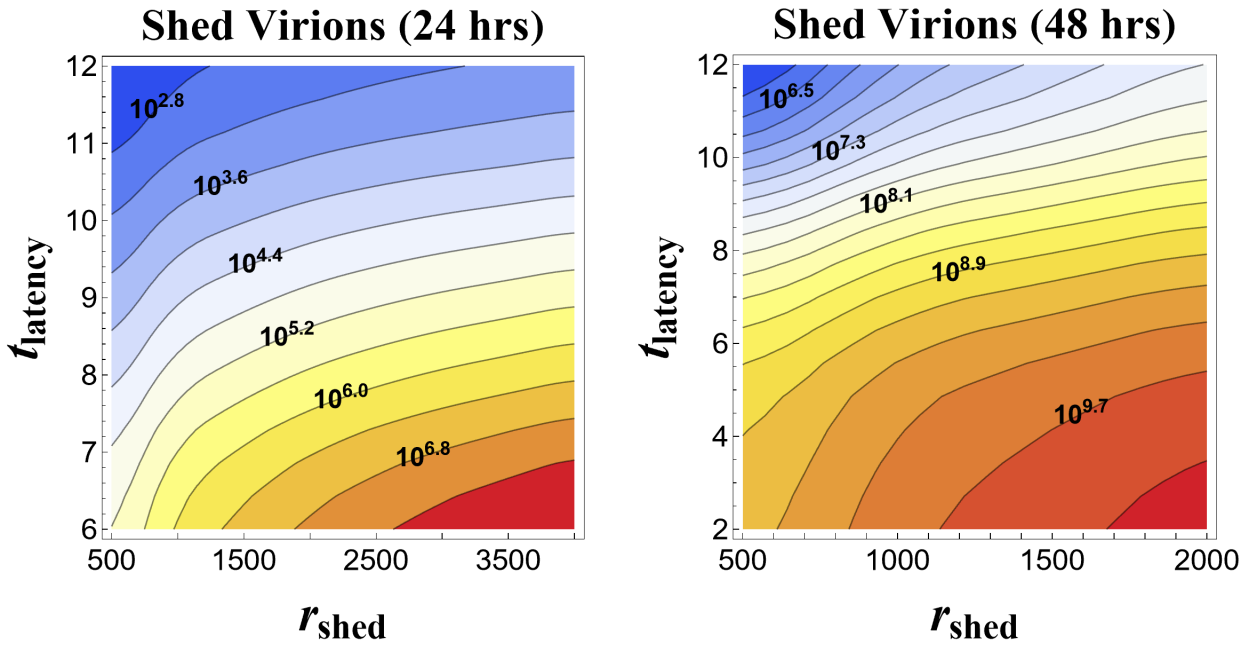

Figure 3.3: State diagrams depicting the level sets of total shed virions from one infected cell at the entry of the nasal passage at 24 hours (left) and 48 hours (right) post cell infection, over ranges of model parameters $(r_{shed}, t_{latency})$. r_{shed} is shedding rate (virions per day), $t_{latency}$ is latency time (hours). Note 24-hour and 48-hour results span different ranges of $(r_{shed}, t_{latency})$. The underlying data are mean shed virions over 100 realizations at each fixed $(r_{shed}, t_{latency})$, with fixed p_{infect}=0.2. Solid curves pass through level sets, with some orders-of-magnitude listed; remaining level sets shown follow a uniform ratio between successive level sets, $10^{0.4}$ at 24 hours and $10^{0.2}$ at 48 hours.

Figure 3.3 reveals dramatic orders-of-magnitude impact of reduced latency times on viral load at 24 and 48 hours post infection from a single cell at the entrance of the nasal passage. Depending on viral shedding rate, e.g., low vs. high, fixed or varying, rises in total shed virions are as high as 4+ orders of magnitude at 24 hours and 3+ orders of magnitude at 48 hours! Figure 3.4 below provides insight into the dynamics

underlying Figure 3.3, showing growth in shed virions (solid) and infected cells (dashed) over 48 hours post single-cell infection for 2, 6, 12-hour latency times and 500, 1000, 1500, 2000 shed virions/day.

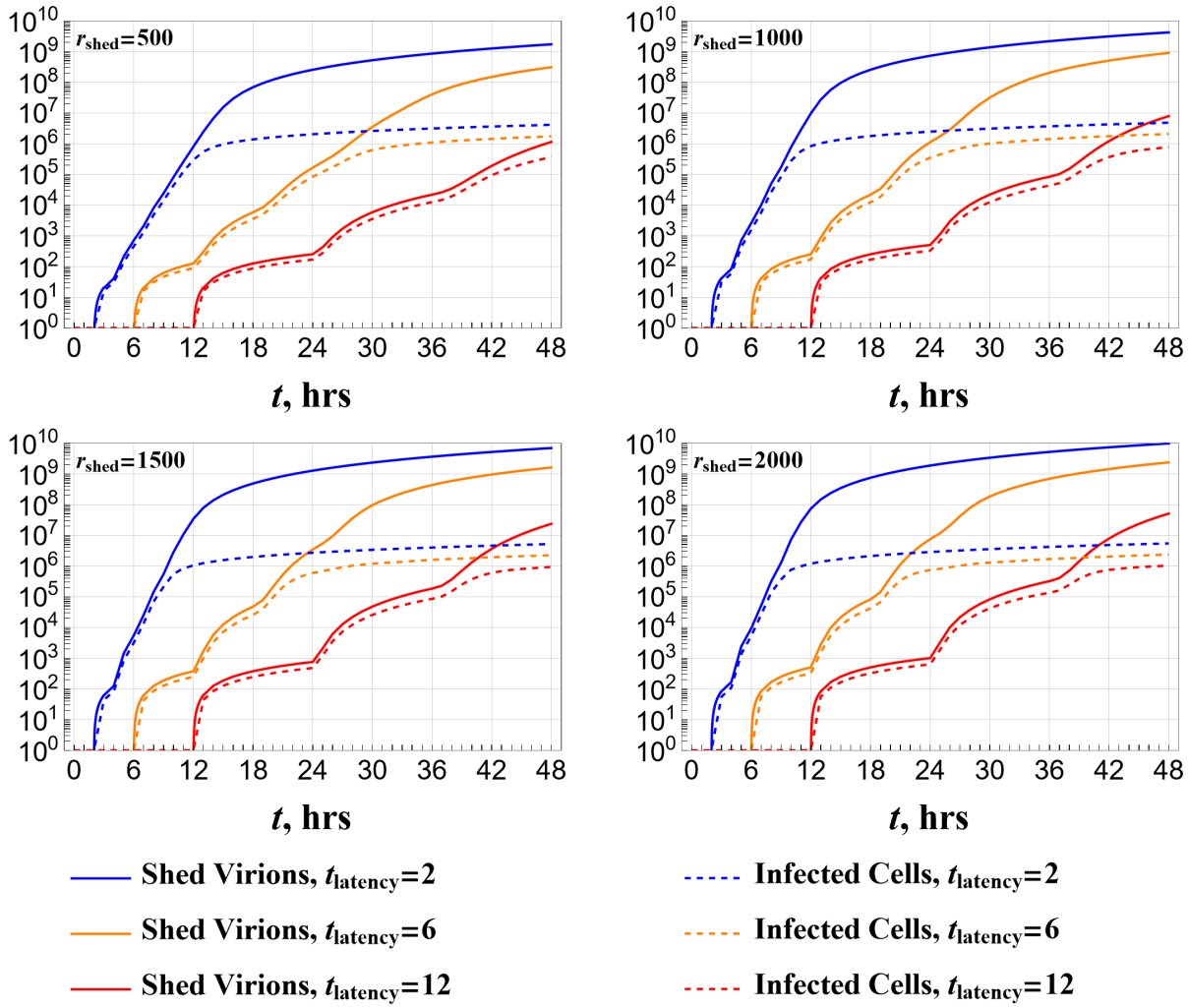

Figure 3.4: The dynamic evolution of infectious shed viral load (solid curves) and infected cells (dashed curves) over two days from a single infected cell in the entrance of the nasal passage, starting from the moment of infection, for latency times $t_{latency}$=12 (red), 6 (orange), 2 (blue) hours, for shedding rates r_{shed} = 500, 1000, 1500, 2000 virions/day, with p_{infect}=0.2 in all simulations.

Figure 3.4 reveals "episodic surges" in shed virions from each ancestral generation of daughter virions. The first surge of shed virions is launched as the latency time (2, 6, 12 hours) of the initial infected cell expires. Those shed virions diffuse, advect with mucus clearance velocity while in the mucus layer, either exit the nasal passage without infecting new cells or encounter and infect new cells at a stochastic sequence

of times and primarily downstream locations in the nasal passage. Each newly infected cell goes through the latency phase before the second surge emerges one latency time later, from the first, second, ..., hundredth, ... infected cell, and so on with the third, then following surges. Mucus advection succeeds in transporting the "streak" of shed virions from the nasal passage into the oropharynx [7]. Infections that initiate farther into the nasal passage each generate streaks of shed virions that grow in the same manner and exit the nasal passage sooner.

Figures 3.3, 3.4 reveal the profound effect of shorter latency times: a faster series of episodic surges of shed virions and therefore greater and faster exponential rise in total shed virions, consistent with clinical data of the delta and omicron variants [3, 40, 11, 16, 49].

Finally, we indicate how to link shed viral load modeling estimates with clinical viral titers. Viral titers count total RNA copies per volume in a nasal swab of mucus and PCL, with no ability to distinguish viable (infectious) virions from unviable RNA copies. Nasal tests are taken at some unknown time relative to nasal infection. As shown here and in [7], shed viral loads rapidly grow by orders of magnitude in the 1-3 days following infection of a single nasal cell. We compute the PCL-mucus volume that contains all shed virions within the nasal passage (the streaks described in [7]), providing an estimate of the viral titer of each streak of shed virions. We can sum the titers from multiple infected cells, locations within the nasal passage, and timelines, and estimate what portion of nasal fluid volume a swab would produce. Note that clinical swabs access far into the nasal passage, a strategy that modeling confirms. Results in [7] for best-known model parameters of the alpha variant easily produce $\approx 10^6$ infectious RNA copies per mL over the 48 hours post infection from less than 10 infected cells in the nasal passage. Further, results in Figures 3.3, 3.4 above easily produce 1000-fold higher titers by reducing latency time from 12 to 6 hours, and another 50-100 times higher viral titer if latency time is reduced from 6 to 2 hours.

3.2: LOW INFECTION PROBABILITY

Our collaborators working on the biology suggested the initial range of infection probabilities investigated in 3.2 were limited. In biological systems, these types of probabilities can range over different orders of magnitude. To investigate the impact this would have on infection, we looked at simulations with probability of infection per encounter-second spanning 4 orders of magnitude, specifically $p_{infect} = 3 \times \{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$. Other parameters were fixed as $r_{shed} = 1000$ and the other baseline values of [7].

These were compared with the effect from varying latency times, $t_{latency} = 2, 6, 12$, which was the parameter with the greatest affect on the model outcomes from 3.1. The results are shown in Figure 3.5.

Overall, the 4 panels shown indicate p_{infect} has little contribution to The infection. An 100-fold increase in its value from the upper-right panel to the lower-right panel shows less than an order of magnitude of change in the final shed virions count. The bottom two panels, corresponding to $p_{infect} = 0.03, 0.3$ respectively are difficult to distinguish, implying reducing the infection probability by an order of magnitude has virtually no impact on the development of an infection.

The upper-left panel, corresponding to the smallest p_{infect} simulated, displays an almost constant separation between shed virions and infected cell curves of differing $t_{latency}$ values. This occurs because the value is small to the point the shed virions have difficulty successfully attaching to another cell prior to being cleared from the nasal passage. Without successful attachment, very few subsequent cell infections occur, and the "episodic surges" seen in the other three panels fail to materialize. A signifant fraction of the shed virions in all three of these curves are shed from the initially infected cell. The only significantly difference between the three curves is the time at which the initial cell began shedding, i.e., their respective latency times.

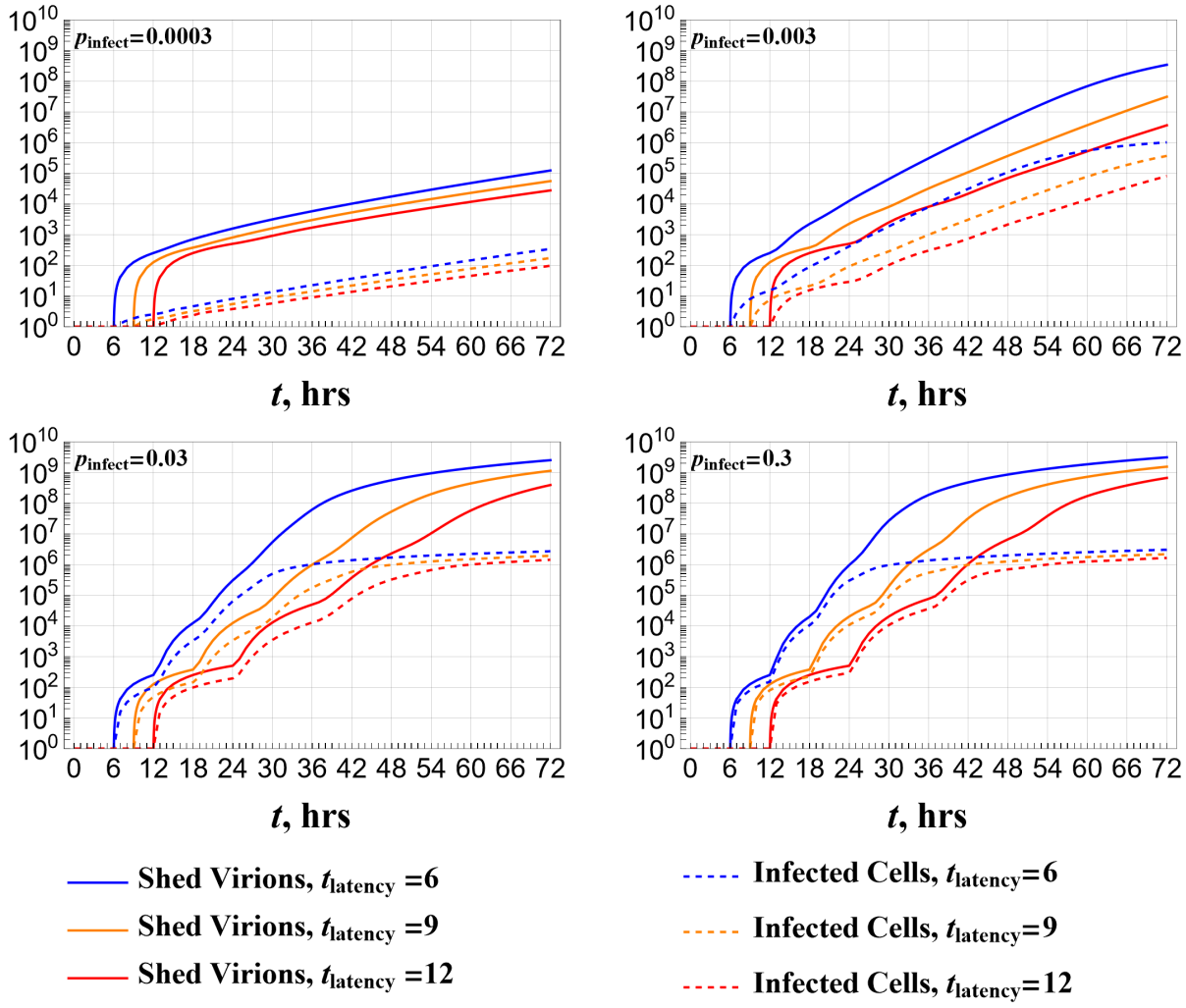

Figure 3.5: Dynamic evoluation of cumulative infectious viral load (solid) and total infected cells (dashed) for decades of p_{infect}

3.3: REDUCED MCC

In [21], it was observed that SARS-CoV-2 infections reduced MCC clearance by 67%. We then wished to simulate the effect reduced MCC, that is, reducing v_{adv} would have on infection outcomes. The mucus advection velocity was chosen as $v_{adv} = 147\frac{\mu m}{s} \times \{16.7\%, 33.3\%, 50\%, 66.7\%, 83.3\%, 100\%\}$ with $147\frac{\mu m}{s}$ being the baseline value from [7], and compared with the dominant input parameter in Sec. 3.1, latency time, with values chosen at $t_{latency} = 6, 9, 12$ in this case.

The shed virion state diagram is shown in the left side of Fig. 3.6. The contour lines on the shed virions (left) plot are all roughly vertical, and the gradient with respect to latency time dominates the gradient of

advection. Over the region depicted, the maximum function change along the advection axis is $\approx 10^{0.2}$, when transitioning from 16.7% of baseline to 50% of baseline. In contrast, traversing the latency time axis shows a function change of $\approx 10^{1.9}$. A possible explanation for the non-monotonic nature of the contour lines is when the advection is small, the virions do not end up traveling as far, and cannot infect cells as efficiently to create a more severe infection. Around 50% of the baseline value in [7] slows clearance, so virions stay inside the nasal passage for longer and cause a worse infection. When advection is large, clearance begins to take effect, and virions are leaving the nasal passage before they are able to cause more severe infections.

The infected cell state diagram is shown in the right side of Fig. 3.6. This diagram shows a more varied regime of behavior. The greatest or least cell infection occurs where the advection velocity is the smallest. The gradients are dominated here again by the latency time contribution.

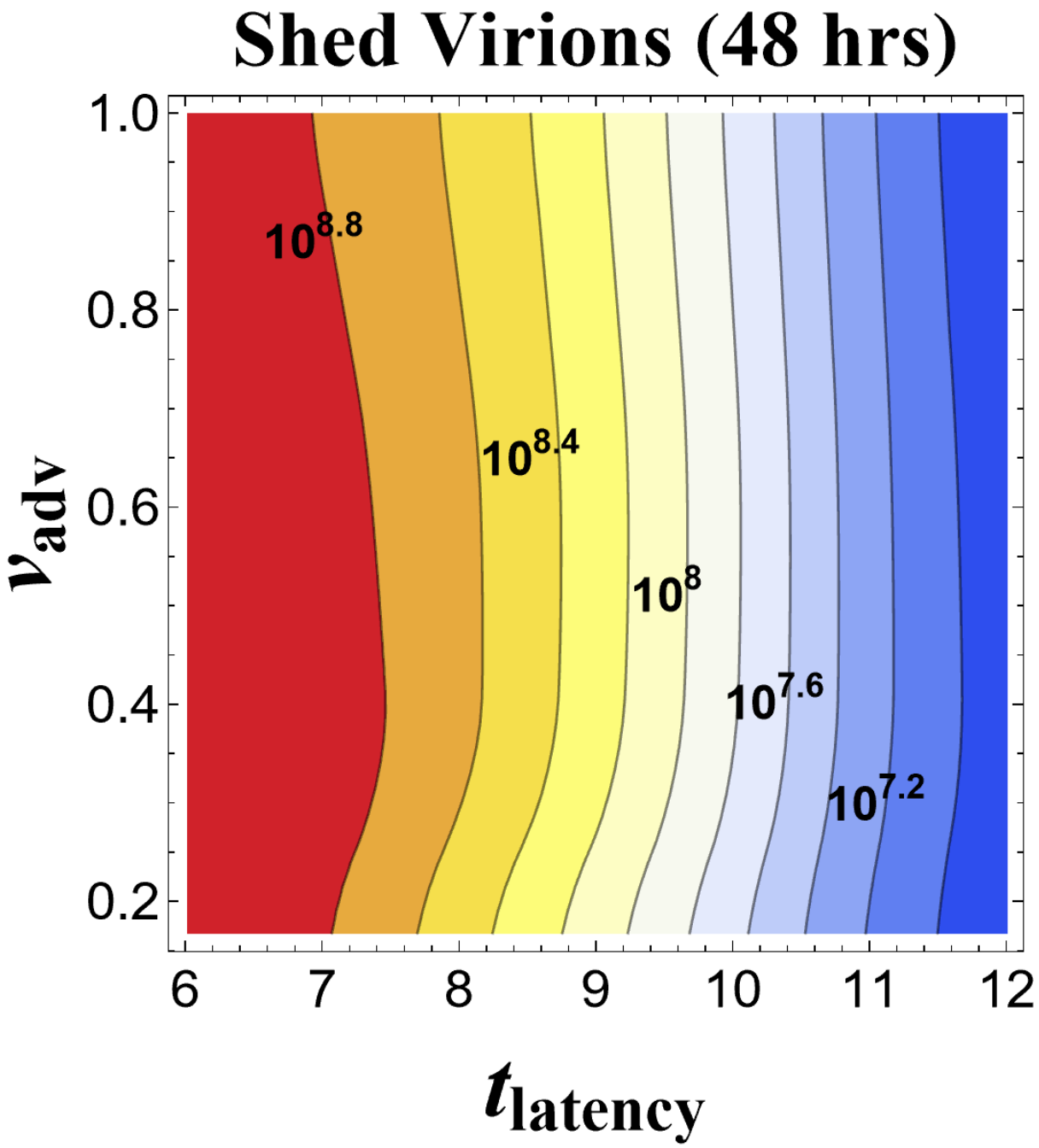

Figure 3.6: State diagrams for low latency time $t_{latency}$ and v_{adv}

The very small impact of a lower v_{adv} value is shown in 3.7. Each panels shows a different fraction of the baseline mucus advection velocity, however, the curves almost identical in all four panes over the lifetime of the simualtions. The infection outcomes are therefore very robust to changes in the mucus advection.

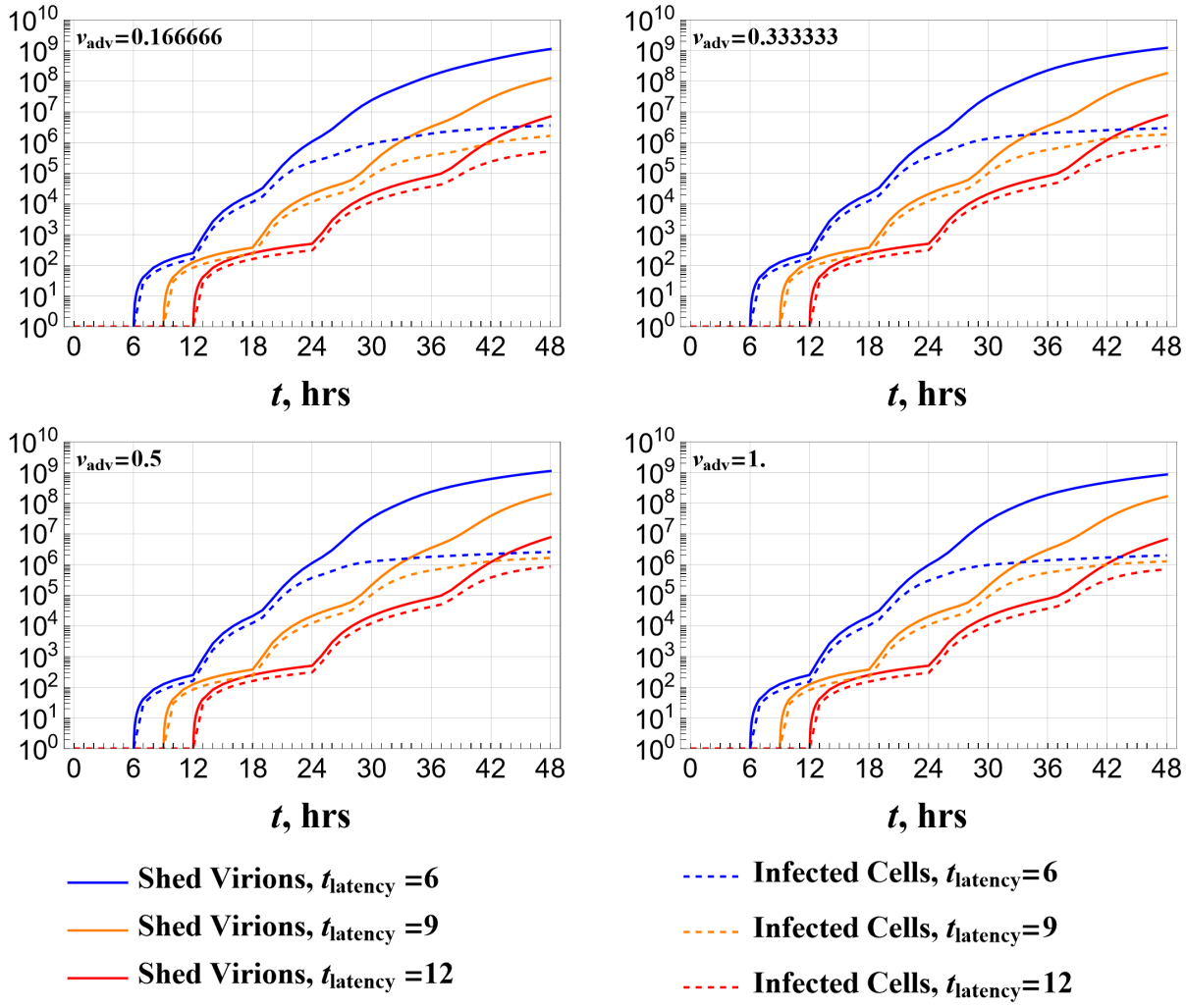

Figure 3.7: Dynamic evoluation of cumulative infectious viral load (solid) and total infected cells (dashed) for a range of mucus velocities, v_{adv}

CHAPTER 4: CONCLUSION

4.1: SENSITIVITY ANALYSIS

A platform was developed to explore the sensitivity of biologically relevant parameters in the human nasal passage in the case of a SARS-CoV-2 infection. Of the parameters considered, modeling clearly identifies latency time $t_{latency}$ as the dominant effect in determining infection severity, followed by the shedding rate r_{shed}. Parameters identified as having having the least impact are probability of infection per encounter p_{infect}, mucus thickness h_{mucus}, advection velocity v_{adv}, and percentage of infectable cells, i_{perc}.

This and future work investigating the sensitivity of these parameters may inform us of future possible variants and viable mechanisms for treatment of these variants.

4.2: RESULTS OBTAINED

The SARS-CoV-2 pandemic has been marked by waves of dramatic rises in viral titers from nasal samples as new dominant variants have emerged. In parallel, numerous mutations of four structural proteins have been sequenced for each of the variants. However, aside from stronger spike-receptor binding affinities of the delta and omicron variants, plus easier pathways for omicron to enter cells, there is very little mechanistic understanding of how mutations have altered infection and replication functions. Employing a mechanistic model [7] of respiratory exposure and infection by SARS-CoV-2, we tested hypotheses for mutation-altered virus-cell functions that might explain clinical outcomes from the alpha, delta, and omicron variants. Simulations presented here reveal three insights. 1. The highly cited stronger spike binding affinities of delta and omicron to cell receptors are not the cause of higher nasal viral titers; in fact, stronger binding, leading to higher cell infectability, slightly lowers viral load. 2. More efficient replication of infectious progeny, e.g., an increase from 500 to 4000 infectious RNA copies per day, can accelerate viral load up to a factor of 100. 3. Reducing the latency phase – from initial cell infection to shedding of daughter viruses – produces the most dramatic acceleration in viral load and by itself is capable of the dramatic rises in nasal viral titers from alpha to delta to omicron. We find that a reduced latency time from 12 hours (best estimate

for the alpha variant) to 6 hours produces 1000-fold rise in nasal viral titer within the first 24-48 hours post infection; further reduction to 2 hours produces an additional 50-100-fold rise. Furthermore, minimal inhaled exposures are sufficient to generate high-titer nasal infections, consistent with results in [22, 24].

These modeling predictions compel experimental verification that the period of latency, or eclipse phase, of infected cells has shortened comparable orders (12 to 6 to 2 hours), and to identify whether shedding rates of infectious progeny have increased (e.g., 500 to 4000 per day) from the alpha to delta to omicron variants. The impact of immune protection in the nasal passage from vaccination or previous infection, not considered in this model in the immediate 24-48 hours post infection, remains to be clarified, especially in light of rising infections from the current omicron sub-variants. Immune protection models [28] will be essential to guide intervention strategies since control of infection in the nasal passage is critical to prevent deep lung infections [7]. We also dispel the recent suggestion from [21] that reduced MCC due to nasal infection is not responsible for significantly higher titers.

4.3: FUTURE WORK

Extensions of this work are numerous – we have not "solved COVID-19". The parameter space P in Sec. 2.3 warrants continued exploration, as many of the parameters to date have yet to be measured, or if so, are not reported. Beyond this fact, viruses are known to mutate, and relevant parameter values may change over time. P itself could be extended to other parameters to be investigated. One avenue not pursed is infant and child physiology. Modeling changes to capture a longer simulation time, say a time scale in which the immune system can respond, is also desired.

Various well-known mechanisms in the immune system and their impact on infections could be explored. The affect of neutralizing and anchoring antibodies is already available in the infection model code, but a sensitivity analysis is yet to be completed. Preliminary results suggest a complex interaction with variable MCC and the other parameters, but more work is needed in this area. To cite a final example, the infection model currently does not incorporate the presence of interferons, proteins which signal to nearby cells a viral infection is taking place. A recent submission by Andres Aristotelous, Alex Chen, and Greg Forest to J. Theoretical Biology explores interferon-induced adaptive immunity in the alveolar region.

The release of virions into the RT is currently modeled as a single virion being released per unit time from just outside a single infected cell. However, it is known that SARS-CoV-2 forms syncytia (fused cell

strctures) out of several nearby cells, and it is these structures which then release daughter virions back into the RT. The release of daughter virions also does not take place one by one, instead releasing in plumes. These plumes have some spatial distribution which could be incorporated into the model.

Other extensions include continued clinical validation (capturing the results of cell-plate culture experiments), realistic exposure scenarios, inclusion of the nasopharynx (and other connecting regions between the nasal passage and the trachea), competing virus types inside one RT and the implementation of formal sensitivity analysis methods.

BIBLIOGRAPHY

[1] J. Allen et al. Coronavirus world map: Tracking the global outbreak. *New York Times*, 2022.

[2] B. Asgharian, W. Hofmann, and F. Miller. Mucociliary clearance of insoluble particles from the tracheobronchial airways of the human lung. *Journal of Aerosol Science*, 32(6):817–832, 2001.

[3] Y. M. Bar-On, A. Flamholz, R. Phillips, and R. Milo. Science forum: Sars-cov-2 (covid-19) by the numbers. *elife*, 9:e57309, 2020.

[4] A. G. Beule. Physiology and pathophysiology of respiratory mucosa of the nose and the paranasal sinuses. *GMS current topics in otorhinolaryngology, head and neck surgery*, 9, 2010.

[5] M. M. Böhmer, U. Buchholz, V. M. Corman, M. Hoch, K. Katz, D. V. Marosevic, S. Böhm, T. Woudenberg, N. Ackermann, R. Konrad, et al. Investigation of a covid-19 outbreak in germany resulting from a single travel-associated primary case: a case series. *The Lancet Infectious Diseases*, 20(8):920–928, 2020.

[6] R. Boucher, M. Lang, M. Rubinstein, T. Elston, R. Tarran, and C. Sandefur. On airway surface layer regulation. *UNC Marsico Lung Institute preprint*, 2021.

[7] A. Chen, T. Wessler, K. Daftari, K. Hinton, R. C. Boucher, R. Pickles, R. Freeman, S. K. Lai, and M. G. Forest. Modeling insights into sars-cov-2 respiratory tract infections prior to immune protection. *Biophysical Journal*, 121(9):1619–1631, 2022.

[8] R.-H. Du, L.-R. Liang, C.-Q. Yang, W. Wang, T.-Z. Cao, M. Li, G.-Y. Guo, J. Du, C.-L. Zheng, Q. Zhu, et al. Predictors of mortality for patients with covid-19 pneumonia caused by sars-cov-2: a prospective cohort study. *European Respiratory Journal*, 55(5), 2020.

[9] S. Gizurarson. Anatomical and histological factors affecting intranasal drug and vaccine delivery. *Current drug delivery*, 9(6):566–582, 2012.

[10] A. Gonçalves, J. Bertrand, R. Ke, E. Comets, X. De Lamballerie, D. Malvy, A. Pizzorno, O. Terrier, M. Rosa Calatrava, F. Mentré, et al. Timing of antiviral treatment initiation is critical to reduce sars-cov-2 viral load. *CPT: pharmacometrics & systems pharmacology*, 9(9):509–514, 2020.

[11] K. P. Hui, J. C. Ho, M.-c. Cheung, K.-c. Ng, R. H. Ching, K.-l. Lai, T. T. Kam, H. Gu, K.-Y. Sit, M. K. Hsin, et al. Sars-cov-2 omicron variant replication in human bronchus and lung ex vivo. *Nature*, pages 1–5, 2022.

[12] B. Iooss and P. Lemaître. Uncertainty management in simulation-optimization of complex systems: algorithms and applications. *A Review on Global Sensitivity Analysis Methods*, pages 101–122, 2015.

[13] R. Ke, C. Zitzmann, D. D. Ho, R. M. Ribeiro, and A. S. Perelson. In vivo kinetics of sars-cov-2 infection and its relationship with a person's infectiousness. *Proceedings of the National Academy of Sciences*, 118(49):e2111477118, 2021.

[14] S. H. Kim, F. L. Kearns, M. A. Rosenfeld, L. Casalino, M. J. Papanikolas, C. Simmerling, R. E. Amaro, and R. Freeman. Glycogrip: Cell surface-inspired universal sensor for betacoronaviruses. *ACS Central Science*, 2021.

[15] M. R. Knowles, R. C. Boucher, et al. Mucus clearance as a primary innate defense mechanism for mammalian airways. *The Journal of clinical investigation*, 109(5):571–577, 2002.

[16] B. Korber, W. M. Fischer, S. Gnanakaran, H. Yoon, J. Theiler, W. Abfalterer, N. Hengartner, E. E. Giorgi, T. Bhattacharya, B. Foley, et al. Tracking changes in sars-cov-2 spike: evidence that d614g increases infectivity of the covid-19 virus. *Cell*, 182(4):812–827, 2020.

[17] S. Kumar, T. S. Thambiraja, K. Karuppanan, and G. Subramaniam. Omicron and delta variant of sars-cov-2: a comparative computational study of spike protein. *Journal of medical virology*, 94(4):1641–1649, 2022.

[18] D. Kurowicka and R. M. Cooke. *Uncertainty analysis with high dimensional dependence modelling*. John Wiley & Sons, 2006.

[19] R. Leander, Y. Wu, W. Ding, D. Nelson, and Z. Sinkala. A model of the innate immune response to sars-cov-2 in the alveolar epithelium. *Royal Society Open Science*, 8(8):210090, 2021.

[20] D. Levin, S. Forrest, S. Banerjee, C. Clay, J. Cannon, M. Moses, and F. Koster. A spatial model of the efficiency of t cell search in the influenza-infected lung. *Journal of theoretical biology*, 398:52–63, 2016.

[21] Q. Li, K. Vijaykumar, S. Phillips, S. S. Hussain, V. N. Huynh, C. M. Fernandez-Petty, J. E. P. Lever, J. B. Foote, J. Ren, J. Campos-Gomez, et al. Mucociliary transport deficiency and disease progression in syrian hamsters with sars-cov-2 infection. *bioRxiv*, 2022.

[22] K. A. Lythgoe, M. Hall, L. Ferretti, M. de Cesare, G. MacIntyre-Cockett, A. Trebes, M. Andersson, N. Otecko, E. L. Wise, N. Moore, et al. Sars-cov-2 within-host diversity and transmission. *Science*, 372(6539):eabg0821, 2021.

[23] D. Mannar, J. W. Saville, X. Zhu, S. S. Srivastava, A. M. Berezuk, K. Tuttle, C. Marquez, I. Sekirov, and S. Subramaniam. Sars-cov-2 omicron variant: Ace2 binding, cryo-em structure of spike protein-ace2 complex and antibody evasion. *BioRxiv*, 2021.

[24] M. A. Martin and K. Koelle. Comment on "genomic epidemiology of superspreading events in austria reveals mutational dynamics and transmission properties of sars-cov-2". *Science translational medicine*, 13(617):eabh1803, 2021.

[25] H. Matsui, S. H. Randell, S. W. Peretti, C. W. Davis, R. C. Boucher, et al. Coordinated clearance of periciliary liquid and mucus from airway surfaces. *The Journal of clinical investigation*, 102(6):1125–1131, 1998.

[26] B. Meng, A. Abdullahi, I. A. Ferreira, N. Goonawardane, A. Saito, I. Kimura, D. Yamasoba, P. P. Gerber, S. Fatihi, S. Rathore, et al. Altered tmprss2 usage by sars-cov-2 omicron impacts infectivity and fusogenicity. *Nature*, 603(7902):706–714, 2022.

[27] H. Mitchell, D. Levin, S. Forrest, C. A. Beauchemin, J. Tipper, J. Knight, N. Donart, R. C. Layton, J. Pyles, P. Gao, et al. Higher level of replication efficiency of 2009 (h1n1) pandemic influenza virus than those of seasonal and avian strains: kinetics from epithelial cell culture and computational modeling. *Journal of virology*, 85(2):1125–1135, 2011.

[28] M. E. Moses, S. Hofmeyr, J. L. Cannon, A. Andrews, R. Gridley, M. Hinga, K. Leyba, A. Pribisova, V. Surjadidjaja, H. Tasnim, et al. Spatially distributed infection increases viral load in a computational model of sars-cov-2 lung infection. *PLoS computational biology*, 17(12):e1009735, 2021.

[29] K. Nguyen, S. Chakraborty, R. A. Mansbach, B. Korber, and S. Gnanakaran. Exploring the role of glycans in the interaction of sars-cov-2 rbd and human receptor ace2. *Viruses*, 13(5):927, 2021.

[30] L. B. Palmer, K. Albulak, S. Fields, A. M. Filkin, S. Simon, and G. C. Smaldone. Oral clearance and pathogenic oropharyngeal colonization in the elderly. *American journal of respiratory and critical care medicine*, 164(3):464–468, 2001.

[31] J. S. Patton and P. R. Byron. Inhaling medicines: delivering drugs to the body through the lungs. *Nature reviews Drug discovery*, 6(1):67–74, 2007.

[32] K. A. Pawelek, G. T. Huynh, M. Quinlivan, A. Cullinane, L. Rong, and A. S. Perelson. Modeling within-host dynamics of influenza virus infection including immune responses. *PLoS computational biology*, 8(6):e1002588, 2012.

[33] T. P. Peacock, J. C. Brown, J. Zhou, N. Thakur, J. Newman, R. Kugathasan, K. Sukhova, M. Kaforou, D. Bailey, and W. S. Barclay. The sars-cov-2 variant, omicron, shows rapid replication in human primary nasal epithelial cultures and efficiently uses the endosomal route of entry. *BioRxiv*, pages 2021–12, 2022.

[34] A. S. Perelson and R. Ke. Mechanistic modeling of sars-cov-2 and other infectious diseases and the effects of therapeutics. *Clinical Pharmacology & Therapeutics*, 109(4):829–840, 2021.

[35] A. S. Perelson and R. M. Ribeiro. Modeling the within-host dynamics of hiv infection. *BMC biology*, 11(1):1–10, 2013.

[36] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007.

[37] C. Quirouette, N. P. Younis, M. B. Reddy, and C. A. Beauchemin. A mathematical model describing the localization and spread of influenza a virus infection within the human respiratory tract. *PLoS computational biology*, 16(4):e1007705, 2020.

[38] A. Rygg and P. W. Longest. Absorption and clearance of pharmaceutical aerosols in the human nose: development of a cfd model. *Journal of aerosol medicine and pulmonary drug delivery*, 29(5):416–431, 2016.

[39] S. Sanyal. How sars-cov-2 (covid-19) spreads within infected hosts—what we know so far. *Emerging topics in life sciences*, 4(4):383–390, 2020.

[40] R. Sender, Y. M. Bar-On, S. Gleizer, B. Bernshtein, A. Flamholz, R. Phillips, and R. Milo. The total number and mass of sars-cov-2 virions. *Proceedings of the National Academy of Sciences*, 118(25), 2021.

[41] Y. Shang, K. Inthavong, D. Qiu, N. Singh, F. He, and J. Tu. Prediction of nasal spray drug absorption influenced by mucociliary clearance. *PLoS One*, 16(1):e0246007, 2021.

[42] Y. Shang, K. Inthavong, and J. Tu. Development of a computational fluid dynamics model for mucociliary clearance in the nasal cavity. *Journal of biomechanics*, 85:74–83, 2019.

[43] X. Shen, H. Tang, R. Pajon, G. Smith, G. M. Glenn, W. Shi, B. Korber, and D. C. Montefiori. Neutralization of sars-cov-2 variants b. 1.429 and b. 1.351. *New England Journal of Medicine*, 384(24):2352–2354, 2021.

[44] A. M. Syed, T. Y. Taha, T. Tabata, I. P. Chen, A. Ciling, M. M. Khalid, B. Sreekumar, P.-Y. Chen, J. M. Hayashi, K. M. Soczek, et al. Rapid assessment of sars-cov-2–evolved variants using virus-like particles. *Science*, 374(6575):1626–1632, 2021.

[45] M. Szmigiera. Impact of the coronavirus pandemic on the global economy - statistics & facts. *Statista*, 2022.

[46] S. Wang, Y. Pan, Q. Wang, H. Miao, A. N. Brown, and L. Rong. Modeling the viral dynamics of sars-cov-2 infection. *Mathematical biosciences*, 328:108438, 2020.

[47] C. Warrender, S. Forrest, and F. Koster. Modeling intercellular interactions in early mycobacterium infection. *Bulletin of mathematical biology*, 68(8):2233–2261, 2006.

[48] T. Watanabe, T. A. Bartrand, M. H. Weir, T. Omura, and C. N. Haas. Development of a dose-response model for sars coronavirus. *Risk Analysis: An International Journal*, 30(7):1129–1138, 2010.

[49] B. J. Willett, J. Grove, O. MacLean, C. Wilkie, N. Logan, G. De Lorenzo, W. Furnon, S. Scott, M. Manali, A. Szemiel, et al. The hyper-transmissible sars-cov-2 omicron variant exhibits significant antigenic change, vaccine escape and a switch in cell entry mechanism. *MedRxiv*, 2022.

[50] A. B. Yoo, M. A. Jette, and M. Grondona. Slurm: Simple linux utility for resource management. In *Workshop on job scheduling strategies for parallel processing*, pages 44–60. Springer, 2003.