Modeling insights into SARS-CoV-2 respiratory tract infections prior to immune protection

Alexander Chen, 1 Timothy Wessler, 2,* Katherine Daftari, 2 Kameryn Hinton, 3 Richard C. Boucher, 4 Raymond Pickles, 4,5 Ronit Freeman, 3 Samuel K. Lai, 5,6,7 and M. Gregory Forest 2,3,6,*

1Department of Mathematics, CSU Dominguez Hills, Carson, California; 2Department of Mathematics, UNC Chapel Hill, Chapel Hill, North Carolina; 3Department of Applied Physical Sciences, UNC Chapel Hill, Chapel Hill, North Carolina; 4Marsico Lung Institute, UNC Chapel Hill, Chapel Hill, North Carolina; 5Department of Microbiology and Immunology, UNC Chapel Hill, Chapel Hill, North Carolina; 6Joint Department of Biomedical Engineering, UNC Chapel Hill and NC State University, Chapel Hill and Raleigh, North Carolina; and 7Division of Pharmacoengineering and Molecular Pharmaceutics, Eshelman School of Pharmacy, UNC Chapel Hill, Chapel Hill, North Carolina

ABSTRACT Mechanistic insights into human respiratory tract (RT) infections from SARS-CoV-2 can inform public awareness as well as guide medical prevention and treatment for COVID-19 disease. Yet the complexity of the RT and the inability to access diverse regions pose fundamental roadblocks to evaluation of potential mechanisms for the onset and progression of infection (and transmission). We present a model that incorporates detailed RT anatomy and physiology, including airway geometry, physical dimensions, thicknesses of airway surface liquids (ASLs), and mucus layer transport by cilia. The model further incorporates SARS-CoV-2 diffusivity in ASLs and best-known data for epithelial cell infection probabilities, and, once infected, duration of eclipse and replication phases, and replication rate of infectious virions. We apply this baseline model in the absence of immune protection to explore immediate, short-term outcomes from novel SARS-CoV-2 depositions onto the air-ASL interface. For each RT location, we compute probability to clear versus infect; per infected cell, we compute dynamics of viral load and cell infection. Results reveal that nasal infections are highly likely within 1–2 days from minimal exposure, and alveolar pneumonia occurs only if infectious virions are deposited directly into alveolar ducts and sacs, not via retrograde propagation to the deep lung. Furthermore, to infect just 1% of the 140 m² of alveolar surface area within 1 week, either 10^3 boluses each with 10^6 infectious virions or 10^6 aerosols with one infectious virion, all physically separated, must be directly deposited. These results strongly suggest that COVID-19 disease occurs in stages: a nasal/upper RT infection, followed by self-transmission of infection to the deep lung. Two mechanisms of self-transmission are persistent aspiration of infected nasal boluses that drain to the deep lung and repeated rupture of nasal aerosols from infected mucosal membranes by speaking, singing, or cheering that are partially inhaled, exhaled, and re-inhaled, to the deep lung.

SIGNIFICANCE We present a respiratory viral infection model that explicitly incorporates the first line of airway defense: mucociliary clearance (MCC). Mucus layers coat the airways and trap inhaled particulates, while the periciliary layer continuously propels mucus to be swallowed. Prior to immune response, MCC is the sole protection from inhaled pathogens. Our model incorporates thickness and transport velocity of all mucus layers, periciliary layer and alveolar liquid thickness, and respiratory tract airway dimensions. The model further incorporates SARS-CoV-2 diffusivity, probabilities to infect epithelial cells, infected cell replication rate and duration of infectious progeny. Model simulations predict outcomes of viral load and infection from inhaled SARS-CoV-2 virions throughout the respiratory tract, shedding insight into clinical outcomes prior to immune response.

INTRODUCTION

SARS-CoV-2 has led to a worldwide pandemic while exposing knowledge gaps in the medical, public health, and scientific communities. We have developed a comprehensive model to simulate outcomes from viral exposures throughout the human respiratory tract (RT), starting here with a baseline model for novel viruses in the absence of...
immune protection. The model incorporates detailed anatomy and physiology of the RT (Table S1) with SARS-CoV-2 mobility in ASL barriers, epithelial cell infection probabilities, replication rates, and duration of virion progeny by infected cells (Table S2). The model is thereby adaptable to any respiratory virus, in particular the mutations in SARS-CoV-2 variants that modify the kinetics of cell infection and replication. We apply the model to gain mechanistic insights into widely observed outcomes: 1) nasal/upper respiratory tract (URT) infections develop within 1–2 days after mild inhaled exposure to SARS-CoV-2; and 2) ~3%–5% of nasal/URT infections progress within 1–2 weeks to alveolar pneumonia.

For a novel and potentially deadly coronavirus, a mechanistic understanding of clinical outcomes is essential to guide medical prevention and intervention strategies. In response to this and future novel respiratory viruses, we assume the absence of innate or adaptive immune responses or delivery of monoclonal antibodies or antiviral drugs and build a within-host model to explore early outcomes from inhaled exposures deposited anywhere in the human RT. One rationale for this starting point is the delay between exposure and ramping up of the immune response to a novel virus exposure. In this paper, we focus on known or best estimates of properties of the SARS-CoV-2 alpha variant. The model developed here provides a baseline for future extensions: to explore current and future SARS-CoV-2 variants or other respiratory viruses; to incorporate immune responses and medicinal interventions; and to develop a platform to design optimal intervention strategies. While we focus model simulations in the immediate hours and days (up to 1 week) after exposure, we acknowledge that some immune response, especially interferon signaling (see (1,2)), is often triggered within these timescales. These effects can be coupled to the model presented by introducing half-life distributions of both viruses and infected cells, as well as reduced replication rates; for SARS-CoV-2, we await this information. Incorporation of immune and therapeutic responses into model extensions and studies of modified outcomes are in progress, but first we build the pre-immunity baseline.

Our RT model for pathogen exposure and outcomes starts with the following fact. When presented with a novel viral exposure in the RT, the primary defense mechanism is mucociliary clearance (MCC) (3–9). Inhaled particulates deposit onto the air-mucus interface, and the mucus layer is continuously propelled by coordinated, beating cilia in the periciliary liquid (PCL) layer between the epithelium and mucus layer (Fig. 1). MCC occurs everywhere in the RT except the alveolar space. This escalator-like advection of the mucus layer channels the mucus layer plus trapped insults to the oropharynx and stomach, transporting up from the lower RT (LRT) and down from the oropharynx. Our within-host model of RT exposure, clearance, and infection from inhaled viruses is unique by explicitly incorporating MCC, and thus is different from and complementary to other successful within-host models of viral infection of the RT (cf. (10–19) and (20–22)).

FIGURE 1 .Schematic of the RT. (A) The nasal and lower RT network (excluding pharynx and larynx) idealized as cylinders plus the alveolar space. (B) Cylindrical airways are multilayered: luminal airspace, mucus layer with variable thickness and advection velocity, uniform 7 µm, nonadvecting, periciliary liquid (PCL), and epithelial cells/tissue. (C) The cylindrical generations can be locally unfolded to a rectangular geometry, with x, y, and z coordinates as shown. Note that the ratio of mucus to PCL thickness shown is highly variable, so that (B) and (C) are loosely representative of generation 10 of the lower RT.
The computational modeling platform incorporates the following (see Tables S1 and S2):

- Anatomy and physiology per branch generation in the LRT: branch length and diameter, airway surface liquid (ASL) thickness (with a uniform 7 μm thickness of the PCL, variable thickness of the mucus layer, and uniform 200 nm alveolar liquid layer thickness), and variable mucus layer advection
- Diffusivity of SARS-CoV-2 virions in ASLs
- Cell infectivity and kinetics of infected cell replication per location in the RT: percent surface area coverage of SARS-CoV-2 infectable cells, probability of cell infection per encounter per second, timescale between virion uptake and cell replication of virions, infected cell replication rates and duration of viable infectious virions (where cell infection probability is a proxy for affinity to, and the density of, angiotensin-converting enzyme (ACE2) or other receptors (23,24) and the endosomal entry pathway (25) for the virus to enter and hijack cellular machinery)

The model, starting from any infectious inoculum or time series of inocula deposited onto the air-mucus or air-alveolar liquid interface, tracks two evolving fronts: the infectious virion front, along with the infectious titer in the ASL in the wake of the front; and the infected epithelial cell front, along with the total infected cell surface area in the wake of the front. These fronts are defined by the leading infectious virion and infected cells farthest from the initial site of deposition, while infection and replication continue in the wake. A critical disease metric is the percentage of infected alveolar type 2 (AT2) cells within the 140 m² total alveolar surface area (26,27), simulated and compiled below versus the number of physically separated, infectious seeds deposited into the alveolar space over a 7-day period. We rely on the rich history of computational modeling and experimental validation of inhaled depositions in the RT (28–33), including deposition of SARS-CoV-2 in the upper RT (URT) (34).

### Anatomy and physiology of the human respiratory tract

The human RT (Fig. 1) consists of the URT and LRT: the URT includes structures outside of the lungs—the nasal cavity, pharynx, and larynx—whereas the LRT, or tracheobronchial tree, consists of the subglottic branching structures of consecutively shorter and narrower, roughly cylindrical airways embedded in highly vascularized tissue, plus the terminal alveolar ducts and sacs where oxygen exchange is performed. Each new branch level is termed a “generation,” for a total of ~24 generations with the trachea the initial (0) and the terminal alveolar sacs the 23rd. Literature estimates (6,28,29,35) for the physical dimensions of each generation, the mucus layer thickness and advection velocity, and alveolar liquid thickness are presented in Table S1. In addition, the PCL layer, ~7 μm thick, persists throughout the RT except in the alveolar ducts and sacs. In the PCL, metachronal waves of coordinated, beating cilia propel the mucus layer toward the esophagus to be swallowed. We note that due to a lack of information about mucociliary advection of ASL in the noncylindrical parts of the URT between the nasal passage and trachea, we do not have results to present here for exposure outcomes of the oropharyngeal tissue. There is a strong correlation between infection in the nasal passages and oropharynx inferred by viral loads in nasal fluids and saliva (36), as further evidenced by effectiveness of testing for infection from either nasal fluids or saliva.

### SARS-CoV-2/ASL/epithelium kinetics: Diffusivity in ASL, infectability of epithelial cells, and infectious virion replication

SARS-CoV-2 is a single-strand, enveloped β-coronavirus with a hydrodynamic diameter of between 65 and 125 nm (37). To encounter infectable cells, virions must diffuse through the ASL. Measured virion diffusivity is 1.27 μm²/s (Table S2). To mediate cell entry, the SARS-CoV-2 spike (S) protein binds to human ACE2 receptors and fuses with the epithelial cell membrane (38,39). The model encodes a probability to infect per encounter per second with an infectable cell, either ciliated or AT2 cells; this probability is a consequence of spike-receptor-binding affinity and the density of receptors on infectable cells. (Recent evidence shows that secretory cells may also be infectable, so results presented here may underestimate infection and spread outside the alveolar region.) Once inside the cell, an eclipse phase entails hijacking cellular machinery; viral RNA is then uncloaked, replicated, transported to the cell membrane, and shed into the ASL (37,40). The eclipse phase (Table S2) for the alpha variant is ~12 h (37,40). Post eclipse, infected cells shed ~50 K progeny per day into the ASL for ~3 days, of which only ~1–2 K per day are infectious (37,40). Note that virion shedding into the ASL dominates direct cell-to-cell infection by virions diffusing in the ASL and being advected while in mucus. These processes amplify nonlocal virion-cell encounters, accelerating both infection and infectious viral load. In a follow-up study (41), we explore sensitivity in the outcomes of viral load and infection due to virion-cell binding and infection probability per encounter, variabilities in eclipse phase duration, and replication rate and duration of infectious virions, mechanistic parameters that have evolved during the pandemic through waves of documented SARS-CoV-2 mutations.

### THE MODELING PLATFORM

**Virion-infection twin kernel for cylindrical generations of the LRT and nasal passage**

The infection kernel, described next, is a numerical algorithm that generates a distribution map of cells infected by
virions shed from an infected cell (Fig. 2). Infected cells shed infectious progeny, and sampling from the infection kernel allows an infection map to be created. Meanwhile, shed infectious virions are passed to the virion kernel, the algorithm that tracks all virions, spatially and dynamically. As virions shed from an infected cell propagate in the ASL, they enter a cylindrical domain. In this context, infec-
table cells and infect with a cell-specific probability. If virions infect, they are removed from the ASL (absorbing boundary condition); if not, they reflect. In the azimuthal direction, the cylindrical domain is modeled as a cylindrical shell, with epithelial cells at the outer boundary of the cylinder (see Fig. 1). Let x denote the radial, y the axial, and z the azimuthal position coordinates (see Fig. 2). Then x = 0 denotes the PCL-epithelium interface, and mucus advection is in the y direction, with total length of each generation given in Table S1. The virion equations of motion are:

\[
x = \sqrt{2D_1} dW_1, \\
y = \sqrt{2D_2} dW_2 + \int s_{\text{adv,gen}} \{ x > PCL_{\text{gen}} \} dt, \\
z = \sqrt{2D_3} dW_3,
\]

where \( dW_1, dW_2, dW_3 \) each denote one-dimensional Brownian motion, \( D_i \) is the diffusion coefficient of a virion, \( s_{\text{adv,gen}} \) is the advection velocity of MCC in the given generation, \( \int \{ x > PCL_{\text{gen}} \} \) is the indicator function for the mucus layer, and \( PCL_{\text{gen}} \) is the PCL thickness, 7 \( \mu \)m. The indicator function on the advection term corresponds to advection of the mucus layer, propelled by beating cilia in the PCL which oscillates but does not advect. Since the azimuthal direction is large in compared with the distance virions diffuse azimuthally in the presence of strong advection (Fig. 1), the cylinder can be unrolled into a rectangular prism domain. Boundary conditions at \( x = 0 \) are absorbing if infection occurs; reflecting otherwise. At \( x = M_{\text{gen}} \), the sum of mucus plus PCL layers per generation, the boundary condition is reflecting. In the azimuthal direction, the boundary condition is periodic; infection does not spread around the circumference within the timescales presented.

As depicted in Fig. 2, the epithelial cell interface is modeled as a rectangular domain with square-shaped cells of side length \( c_{\text{cell diam}} \). Since not all cells are SARS-CoV-2 infectable (~50% above and ~20% within the alveolar region, which could also be as high as 50%), each cell is modeled as infectable with probability \( c_{\text{infectable}} \). Certain subpopulations have higher or lower cell infectabilities, which drugs can target (42,43). In addition, infection of a cell depends on virion-cell binding interactions, dictated by the binding affinity and avidity of spike-receptor-binding domains to ACE2 plus other receptors and their density, more generic targets for drugs (44,45). In this context, infection is modeled as \( p_{\text{infection}} \), a probability of infection per encounter per second while a virion is within the distance that allows a spike-receptor binding event. When a virion reaches position \( x \leq 0 \), it encounters a cell; infection occurs...
with a prescribed probability. If infection occurs, the virion is removed and the \((y, z)\) position of the cell and time of infection \(t_{\text{infect}}\) are recorded. Otherwise, the virion is reflected into the ASL at position \(x\). Virions are tracked until they infect. The result is an infection map \(\{(y_i, z_i)\}_{i=1}^{N}\) of cells relative to the initial infection and corresponding infection times \(\{t_{\text{infect}}^{i}\}_{i=1}^{N}\). After time \(t_{\text{latency}}\) infected cells start shedding infectious progeny whose diffusive PCL transport and advective-diffusive mucus layer transport are tracked by the virion kernel. For virion progeny, the infection kernel is sampled repeatedly to obtain locations of new generations of infected cells. From a shedding cell \(c_i\), the infection kernel can be sampled to obtain a list of all cells \(\{c_{ij}\}_{j=1}^{N_i}\) that were infected by virions originating at \(c_i\); \(N_i\) is the total number of cells infected by virions originating from cell \(c_i\), up to the total simulation time \(T\). The infection kernel also provides times \(\{t_{\text{infect}}^{i}, N_i\}_{j=1}^{N_i}\) at which infections occur, relative to the shedding time of the virion. From an infected cell, shedding is assumed to occur at equal time increments such that \(\nu_{\text{prod}} \approx 2000\) infectious virions are produced per day per cell for \(\sim 3\) days (after which time we allow the cell to expire and be replaced by an uninfected cell). Then for the \(j\)th virion shed, \(t_{ij} = t_{\text{infect}}^{i} + \frac{1}{\nu_{\text{prod}}} + t_i + t_{\text{latency}}\). Care is taken to efficiently store the propagation of infected cells and infectious virions. Infected cells are allowed to continue to bind and uptake infectious virions, since we do not yet know bounds on uptake, but once infected their replication history does not change.

**Virion-infection twin kernel for the alveolar space**

In the alveolar space, two important issues are: 1) the map between the number of spatially separated infectious seeds and the total and percent surface area of infected AT2 cells; and 2) whether the sizes and titer of the infectious seeds (inocula) are important. The alveolar liquid is modeled as a large \(I \times J\) rectangular region in which virions diffuse absent of advection. To simplify calculations with up to \(10^5\) seeds, the infection kernel is run to obtain the farthest infected cell from the cell of origin. The latency period (12 h) before cells shed virions is utilized as a lower bound for the time period of the spread of virions for a single infected cell. After \(T\) days, the infection from an initial seed will thus have advanced at most a distance of \(R_{\text{spread}} \approx 3\). The result is a set of circular regions \(\{C_k\}\) where \(C_k = \left\{B \left( c_k, r_k + \frac{T_{\text{spread}}}{r_{\text{spread}}} \right) \right\}_{k=1}^{N}\) where \(c_k, r_k\) are the center and original radius of the seed, respectively, \(N\) is the total number of seeds, and \(B\) is a circle with the given center and radius. By taking \(\bigcup_{k=1}^{N} C_k\), the total area coverage of the infectious seeds can be calculated.

Our modeling platform leverages a parallelization of deposition events and outcomes in both space and time. One cannot simulate, store, and post-process data everywhere throughout the RT, of all airway surface liquids, all epithelial cells, all infectious virions in ASLs, all infected cells and their shedding of infectious virions, and all uninfected cells available to infect, over time, from the first and all subsequent inhaled exposures. Thus, we parallelize. We model and simulate the outcome from any initial condition: the onset of infection of one cell, anywhere in the RT. The modeling outcome per initial condition is the propagation, in time and from the location of the RT, of infected epithelial cells (infection kernel output) and of infectious virions in ASLs (virion kernel output). We then superimpose deposition-infection events, in time and locations, specific to the exposure scenarios.

**Aggregation of the twin kernel to create maps of viral load and infected cells**

The virion kernel simulates, dynamically tracks, and stores data for all infectious virions in the ASL arising from each specified virion seed-deposition-infection event, including the leading fronts of infectious virions in the ASL and the total number of infectious virions. When a virion passes through the ASL, encounters infectable cells, and successfully infects, it triggers removal of that virion from the ASL and a call to the infection kernel; virion-cell encounters that do not infect trigger virion reflection back into the ASL. Note that virions at the leading fronts and within the expanding, infected ASL volume continue to be tracked and stored separately as free (in ASL) and within cells (total minus free).

The infection kernel simulates, dynamically tracks, and stores data for all cells that become infected after the “first” virion seed-deposition-infection event. Infected cells have an initial 12 h eclipse phase during which cell machinery is hijacked, then a shedding phase with 2000 infectious virions per day for 3 days shed into the ASL; these data are fed to the virion kernel.

For each virion seed-deposition-infection event, we separately track the propagation of infected cells in eclipse and shedding phases. From these data, it is possible to sum over all spatially separated depositions, on whatever time schedule the seed depositions are posited, and wherever the depositions occur. In the model simulations and results section, we simulate single deposition-infection events in the nasal passage, trachea, and generations 5, 10, and 15 of the LRT, as well as the alveolar space. These results can be superimposed over any schedule of depositions; we defer details for diverse exposure scenarios to future studies. The trachea (generation 0) and generations 1–19 in the LRT as well as the nasal passage have a cylindrical geometry with dimensions per generation of the length, air core, mucus layer and PCL thickness, and epithelial circumference (Table S1); the percent surface area of infectable (ciliated) cells and infectability and replication kinetics are presented in Table S2.
The code for cylindrical generations can be accessed at https://github.com/mathalexchen/SARS-CoV-2_model.

MODEL SIMULATIONS AND RESULTS

Table 1, columns 1–6 recapitulate data from Tables S1 and S2 for the nasal passage and generations 0–19 of the LRT: length, diameter, mucus layer advection, length/advection = clearance time of the mucus escalator to traverse the full generation length, and lower and upper estimates of mucus layer depth (the mean is used in simulations). Recall that each of these generations has a 7 μm PCL. For each generation, we simulate 10^7 realizations at locations of deposition: the upstream end (column 6, so that mucus advects into the generation); the midpoint (column 7); and uniformly distributed within the inner second and third quartiles of each generation (column 8). From the simulated data we deduce the probability to infect a cell within the generation.

Results from Table 1, column 7 show that the midpoint probabilities to infect within the generation are high in the nasal passage (~69%), decrease very rapidly from the trachea to generation 2 (~7%), then reverse and reach probability 1 in generations 8–19. The LRT certainty to infect is a consequence of multiple effects. Mucus advection gets weaker, the mucus layer becomes thinner than the nonadvecting PCL, and therefore virion-cell encounters are so frequent that they infect before diffusing out of the generation.

Results from Table 1, column 8 show a uniform deposition within the middle two quartiles of each generation, quite close to the midpoint results in column 7.

Results from Table 1, column 6 show that the upstream-end probabilities to infect within the generation are extremely high in the nasal passage (~91%), rapidly decrease from the trachea (~42%) to generation 2 (~23%), then reverse and increase to probability ~1 in generations 7–9. Since virion diffusion and cell infection spread upstream as well as downstream in generations 10–19 due to weak mucus advection, a percentage of virions infect the lower, i.e., upstream, generation, affecting the within-generation count. This foreshadows results to come for extremely slow propagation of infectious virions and infected cells in the deep lung generations.

Table 1 data suffice to approximate, per generation, infection within versus clearance from a given viral load deposited at the air-mucus interface. One can convolve data from successive generations to calculate probabilities to clear the RT and statistical outcomes of onset and propagation of infection. To do this analysis, another layer of complexity is necessary. One must generalize Table 1, column 6 to virions entering the generation at any height within the ASL from the previous generation; such results are deferred to another study. These results empower use of a simpler initial condition per generation, employed below: an infected cell as it emerges from the eclipse phase, shedding 2000 infectious virions per day for 3 days. (In a subsequent study (37), we generalize Table 1 and explore sensitivity of the results to duration of the eclipse phase, and rate and duration of replication by infected cells. These results are informative with respect to outcomes from SARS-CoV-2 variants.)

We next present details of infected cell outcomes (Fig. 3) and infectious virion outcomes (Fig. 4) from one realization of the model in the nasal passage and LRT generations 0 (the trachea), 5, 10, and 15, with initial condition at t = 0 of one infected cell that has just emerged from the eclipse phase, located at the upstream end of each generation. Fig. 3 presents snapshots of infected cells, both shedding (red) and not shedding (yellow), at t = 12, 24, and 36 h. Recall that mucus advection is downward in the nasal passage and upward in the LRT, always toward the esophagus.

Fig. 4 shows infectious virion outcomes from row 3 in Fig. 3: the total number of infectious virions replicated in 36 h and the total number of “free virions” in the ASL. We note that in the nasal passage, trachea, and generation 5, where advection is significant, after 36 h the free virion and infected cell fronts have escaped the generation, consistent with results in Table 1.

Results from Figs. 3 and 4 show that in the URT and LRT above generation 10 where mucus advection dominates virion diffusion, infection of epithelial cells and infectious virions propagate strongly toward the esophagus, away from the alveolar generations. Infected cells, within 36 h post emergence from the eclipse phase, create streaks of infection that exceed the entire lengths of the nasal passage, trachea, and generation 5, while spanning only a fraction of the circumference. Meanwhile, the total viral load produced and shed into the ASL over the same 36 h from one infected cell is on the order of 10^8, denoted O (10^8). Also depicted, O (10^7) virions have cleared the generation, O (10^5–10^6) remain free in the ASL, and all remaining virions are within infected ciliated cells. (We did not limit how many virions can bind to and become internalized in an infected cell, since we do not know that information. Such limits will only amplify the spread of infectious virions and infected cells.)

In generations 10–19 of the LRT where mucus advection is sufficiently weak and the mucus layer is sufficiently thin, virions frequently encounter and infect epithelial cells. Therefore, even though advection does not strongly bias the spread of infected cells and infectious virions away from the deep lung, the spread of both infectious virions and infected cells is extremely slow, with negligible progression toward alveolar generations 20–23.

The above results reveal a critical insight: infection of the alveolar ducts and sacs requires deposition of infectious seeds directly into those LRT generations.

Next, we explore how important are the size and infectious virion count of infectious seeds deposited in the alveolar ducts and sacs. We note that many persons with COVID-19 alveolar pneumonia aspirate nasal/oropharyngeal fluids into their lungs, especially while sleeping. The aspirated
<table>
<thead>
<tr>
<th>Generation</th>
<th>Assumed length (cm)</th>
<th>Assumed mucus layer thickness (µm)</th>
<th>Assumed advection (mm/min)</th>
<th>Time for mucus layer to advect the full length (h)</th>
<th>Probability of infection (starting at upstream end)</th>
<th>Probability of infection (starting at midpoint)</th>
<th>Probability of infection (starting in middle two quartiles)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasal</td>
<td>$1.30 \times 10^1$</td>
<td>17.0</td>
<td>8.8000</td>
<td>0.246</td>
<td>0.9093</td>
<td>0.6863</td>
<td>0.6721</td>
</tr>
<tr>
<td>0</td>
<td>$1.00 \times 10^1$</td>
<td>50.0</td>
<td>5.5000</td>
<td>0.303</td>
<td>0.4177</td>
<td>0.1733</td>
<td>0.1727</td>
</tr>
<tr>
<td>1</td>
<td>$4.36 \times 10^0$</td>
<td>44.4</td>
<td>3.9100</td>
<td>0.186</td>
<td>0.3000</td>
<td>0.0989</td>
<td>0.1056</td>
</tr>
<tr>
<td>2</td>
<td>$1.78 \times 10^0$</td>
<td>38.8</td>
<td>2.4900</td>
<td>0.119</td>
<td>0.2229</td>
<td>0.0644</td>
<td>0.0700</td>
</tr>
<tr>
<td>3</td>
<td>$9.65 \times 10^{-1}$</td>
<td>33.1</td>
<td>1.5400</td>
<td>0.104</td>
<td>0.2619</td>
<td>0.0868</td>
<td>0.0880</td>
</tr>
<tr>
<td>4</td>
<td>$9.95 \times 10^{-1}$</td>
<td>27.5</td>
<td>0.8890</td>
<td>0.187</td>
<td>0.6024</td>
<td>0.3164</td>
<td>0.3120</td>
</tr>
<tr>
<td>5</td>
<td>$1.01 \times 10^0$</td>
<td>21.9</td>
<td>0.4960</td>
<td>0.339</td>
<td>0.9046</td>
<td>0.6857</td>
<td>0.6625</td>
</tr>
<tr>
<td>6</td>
<td>$8.90 \times 10^{-1}$</td>
<td>19.6</td>
<td>0.2960</td>
<td>0.501</td>
<td>0.9857</td>
<td>0.8693</td>
<td>0.8463</td>
</tr>
<tr>
<td>7</td>
<td>$9.62 \times 10^{-1}$</td>
<td>17.3</td>
<td>0.1670</td>
<td>0.960</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9999</td>
</tr>
<tr>
<td>8</td>
<td>$8.67 \times 10^{-1}$</td>
<td>15.0</td>
<td>0.1010</td>
<td>1.43</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>9</td>
<td>$6.67 \times 10^{-1}$</td>
<td>12.7</td>
<td>0.0616</td>
<td>1.80</td>
<td>0.9988</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>10</td>
<td>$5.56 \times 10^{-1}$</td>
<td>10.4</td>
<td>0.0396</td>
<td>2.34</td>
<td>0.9901</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>11</td>
<td>$4.46 \times 10^{-1}$</td>
<td>9.32</td>
<td>0.0252</td>
<td>2.95</td>
<td>0.9575</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>12</td>
<td>$3.59 \times 10^{-1}$</td>
<td>8.29</td>
<td>0.0165</td>
<td>3.63</td>
<td>0.8973</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>13</td>
<td>$2.75 \times 10^{-1}$</td>
<td>7.26</td>
<td>0.0113</td>
<td>4.06</td>
<td>0.8029</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>14</td>
<td>$2.12 \times 10^{-1}$</td>
<td>6.23</td>
<td>0.0080</td>
<td>4.44</td>
<td>0.7757</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>15</td>
<td>$1.68 \times 10^{-1}$</td>
<td>5.20</td>
<td>0.0057</td>
<td>4.93</td>
<td>0.7188</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>16</td>
<td>$1.34 \times 10^{-1}$</td>
<td>4.70</td>
<td>0.0041</td>
<td>5.47</td>
<td>0.6548</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>17</td>
<td>$1.20 \times 10^{-1}$</td>
<td>4.21</td>
<td>0.0028</td>
<td>7.07</td>
<td>0.6509</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>18</td>
<td>$9.20 \times 10^{-2}$</td>
<td>3.72</td>
<td>0.0014</td>
<td>11.0</td>
<td>0.5678</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>19</td>
<td>$8.00 \times 10^{-2}$</td>
<td>3.22</td>
<td>0.0010</td>
<td>13.9</td>
<td>0.5957</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

The total airway surface liquid thickness of these generations is the mucus thickness (column 3), which varies dramatically, plus the 7 µm periciliary liquid layer, which has a uniform thickness throughout the RT.
of alveolar seed deposition: a 1 mL bolus with $10^6$ infectious infected AT2 surface area after 7 days between two extremes

First, we have demonstrated the double-edged nature of alveolar seed deposition: a 1 mL bolus with $10^6$ infectious virions versus a submicrometer aerosol with one infectious virion. Since the alveolar fluid thickness is comparable with SARS-CoV-2 virion diameter and there is negligible advection, virion-AT2 cell encounters occur extremely frequently, so to emphasize the point of the next result we exaggerate the outcome of the bolus by assuming that all virions infect immediately.

Results from Fig. 5 show that the infected surface area after 7 days is $\sim 4.4 \times 10^{-3} \text{ m}^2$ from a 1 mL bolus with $10^6$ infectious virions and $\sim 4.0 \times 10^{-16} \text{ m}^2$ from a submicrometer aerosol with one infectious virion, each on their own a negligible percentage of the 140 m$^2$ of alveolar surface area. The potential sources of these seeds are discussed below.

These alveolar deep lung results are conservative estimates of within-host outcomes from inhaled exposure to SARS-CoV-2, since we have not considered any protection from macrophages or T cells that would require even more infectious alveolar seeds to produce the same degree of infection.

CONCLUDING REMARKS

The pre-immunity model developed and applied here provides a mechanistic tool to explore within-host exposures to SARS-CoV-2 in the RT and to connect exposures to outcomes ranging from clearance to infection and progression of infection to COVID-19 disease. This baseline model assumes that the RT does not have immune protection for the timescales explored in this paper; all sources and mechanisms of protection and timing of protection will be integrated and explored in subsequent studies. Results presented yield the following insights into observed outcomes from exposure to SARS-CoV-2.

First, we have demonstrated the double-edged nature of the mucus barrier and MCC. If present in mucus, pathogen-specific antibodies bind to viral spikes to neutralize the virion. Furthermore, antibodies may also bind to mucin polymers, creating virion-antibody-mucin crosslinks that suppress viral (and bacterial) mobility in mucus barriers on all internal organs, amplifying the protection by natural clearance mechanisms of mucus barriers (46–54). However, when novel viruses deposit at the air-mucus interface and there are no virus-specific antibodies in the mucus layer, results presented show that MCC in the URT and upper generations of the LRT boosts the spread and growth of infection and viral load. This result is consistent with the widespread occurrence of nasal/URT infections from mild SARS-CoV-2 exposures. Indeed, the same scenario applies to any new strain of seasonal influenza or respiratory virus, where predictions require using virus-specific parameters for cell infection, eclipse phase prior to replication, and replication rate and duration of infectious virus. Any environment where aerosols from infected persons are exhaled and remain suspended due to poor ventilation will provide the opportunity for inhaled exposure and deposition of infectious aerosols with at least one SARS-CoV-2 virion. Strikingly, experimental and modeling data provide compelling evidence that an exceptionally low number of
infectious aerosol seeds are sufficient to launch a nasal/URT infection, ranging from 100 (34) to as low as 1–8 (55), 1–3 (56), and even 1 (57)!

We also remark on the relationship between the scalar statistic for within-host spread of infection, called $R_0$, within, and our predictions of spread of within-host infection. The scalar parameter $R_0$, within is a robust statistic for analysis of cell-culture experiments (59) where an infectious virus (or viral load) is placed on a stationary liquid layer about cells, after which the spread of infection per infected cell is tracked over short timescales (hours to a day), comparable with results presented in this paper. Such cell-culture results, however, are only relatable to generations deep in the LRT where MCC is negligible. Indeed, dramatic differences in spread of infection from a single infected cell are revealed in Fig. 3 from the nasal passage to generations spanning the LRT, indicating that in vivo values of $R_0$, within are generation specific in the RT and furthermore extremely variable. A different, and much longer timescale (days to 1 month), approach to estimation of $R_0$, within was applied by (58) to data from a 2020 National Basketball Association tournament and a German dataset from an early 2020 outbreak (60). Their analysis gives an estimate of $R_0$, within of between 2.6 and 14.9 for SARS-CoV-2 infection in 2020, much lower than we predict in the nasal passage in the immediate hours and days from pre-immunity exposure and infection. Many factors come into play on these longer timescales, including immune responses as well as the

![FIGURE 4](image.png)

**FIGURE 4** Tracking virions from one infected cell over 36 h after shedding begins. From the same simulations in Fig. 3, plots of the rise over time in total number of infectious virions shed (blue), total number of virions exiting the region/generation (yellow), and total number of free virions at any moment inside the region/generation (red). The time for the mucus layer to advect the full length of a given generation, column 5 of Table 1, is labeled “escalator clear time” (solid vertical gray line). This timescale for trapped virions to clear the generation is extremely short and not visible on this scale in the top panels while being visible in the bottom panels. The time for the first virion to exit the region/generation (green dotted line) is short yet visible in the top panels, whereas no virions exit generations 10 and 15 within 36 h.

![FIGURE 5](image.png)

**FIGURE 5** Spread of alveolar infection from a 1 mL bolus with $10^6$ infectious virions versus a submicrometer aerosol with one infectious virion. (A) The purple circle represents a 1 mL bolus (radius = 37,550 μm), assumed to infect all AT2 cells below the patch at the start of day 0. The orange circle that slightly expands the purple circle represents the region of cells the bolus infects after 7 days (radius=38,670 μm). (B) A single dot in the 10 x 10 cm box represents an aerosol with one infectious virion assumed to infect a single cell at the start of day 0. The orange circle represents the region of cells the single virion infects after 7 days (radius=1120 μm). The inset image expands a 3 x 3 mm region of the 10 x 10 cm square, and at this magnification the purple dot at the center is still nearly imperceptible.
fact that upon first exposure, virions encounter uninfected cells exclusively as assumed here, whereas on longer timescales virions encounter already infected cells.

A second insight is gained by quantifying the first line of defense provided by MCC, the escalator-like transport by coordinated cilia of RT mucus layers. MCC biases both infectious virions and infected cells to spread away from the alveolar region of the deep lung. The implication of this innate protection mechanism is that alveolar infection arises exclusively from direct deposition of infectious seeds into the deep lung. Furthermore, we show that at least $10^6$ or more spatially separated aerosols with at least one infectious virion, or $10^3$ or more 1 mL boluses with $10^6$ infectious virions, must be deposited directly into the deep lung to infect 1% or more of the 140 m² of alveolar surface area. These results point to the extreme likelihood that alveolar pneumonia occurs from transmission of infectious aerosols or droplets ranging in size and viral titer that originate from the host’s URT, i.e., self-transmission. The mechanisms by which self-transmission most likely does and does not occur are discussed below. Furthermore, this self-transmission from a nasal/oropharyngeal region to the percentage of infected alveolar surface area. Left: percent surface area infected after 1, 2, ..., 10 days for $10^k$ seeds, $k = 0, 1, ..., 9$ unrealistically deposited at $t = 0$. Right: 7-day snapshots for evenly distributed depositions (orange) versus all seeds deposited at $t = 0$ (blue).

We note that a significant percentage of COVID-19 hospital patients have pre-existing conditions that promote aspiration of nasal/oropharyngeal fluid boluses into the lung. Nasal swabs reveal that 1–1.5 mL of nasal fluids of infected persons with $O(10^6)$ infectious SARS-CoV-2 virions, and a simple calculation, reveal that a 1.5 mL and 1.0 mL bolus has to divide starting at generation 2 and generation 3, respectively, in the LRT in order to drain further. As the nasal bolus laden with infectious virions continues to drain, it will divide up to 17 additional times before reaching the alveolar ducts and sacs, effectively aerosolizing each bolus into many thousands (17 divisions generates $O(10^5)$) physically separated, infectious seeds. Therefore, risk factors that promote aspiration of nasal fluids easily produce sufficient infectious alveolar seeds over 1–2 weeks. This scenario is similar to how colds or influenza settle from the nasal/oropharyngeal region deeper into the lung (26); Alfi et al. (61) recently reported sharp contrasts between the early stages of immune response to SARS-CoV-2 and influenza viruses in the nasal mucosa versus LRT. We also note that the results presented are consistent with the “infection gradient” cited from autopsies of COVID-19 patients (21), with initial high-titer nasal/oropharyngeal infections, very patchy infection throughout the LRT, and high levels of inflammation and infection in the alveolar ducts and sacs. Furthermore, Basu et al. (64) have very recently reinforced our argument that the nasopharynx acts as a seeding zone for transmission of a nasal infection into the lower airways by aspiration of boluses.

Additional mechanisms for self-transmission continue to emerge. Johnson and Morawska (62) and Morawska et al. (63), consistent with long-standing work by the within-host aerosol modeling community, present data that normal breathing does not dislodge or rupture fluid from the nasopharynx or mouth that is then exhaled into the surrounding atmosphere. However, experimental evidence by Kushalnagar et al. (65) and Yang et al. (66) clearly point to speech and singing, whereby wet membranes close and open and rupture fluids from the mouth and vocal cords, as a significant mechanism for exhalation of exhaled infectious aerosols into the surrounding air. We make a simple observation that individuals who exhale infectious aerosols—the obvious source of host-to-host transmission—are continuously exposed to their own ambient airspace, creating the potential for individuals with URT infections to exhale and re-inhale massive numbers of aerosols in a poorly ventilated environment when they are actively speaking, eating, singing, or engaging in other vocal activities such as sporting events. This self-transmission mechanism has been proposed previously and supported by clinical data from the deaf community that linked disease severity to the amount of speech activity (65). We further know from the aerosol modeling community (see (28,32,33)) that a nontrivial percentage of inhaled submicrometer aerosols deposit directly into the deep lung.

In work recently posted (41), we explore the wave of SARS-CoV-2 variants, including the delta and omicron variants (67), illustrating how to adapt the current baseline model to any novel virus or coronavirus in the early hours and days.
prior to immune responses, e.g., in the unvaccinated or previously uninfected population. In (41), we translate documented mutations on the spike, membrane, envelope, and nucleocapsid proteins of SARS-CoV-2 (see (42,68–71)) into implications for three mechanistic model parameters: binding affinity and avidity to ACE2 receptors (thus the probability to bind per encounter per second, apparently higher for the variants (68)), the duration of the eclipse phase (unreported to our knowledge), and the efficacy of replication of infectious progeny virions (also not reported thus far). This work (41) is part of an ongoing parameter sensitivity study to understand the most influential mechanistic parameters with respect to outcomes of viral load and infection from pre-immune exposure. We are also actively expanding the baseline model in this paper to include immune responses, e.g., interferon signaling to activate macrophages and T cells (modeled by introducing half-life distributions of virions and infected cells) and to incorporate the presence of antibodies in mucosal barriers from either vaccinations, previous infection, or drug inhalation.

There are clear public health implications for the findings presented here. Simple protective measures such as social distancing and wearing a mask are effective both for oneself and others! This is especially true in any environment where submicrometer aerosols are actively generated by speaking, singing, cheering, or eating, and remain suspended in the air for hours, especially indoors with poor ventilation and where specific activities are prevalent. Masks suppress the number of aerosols exhaled and inhaled. For any person with a nasal/oropharyngeal infection who engages in speaking, eating, screaming, and singing, the wet membranes in the mouth, nose, pharynx, vocal cords, and epiglottis are constantly closing and opening, thereby rupturing submicrometer aerosols (66,72–75). In poorly ventilated conditions, aerosols are partially inhaled, exhaled, and re-inhaled, as well as inhaled by others, many transporting to the deep lung.

A mechanism also exists for direct transport of aerosol seeds from a deep lung infection to the nasal/URT and external airspace: breathing! Scores of authors for decades (see (62,65,76–81), the comprehensive review (82), and recent article (83)) have articulated that the small bronchioles (generations just prior to the alveolar ducts, see Fig. 1) close and open with each breath, rupturing aerosols that then travel, either into the alveolar ducts and sacs during inhalation, or a significant fraction travels all the way up to the nasal passages or ambient air during exhalation. The remaining fraction contacts the mucus layers and presents a mechanism to self-transmit alveolar infection to the LRT.

We close with comments regarding access to the software used to generate the results presented herein. The code for cylindrical generations can be accessed at https://github.com/mathalexchen/SARS-CoV-2_model; we will post analogous code for alveolar generations as well as extensions of the code to achieve optimal performance as they become suited for stable implementation. Due to the rapid escalation of infectious virions and infected cells in the nasal passage and upper branches of the LRT, more efficient software has been developed for purposes of exploring parameter sensitivity of results presented here, in particular in (41) for exploring the likely mechanistic impacts of the delta and omicron variants of SARS-CoV-2. Another question arises as to whether a simpler, coarse-grained model and code can be developed, validated by the present model and simulations and all clinical data, and then made publicly available. Indeed, there are several versions of simpler models that would be valuable for experimentalists to use, tailored to ex vivo experiments or in vivo outcomes, specific to the degree of exposure, early versus advanced infection, and with degrees of immune responses or monoclonal antibodies and antiviral drugs delivered directly to sites of infection or in the bloodstream. All of these advances of the presented model and software are the focus of current efforts by the modeling team, with plans to make all software publicly available.

SUPPORTING MATERIAL

Supporting material can be found online at https://doi.org/10.1016/j.bpj.2022.04.003.

AUTHOR CONTRIBUTIONS


ACKNOWLEDGMENTS

Research leading to the reported results was supported in part by: RSCA grant from CSU Dominguez Hills (A.C.), NSF DMS-1929298, DMS-2028758(M.G.F., R.F., S.K.L.), Cystic Fibrosis Foundation BOUCHE9R0, BOUCHE9X00, FREEMAI9GO, National Institutes of Health grants R43AI149894-01A1 (R.P., S.K.L.), R44AI157661-01(R.P., S.K.L.), UL1TR002489 (S.K.L.), UH3 HL123645, R01 HL136961, P30 DK 065988, and P01 HL108808, the David and Lucile Packard Foundation(2013-39274; S.K.L.), the National Center for Advancing Translational Sciences (S.K.L.), the Eshelman Institute of Innovation (S.K.L.), and the North Carolina Policy Collaboratory at the University of North Carolina and the North Carolina Coronavirus Relief Fund established and appropriated by the North Carolina General Assembly (R.F., R.B., S.K.L., M.G.F.). R.F. acknowledges a Research Corporation for Science Advancement (COVID Initiative grant nos. 27350, 27352) awards, and the UNC Institute for Convergent Science Director’s Machete Award. M.G.F. acknowledges valuable communications with Bahman Asgharian, Julie Kimbell, Saikat Basu, and multiple members of the IMAGiMSM Working Group led by James Glazier and Reinhard Laubenbacher.

REFERENCES


