



OPEN The effect of random virus failure following cell entry on infection outcome and the success of antiviral therapy

Christian Quirouette¹, Daniel Cresta¹, Jizhou Li², Kathleen P. Wilkie³, Haozhao Liang^{4,5} & Catherine A. A. Beauchemin^{1,2}✉

A virus infection can be initiated with very few or even a single infectious virion, and as such can become extinct, i.e. stochastically fail to take hold or spread significantly. There are many ways that a fully competent infectious virion, having successfully entered a cell, can fail to cause a productive infection, i.e. one that yields infectious virus progeny. Though many stochastic models (SMs) have been developed and used to estimate a virus infection's establishment probability, these typically neglect infection failure post virus entry. The SM presented herein introduces parameter $\gamma \in (0, 1]$ which corresponds to the probability that a virion's entry into a cell will result in a productive cell infection. We derive an expression for the likelihood of infection establishment in this new SM, and find that prophylactic therapy with an antiviral reducing γ is at least as good or better at decreasing the establishment probability, compared to antivirals reducing the rates of virus production or virus entry into cells, irrespective of the SM parameters. We investigate the difference in the fraction of cells consumed by so-called extinct versus established virus infections, and find that this distinction becomes biologically meaningless as the probability of establishment approaches zero. We explain why the release of virions continuously over an infectious cell's lifespan, rather than as a single burst at the end of the cell's lifespan, does not result in an increased risk of infection extinction. We show, instead, that the number of virus released, not the timing of the release, affects infection establishment and associated critical antiviral efficacy.

Typically, mathematical models describing the course of a virus infection within a host (in vivo) or a cell culture (in vitro) express the number of cells and infectious virions (virus particles) as real positive numbers (continuous), and infection events as deterministic, e.g. one virion infects 0.2 cell. By nature, however, the number of cells and virions are whole numbers (discrete) and infection events are stochastic, e.g. one virion will infect one cell 20% of the time. When dealing with large numbers of infectious virions and cells, stochastic fluctuations can become negligible, and the continuous, deterministic, mean-field approach can provide an accurate representation of the kinetics of interest. Yet the small number regime arises commonly. For example, antiviral therapy can reduce the effective number of infection-capable virions to near or below unity. In such cases, random fluctuations could have an important effect on the time course and outcome of an infection.

Several stochastic models (SMs) of virus infection kinetics have been proposed to date, e.g.¹⁻⁷. Notably, Heldt et al.⁴ predicted with an intracellular SM that most cells inoculated with a single infectious virion would result in a non-productive cell infection. But extracellular SMs typically do not account for cell infection failure post virus entry. When it has been included only some lethal replication errors were represented, e.g. lethal reverse transcription errors only⁶. In fact, there are many ways that infectious virions post cell entry can fail to cause a cell infection that will produce fully infectious progeny. Fig. 1 illustrates early steps of human immunodeficiency virus (HIV) and influenza A virus (IAV) replication and associated opportunities for abortive cell infections.

For HIV, entry begins with successful binding to the cell surface, followed by fusion of the virus membrane with the cell surface. Yet sometimes virions can be taken in via the endocytic route⁸. In this case, virions can

¹Department of Physics, Toronto Metropolitan University, Toronto, Canada. ²Interdisciplinary Theoretical and Mathematical Sciences (iTHEMS), RIKEN, Wako, Japan. ³Department of Mathematics, Toronto Metropolitan University, Toronto, Canada. ⁴Nishina Center for Accelerator-Based Science (RNC), RIKEN, Wako, Japan. ⁵Present address: Department of Physics, University of Tokyo, Tokyo, Japan. ✉email: cbeau@torontomu.ca

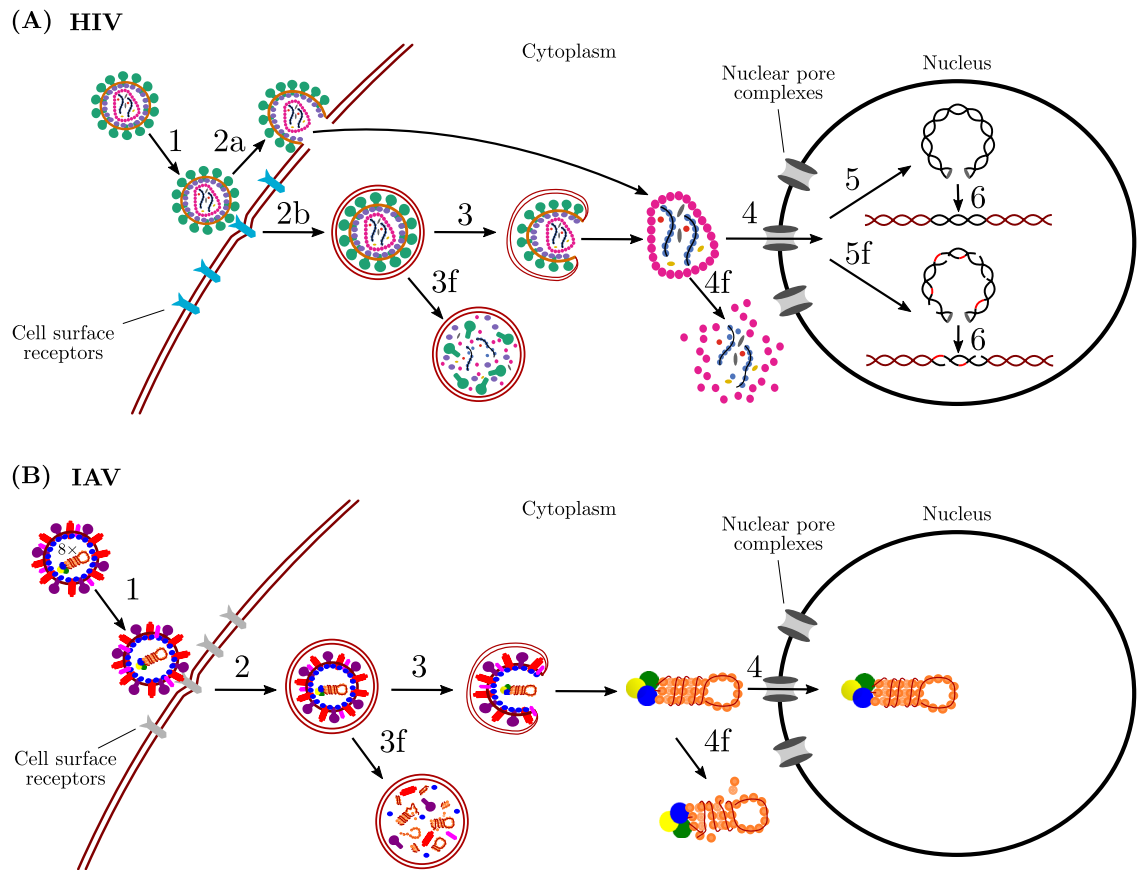


Figure 1. Early steps of HIV and IAV replication and possible routes of abortive cell infections. **(A)** An HIV virion attaches to the cell surface (Step 1), and can enter by direct fusion of the virus membrane with the cell surface (Step 2a) or be endocytosed (Step 2b). Endosomal fusion can be successful and the virus genomic material is released into the cell (Step 3) or fails and the virus-containing endosome is eventually degraded (Step 3f). Following successful entry, the virus genomic material can be imported into the nucleus (Step 4) or fail to do so and eventually be subject to degradation (Step 4f). Reverse transcription, i.e. viral RNA is converted to viral DNA, either proceeds successfully (Step 5) or fatal errors are introduced into the viral DNA during reverse transcription (Step 5f). Once in the nucleus, viral DNA is integrated into the cell DNA (Step 6) and can later be used for virus replication. **(B)** An IAV virion attaches to the cell surface (Step 1). The virion then enters by endocytosis (Step 2). Next, endosomal fusion can be successful (Step 3) or fail (Step 3f). If fusion is successful, the released viral ribonucleoproteins (vRNPs) containing the viral RNA segments can either all be imported into the nucleus (Step 4) or some might fail to enter the nucleus and be degraded (Step 4f). Once vRNPs are inside the nucleus, virus replication is then initiated. For both viruses, as the number of copies of viral RNA and proteins grow during replication, it becomes increasingly unlikely that the cell infection will fail to produce any infectious progeny.

fail to fuse with late endosomes and eventually be degraded⁹. Or, if there is successful entry, virions can still be degraded in the cytosol¹⁰. The virus can also fail to be imported into the nucleus^{11,12}. In addition, fatal mutations can be acquired during reverse transcription¹³, when viral RNA is transcribed into complementary DNA (cDNA) which is used to make typically a single copy of viral DNA for integration. Finally, failure may also be the result of host cell mechanisms that interfere with early replication steps^{14,15}.

For IAV, entry begins with successful binding to the cell surface, followed by internalization by endocytosis. Failure to undergo fusion with late endosomes can then occur^{16,17}. If uncoating happens, released viral ribonucleoproteins (vRNPs) containing the viral RNA segments could be subject to RNA degradation in the cytosol¹⁸ or fail to enter the nucleus^{17,19}. Degradation of one or more genome segments, following nuclear import and prior to transcription could also occur⁴. Since each genome segment encodes at least one viral protein, degradation of only one of them means failure to cause a productive cell infection. There is also a multitude of host cell mechanisms that can impair early steps of virus replication^{20–22} and these may be yet more potential sources for failure.

These are just some of the ways in which otherwise fully physically infectious virions could, through random chance, fail to complete a key step following cell entry. In addition to such stochastic occurrences in fully functional virions, a number of entry capable virions could have physical defects that prevents them from completing one or more key replication steps, leaving them physically unable to cause a productive infection. As such, failure of productive cell infection post virus entry is a combination of both the stochastic post-entry failure of otherwise fully infectious virions, and the inevitable failure of replication defective, entry capable virions.

A common, important application of SMs is to estimate the extinction probability of an infection: the likelihood that the infection will fail to take hold or spread significantly. It typically depends on both the replication capabilities of the virus (i.e. infection parameters) and the initial number of infectious virions or cells. The extinction probability is an important quantity to derive as it can, for example, be used to evaluate the probability of success of antiviral therapy^{3,23}. In particular, Czuppon et al.²³ compared the ability of prophylactic antivirals acting either on the rate of virus entry into cell or the rate of virus production, to reduce the establishment probability, or $1 - (\text{extinction probability})$, of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection in vivo. However, the effect of certain classes of antivirals, such as endosomal fusion inhibitors, would be better represented as reducing the probability that, having successfully entered a cell, a virion will go on to cause the productive infection of that cell. For IAV, it would be a more appropriate way to represent the mode of action of adamantanes, such as amantadine and rimantadine which block the M2 ion channel necessary for successful fusion of IAV with the endosome^{24,25}. For SARS-CoV-2, it has been suggested that cathepsin L inhibitors can reduce the probability of successful endosomal fusion²⁶. Failure to undergo endosomal fusion would lead to loss of the infectious virion, but would not result in a productive cell infection.

Although the extinction probability is often an important consideration in comparing prophylactic antivirals, what is never discussed is the number or fraction of cells that are actually consumed by infections that are said to have gone “extinct” or to have “established”. It is possible that antivirals with different modes of action, even with the same extinction probability, could result in a very different fraction of cells consumed by so-called established or extinct infections. This would have important implications for how one should interpret the probability of success of a particular antiviral therapy.

In addition, nearly all SMs, including past works that used the extinction probability to evaluate the probability of success of antiviral therapy^{3,23}, assume the duration of the infectious phase, the period during which infected cells are producing and releasing virus progeny, to be exponentially distributed. Careful pairing of mathematical models and experimental measurements has established that the duration of the infectious phase in vitro for cells infected with IAV²⁷, simian HIV²⁸, or Ebola virus²⁹, follows a log-normal or normal-like distribution, and has clearly rejected the probability of an exponentially distributed infectious phase duration. Yan et al.⁵ estimated the extinction probability for a SM of IAV infection in vivo that allowed the lifespan of infectious cells to follow an Erlang distribution which, via its shape parameter (k), can capture exponential ($k = 1$), log-normal-like ($k \sim [1, 6]$), normal-like ($k > 10$), and even Dirac delta-like ($k \rightarrow \infty$) distributions. Yan et al.⁵ have shown that increasing the shape parameter (k) from exponential to log-normal leads to a decrease in the extinction probability given an infection initiated with a number of infectious virions. Hence, the distribution of the infectious phase duration is also expected to affect the likelihood that an infection will become established under antiviral therapy.

In this work, we construct a SM for virus infection through the physical consideration of key infection steps, rather than through a systematic, direct mathematical conversion of our mean-field model. Our SM explicitly represents the probability that a virion, after having successfully entered a cell, will fail to result in the productive infection of that cell. The SM is first used to estimate the extinction probability of an infection. Extending work by Czuppon et al.²³, we show that prophylactic therapy with an antiviral that blocks productive cell infection *after* virus entry, is better at reducing the establishment probability, than one reducing either the virus production rate or the rate of virus entry into cells. In addition, we investigate the difference in the fraction of cells consumed by so-called extinct versus established infections, and re-visit the comparison of antivirals through this new lens. Finally, we demonstrate how the antiviral efficacy required to achieve a desired infection extinction probability critically depends on the assumed distribution of the infectious phase duration.

Results

The mathematical models

The mean-field mathematical model (MFM), its stochastic counterpart, and associated infection parameters are illustrated in Fig. 2. Mathematically, the MFM and SM are given by

$$\begin{aligned}
 \frac{dT}{dt} &= -\gamma\beta TV/s & T^{t+1} &= T^t - N^{\text{inf}} \\
 \frac{dE_1}{dt} &= \gamma\beta TV/s - \frac{n_E}{\tau_E} E_1 & E_1^{t+1} &= E_1^t + N^{\text{inf}} - E_1^{\text{out}} \\
 \frac{dE_i}{dt} &= \frac{n_E}{\tau_E} E_{i-1} - \frac{n_E}{\tau_E} E_i & E_i^{t+1} &= E_i^t + E_{i-1}^{\text{out}} - E_i^{\text{out}} \quad i = 2, 3, \dots, n_E \\
 \frac{dI_1}{dt} &= \frac{n_E}{\tau_E} E_{n_E} - \frac{n_I}{\tau_I} I_1 & I_1^{t+1} &= I_1^t + E_{n_E}^{\text{out}} - I_1^{\text{out}} \\
 \frac{dI_j}{dt} &= \frac{n_I}{\tau_I} I_{j-1} - \frac{n_I}{\tau_I} I_j & I_j^{t+1} &= I_j^t + I_{j-1}^{\text{out}} - I_j^{\text{out}} \quad j = 2, 3, \dots, n_I \\
 \frac{dV}{dt} &= p \sum_{j=1}^{n_I} I_j - cV - \beta TV/s & V^{t+1} &= V^t + V^{\text{prod}} - V^{\text{decay}} - V^{\text{enter}}
 \end{aligned} \tag{1}$$

Initially, all cells are uninfected, susceptible target cells, T , i.e. $T(t = 0) = N_{\text{cells}}$, which are exposed to an initial dose of $V(t = 0) = V_0$ infectious virions. As target cells, T , encounter infectious virions, V , in the supernatant of a cell culture, a tissue, or an organ compartment of volume s , some become infected ($T \rightarrow E_1$). The rate of successful cell infections by infectious virions is $\gamma\beta V/s$, which depends on the concentration of infectious

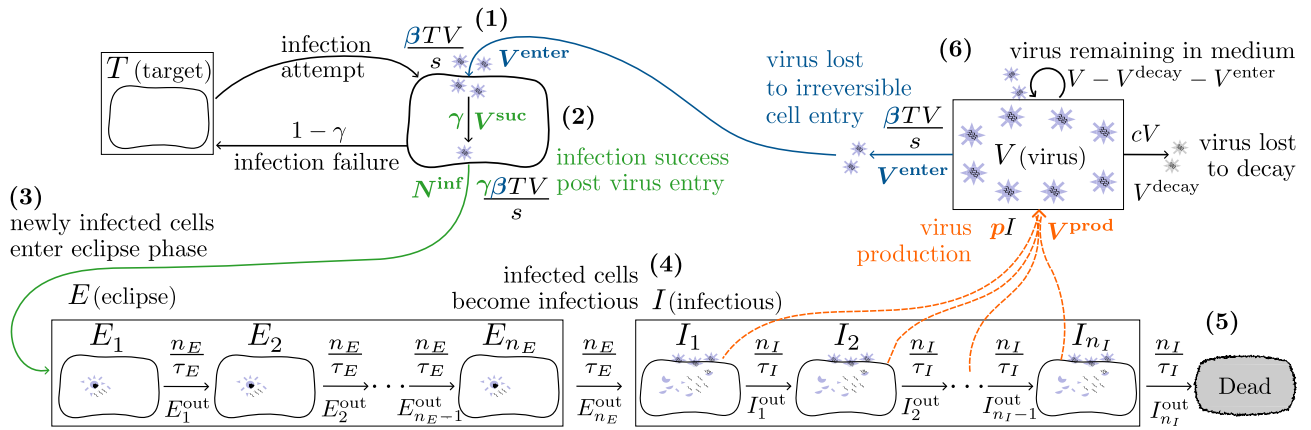


Figure 2. Representation of the mathematical models. (1) As uninfected, susceptible target cells (T) interact with infectious virions (V) over time, some portion of the latter (V^{enter}) irreversibly enter some of the available target cells, at rate $\beta TV/s$. (2) Out of these, a fraction $\gamma \in (0, 1]$ will cause the successful infection of N^{inf} target cells, while infection failure of the remaining fraction $(1 - \gamma)$ leaves cells in their uninfected state. (3) A newly infected cell will first enter the eclipse phase, and (4) then transition into the infectious phase during which it produces infectious virions at an average rate p , before (5) ultimately ceasing virus production and possibly undergoing apoptosis. (6) At each time step in the SM, a random number of newly produced infectious virions (V^{prod}) are released into the medium or organ compartment of volume s , increasing the infectious virus concentration already present, while V^{decay} infectious virions will lose infectivity, and another V^{enter} will be lost to irreversible cell entry. Terms like $\beta TV/s$, n_E/τ_E , n_I/τ_I , pI , and cV represent the rates at which events take place in the MFM, whereas terms like V^{enter} , E_j^{out} , I_j^{out} , V^{prod} , and V^{decay} correspond to the random number of such events over one time step of the SM (see Table 1).

virions V/s . When a cell becomes infected, it enters ($T \rightarrow E_1$) and traverses ($E_1 \rightarrow E_2 \rightarrow \dots \rightarrow E_{n_E}$) the n_E compartments of the eclipse phase, during which it is infected but not yet producing infectious virions. The infected cell then enters ($E_{n_E} \rightarrow I_1$) and traverses ($I_1 \rightarrow I_2 \rightarrow \dots \rightarrow I_{n_I}$) the infectious phase, during which it produces infectious virions at constant rate p . When an infected cell leaves the last compartment (I_{n_I}), it ceases virus production and thus ceases to contribute to the infection kinetics, and possibly undergoes apoptosis. The exponentially distributed durations of the n_E eclipse (or n_I infectious) phase compartments together yield an Erlang distributed total duration for the eclipse (or infectious) phase of mean duration τ_E (or τ_I), and standard deviation $\tau_E/\sqrt{n_E}$ (or $\tau_I/\sqrt{n_I}$), where n_E (or n_I) corresponds to the shape parameter of the Erlang distribution. These compartments are not meant to correspond to particular biological states. Rather, they offer a mathematically and computationally expeditious way (compared to delayed, partial or integro-differential equations^{27,30}) to implement biologically realistic durations for the time spent by infected cells in the eclipse and infectious phases (e.g., normal and lognormal distributions^{28,29,31}).

Infectious virions (V) are produced at a constant rate of p per infectious cell, and are lost either through loss of infectivity at rate c , or irreversible commitment to entry into susceptible cells at rate $\beta T/s$. Here, $\beta T/s$ is not meant to represent the rate of virus loss due to cell attachment because attachment is a reversible process: virions can still either detach and re-join the pool of virions in the extracellular space or irreversibly enter the cell. Instead, given the ongoing kinetics of virus attachment/detachment, $\beta T/s$ represents the net rate at which virions irreversibly commit to cell entry, i.e. pass the point of no return from which they cannot re-enter the extracellular space. For brevity, we will refer to β hereafter simply as the rate of virus loss due to cell entry.

This MFM is similar to that widely validated and applied to analyze and predict the course of in vitro infections with IAV^{32–34}, simian HIV²⁸, Ebola virus²⁹, RSV³⁵ and rotavirus³⁶. It differs from our previous MFM by explicitly accounting for the loss of infectious virions due to cell entry (term $-\beta TV/s$ in the dV/dt equation), previously considered by others^{2,5,37,38}. This loss of infectious virions due to cell entry causes a corresponding loss of uninfected target cells becoming infected ($T \rightarrow E_1$) at rate $\gamma\beta TV/s$. Parameter γ therefore has units of cell per infectious virion (IV), where $\gamma \in (0, 1]$ cell/IV, and corresponds to the average fraction of infectious virion entries into a cell that results in the successful, productive infection of that cell ($T \rightarrow E_1$). This additional parameter γ distinguishes our work from past work^{2,5,37,38} where it has been assumed that the loss of one infectious virion due to cell entry always inevitably leads to the successful infection of one cell, i.e. $\gamma = 1$ infected cell per infectious virion entry. We use this assumption as our base parameter value for γ throughout the work presented herein.

Biologically, γ accounts for several different causes of infection failure after a virion has irreversibly committed to entry into a cell, e.g. Fig. 1. It can represent semi-infectious virions that are entry-competent, but are defective in their ability to complete another downstream step, e.g. are missing one or more viral genome segment or have deleterious genetic mutations. It can also represent fully infectious virions that are not defective but rather, through random chance, fail to achieve a key step they could have functionally achieved, e.g. 50% of influenza A virions failing to fuse with endosome membrane following cell entry^{16,39}. Herein it is assumed that when an infectious virion enters a cell, the cell will either be successfully infected ($T \rightarrow E_1$) with probability γ , or remain

| Random variable | Random number generator |
|---|---|
| E_i^{out} | Binomial($n = E_i^t$, $p_E = \Delta t \cdot n_E / \tau_E$) where $i = 1, 2, \dots, n_E$ |
| I_j^{out} | Binomial($n = I_j^t$, $p_I = \Delta t \cdot n_I / \tau_I$) where $j = 1, 2, \dots, n_I$ |
| V^{prod} | Poisson($\lambda = \Delta t \cdot p \sum_{j=1}^{n_I} I_j^t$) |
| V^{decay} , V^{enter} , otherwise | Trinomial($n = V^t$, $p_1 = \Delta t \cdot c$, $p_2 = \Delta t \cdot \beta T^t / s$, $p_3 = 1 - p_1 - p_2$) |
| N^{inf} | $ \{x_i x_i \in \mathcal{U}\{a = 1, b = T^t\}, 1 \leq i \leq V^{\text{suc}}\} ^\dagger$ where $V^{\text{suc}} = \text{Binomial}(n = V^{\text{enter}}, p_V = \gamma)$ |

Table 1. Random variables of the stochastic model. $^\dagger N^{\text{inf}}$ is equal to the cardinality, i.e. the number of unique elements, in the set of V^{suc} random numbers drawn from the discrete uniform distribution over the interval $[1, T^t]$.

uninfected (T) with probability $1 - \gamma$. Failed infections are assumed to have no significant effect, leaving the uninfected cell exactly as they found it in state T . Upon infection, we assume the cell becomes instantaneously unavailable for co-infection, i.e. virions are lost through entry into target cells only, and cannot enter infected cells. As such, we do not attempt to represent the case of co-infection of a cell by semi-infectious, segmented RNA virions, each with an incomplete set of functional genome segments, which together have a complete set and are thus able to cause a productive cell infection^{40,41}. We also do not include defective interfering particle production⁴². Since our work focuses on infections starting from very few infectious virions, where randomness plays a meaningful role, the likelihood of cell co-infection and these associated effects should be negligible.

The SM is largely analogous to the MFM, where SM variables, e.g. T^t the number of target cells at time t , each denoted with a superscript t , are whole numbers. The remaining terms, corresponding to changes in these variables, are random whole numbers drawn at each time step from distributions thought to best represent the corresponding underlying biological process (see [Methods](#)). Table 1 summarizes how each random variable is generated.

Important biological quantities

Following Pearson et al.², let us define $\mathcal{P}_{V \rightarrow \text{Extinction}}$ and $\mathcal{P}_{I \rightarrow \text{Extinction}}$ as the probability of infection extinction given an infection initiated with either only one infectious virion or only one infectious cell, respectively. We now rewrite the derivation from Pearson et al.² for these quantities, in the context of our SM and its associated parameter notation. Unless otherwise stated, equivalent equations can be found in Pearson et al.²

The extinction probability given any initial number V_0 of infectious virions and I_0 of infectious cells is $(\mathcal{P}_{V \rightarrow \text{Extinction}})^{V_0} \cdot (\mathcal{P}_{I \rightarrow \text{Extinction}})^{I_0}$. If initially there is only one infectious virion, the infection can fail to spread if either that initial infectious virion fails to cause a successful cell infection with probability $1 - \mathcal{P}_{V \rightarrow I}$, or if it does cause a successful cell infection with probability $\mathcal{P}_{V \rightarrow I}$ but that cell infection subsequently leads to extinction with probability $\mathcal{P}_{I \rightarrow \text{Extinction}}$. We can therefore write

$$\mathcal{P}_{V \rightarrow \text{Extinction}} = (1 - \mathcal{P}_{V \rightarrow I}) + \mathcal{P}_{V \rightarrow I} \cdot \mathcal{P}_{I \rightarrow \text{Extinction}} \quad (2)$$

If initially there is only one infectious cell, the cell will produce m infectious virions over its lifespan with probability $\mathcal{P}_{I \rightarrow mV}$. Each one of these produced infectious virions can be treated as an independent infection event such that,

$$\mathcal{P}_{I \rightarrow \text{Extinction}} = \sum_{m=0}^{\infty} \mathcal{P}_{I \rightarrow mV} \cdot (\mathcal{P}_{V \rightarrow \text{Extinction}})^m \quad (3)$$

Taken together, the extinction probability given an infection initiated with only one infectious virion, $\mathcal{P}_{V \rightarrow \text{Extinction}}$, yields the implicit relation

$$\mathcal{P}_{V \rightarrow \text{Extinction}} = (1 - \mathcal{P}_{V \rightarrow I}) + \mathcal{P}_{V \rightarrow I} \cdot \sum_{m=0}^{\infty} \mathcal{P}_{I \rightarrow mV} \cdot (\mathcal{P}_{V \rightarrow \text{Extinction}})^m \quad (4)$$

The probability that an infectious virion is successful at causing a productive cell infection, $\mathcal{P}_{V \rightarrow I}$, can be expressed as the ratio between the rate of successful cell infection per infectious virion $\gamma \beta N_{\text{cells}} / s$ and the rate of virion loss through loss of infectivity plus cell entry, $c + \beta N_{\text{cells}} / s$ (see [Methods](#) for derivation), namely,

$$\mathcal{P}_{V \rightarrow I} = \frac{\gamma \beta N_{\text{cells}} / s}{c + \beta N_{\text{cells}} / s} = \frac{\gamma}{1 + c / (\beta N_{\text{cells}} / s)} \quad (5)$$

This expression reduces to that presented in Pearson et al.² when one assumes that the loss of one infectious virion due to cell entry always results in the infection of one cell, i.e. $\gamma = 1$ cell/IV.

The probability that an infectious cell produces m infectious virions over its lifespan, $\mathcal{P}_{I \rightarrow mV}$, can be expressed as the marginal probability distribution of the probability that an infectious cell produces m infectious virions given a lifespan of length t , Poisson($m | \lambda = pt$), and the probability that the lifespan is of length t , Erlang($t | k = n_I, \lambda = n_I / \tau_I$). Therefore,

$$\begin{aligned} \mathcal{P}_{I \rightarrow mV} &= \int_0^\infty \text{Poisson}(m|\lambda = pt) \cdot \text{Erlang}(t|k = n_I, \lambda = n_I/\tau_I) dt \\ &= \frac{(m+r-1)!}{m!(r-1)!} (1-p_B)^r (p_B)^m = \text{NB}(m|r = n_I, p_B = \mathcal{B}/(n_I + \mathcal{B})) \end{aligned} \quad (6)$$

where NB stands for the negative binomial (or Pascal) distribution and describes the probability that m successes will have occurred by the time one has observed $r = n_I$ failures from a series of independent Bernoulli trials with a probability of success p_B (see [Methods](#) for derivation). The mean of the distribution corresponds to the average burst size, $\mathcal{B} = p\tau_I$, i.e. the average number of infectious virions produced by a productively infected cell over its lifespan, with p the rate of infectious virion production per cell and τ_I the average infectious cell lifespan. This expression for $\mathcal{P}_{I \rightarrow mV}$ is novel, and it does not have a corresponding equation in Pearson et al.²

Combining the mean number of infectious virions produced by a productively infected cell, \mathcal{B} , and the probability that each of them causes the infection of a cell, $\mathcal{P}_{V \rightarrow I}$ (Eq. 5), we obtain the average basic reproductive number (R_0), defined as the average number of productive secondary infections caused by an infected cell over its lifespan when placed within a fully susceptible and uninfected cell population,

$$R_0 = \mathcal{B} \cdot \mathcal{P}_{V \rightarrow I} = p\tau_I \cdot \frac{\gamma\beta N_{\text{cells}}/s}{c + \beta N_{\text{cells}}/s} = p\tau_I \cdot \frac{\gamma}{1 + c/(\beta N_{\text{cells}}/s)}. \quad (7)$$

Substituting Eqs. (5) and (6) into Eq. (4) results in the following expression for $\mathcal{P}_{V \rightarrow \text{Extinction}}$

$$\mathcal{P}_{V \rightarrow \text{Extinction}} = \left[1 - \frac{\gamma}{1 + c/(\beta N_{\text{cells}}/s)} \right] + \frac{\gamma}{1 + c/(\beta N_{\text{cells}}/s)} \left[\frac{\mathcal{B}(1 - \mathcal{P}_{V \rightarrow \text{Extinction}})}{n_I} + 1 \right]^{-n_I} \quad (8)$$

While Eq. (8) does not lead to an analytical solution for $\mathcal{P}_{V \rightarrow \text{Extinction}}$, a solution can easily be obtained numerically (see [Methods](#)). The probability of infection establishment given an infection initiated with only one infectious virion, $\mathcal{P}_{V \rightarrow \text{Establishment}}$, is simply $1 - \mathcal{P}_{V \rightarrow \text{Extinction}}$.

Effectiveness of antivirals to reduce an infection's establishment probability

Prophylactic antiviral therapies are often characterized and compared in terms of their ability to reduce an infection's establishment probability. This is because for natural infections in a host, the initial virus inoculum can be sufficiently small that prophylactic antiviral therapy can prevent infection, i.e. induce infection extinction.

Recently, Czuppon et al.²³ used a SM to evaluate the effectiveness of prophylactic treatment with antivirals reducing the rate of virus entry into cells sometimes called the virus infectivity rate, herein β , or the virus production rate, p , to reduce the establishment probability for a SARS-CoV-2 infection in vivo initiated by a small number of infectious virions. But an endosomal fusion inhibitor (e.g. cathepsin L inhibitors for SARS-CoV-2²⁶) would not affect the rate of removal of infectious virions from the medium (i.e. it leaves β unaffected), but by enhancing fusion failure post cell entry, it would prevent the cell from progressing to a productively infectious state, hence blocking 100% of virion progeny production. Using our SM, which can explicitly represent productive infection failure post virus entry via parameter γ , we extend Czuppon's investigation to evaluate how such an antiviral would compare against those reducing β or p . Fig. 3 illustrates the different antiviral modes of action considered herein.

Figure 4 shows the establishment probability given an infection initiated with only one infectious virion as a function of the efficacy (ε) for antivirals reducing β , p or γ . As antiviral efficacy ε increases, the establishment probability decreases at first slowly and then abruptly as it approaches zero (as the basic reproductive number, R_0 , approaches one). As others before us^{3,23}, we assume that these antivirals have negligible cytotoxic effects, i.e. do not cause cell damage or death, over the range of concentrations (efficacies) explored, and that all antivirals can achieve all efficacies. In reality, antivirals with different mechanisms of action would likely differ in their cytotoxicity profile, and in the maximum efficacy they can achieve following the E_{max} model⁴³, which relates drug concentration to efficacy. The parameter values used to generate this figure and all other figures herein are provided in the [Methods](#). They were taken from Czuppon et al.²³ for SARS-CoV-2 infections in patients, and in particular assume an exponentially distributed duration for the infectious phase ($n_I = 1$). For $n_I = 1$, the expression for the establishment probability in Eq. (8) reduces to

$$\mathcal{P}_{V \rightarrow \text{Establishment}} = \mathcal{P}_{V \rightarrow I} - \frac{1}{\mathcal{B}} = \frac{R_0}{\mathcal{B}} - \frac{1}{\mathcal{B}} = \frac{\gamma}{1 + c/(\beta N_{\text{cells}}/s)} - \frac{1}{p\tau_I} \quad (9)$$

where $\mathcal{B} = p\tau_I$ is the burst size (see [Methods](#)). Since β and c appear as a ratio, $c/(\beta N_{\text{cells}}/s)$, in the expression for the establishment probability (Eq. (9)), an antiviral reducing β and increasing c would have the same effect on the establishment probability as one only reducing β with greater efficacy. This is particularly interesting because many antivirals that reduce the ability of infectious virions to enter cells (reduce β), such as monoclonal antibodies, often target the virions directly by binding to and disabling their surface protein, thus also enhancing the rate at which virions lose infectivity (increasing c).

At relatively low efficacy, e.g. $\varepsilon = 0.8$ (see line ① in Fig. 4), for the parameters chosen, an antiviral reducing γ is more effective than one reducing β , which is more effective than one reducing p at reducing the establishment probability. The infection parameters used by Czuppon et al.²³ correspond to a sufficiently large burst size ($\mathcal{B} = p\tau_I = 18.8$ IV/cell, see [Methods](#)) such that $\mathcal{P}_{V \rightarrow \text{Establishment}} \approx \mathcal{P}_{V \rightarrow I}$ ($= 0.409$). As such, the bottleneck to infection establishment is the probability that the single initial infectious virion causes a productive cell infection, $\mathcal{P}_{V \rightarrow I}$, which depends on β and γ , but not p . While an antiviral reducing β reduces both the numerator

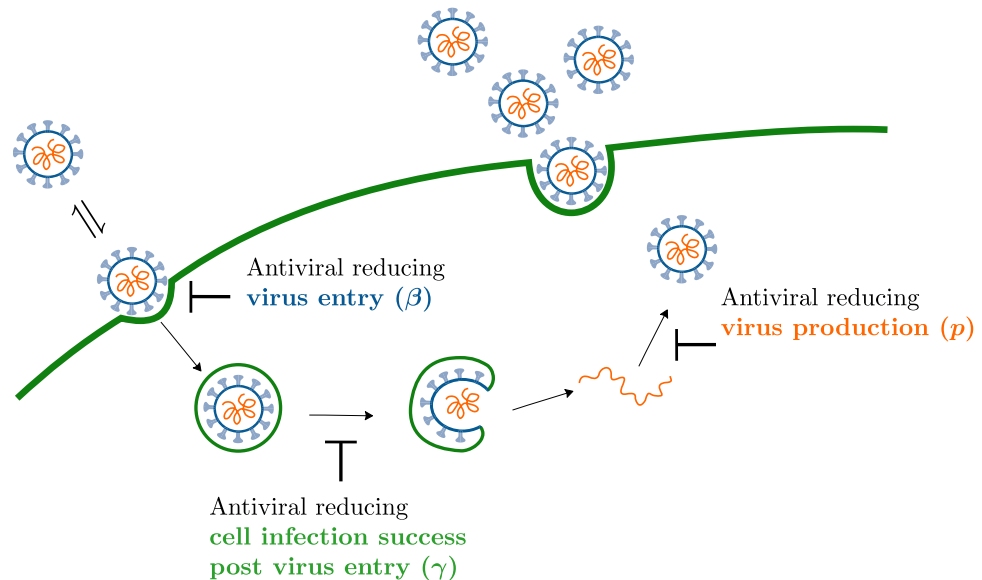


Figure 3. Antiviral modes of action. Antiviral reducing (left) the virus entry rate β , (centre) the probability of a successful cell infection post virus entry γ , or (right) the virus production rate p . If an antiviral reducing β prevents an infectious virion from entering a cell, the virion could try again later on and succeed. Whereas, if an antiviral reducing γ prevents an infectious virion from completing a successful cell infection post cell entry, the virion would be removed from the overall infection. Also, an antiviral reducing p would reduce virus production rather than totally block it, as would be the case if an antiviral reducing γ prevents a productive cell infection. These differences between the modes of actions previously studied (reducing β or p) and newly considered herein (reducing γ) highlight why the latter might prove advantageous.

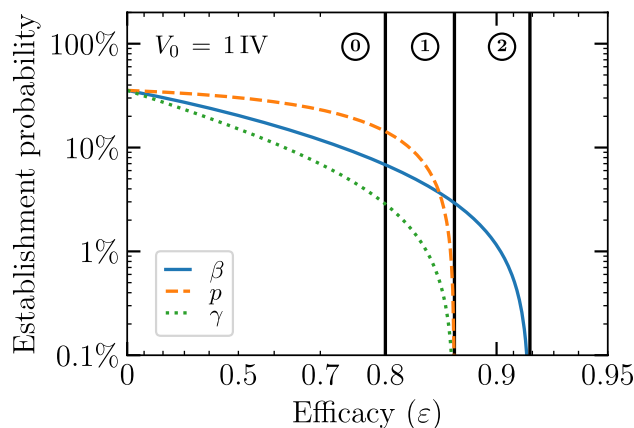


Figure 4. Effect of different antivirals on the probability of infection establishment. The establishment probability given an infection initiated with only one infectious virion as a function of efficacy (ϵ) for antivirals acting either to reduce the virus entry rate, $\beta \rightarrow (1 - \epsilon)\beta$, the virus production rate, $p \rightarrow (1 - \epsilon)p$, or the probability of a successful cell infection post virus entry, $\gamma \rightarrow (1 - \epsilon)\gamma$. The line labelled ① indicates a value of lower efficacy ($\epsilon = 0.8$), at line ① the establishment probability goes to zero ($R_0 = 1$) for an antiviral reducing p or γ ($\epsilon \approx 0.87$), and at ② it goes to zero for an antiviral reducing β ($\epsilon \approx 0.92$). The infection parameters are from Czuppon et al.²³ (see [Methods](#)).

and denominator of $\mathcal{P}_{V \rightarrow I}$, one reducing γ reduces only the numerator and therefore has a greater effect. At this relatively low antiviral efficacy ($\epsilon \approx 0.8$), as the initial number of infectious virions increases and as the establishment probability approaches 100%, differences in the efficacy of antivirals reducing p , β or γ to reduce the establishment probability vanishes (see Supplementary Material, Section S2).

As the antiviral efficacy is increased further, it eventually reaches the critical point where $R_0 = 1$, and the establishment probability, $(R_0 - 1)/\mathcal{B}$, equals zero. The value of the antiviral efficacy at which this is achieved is the same for antivirals reducing p or γ (line ① in Fig. 4) since they affect R_0 (Eq. (7)) identically. The efficacy of an antiviral reducing β must be much higher in order to reach this critical point (line ② in Fig. 4), because reducing β reduces both the numerator and denominator of R_0 .

Therefore, as Czuppon et al.²³ before us, and for their choice of parameters, we find that at lower efficacies, an antiviral reducing β is better than one reducing p at reducing the establishment probability, whereas at higher efficacies it is the opposite. However, here we show that an antiviral reducing γ at any efficacy is best, better than one reducing p or β , at reducing the establishment probability.

Distinction between infection extinction and establishment

The mathematical expression for the probability of extinction of an infection, $\mathcal{P}_{V \rightarrow \text{Extinction}}$, gives the probability, starting from a single infectious virion, that this virion fails to infect a cell, or should it succeed, the probability that this cell's virion progeny all go on to fail, or should any succeed, the probability that the cells they infect in turn produce virions that all fail, or should any succeed ..., and so on. The recursive nature of this expression makes it difficult to accurately define what constitutes the establishment or extinction of an infection in biological terms. Can one look at a given biological or SM infection outcome and classify it as an extinct versus an established infection?

Let us look at the variety of infection outcomes that the mathematical expression, $\mathcal{P}_{V \rightarrow \text{Extinction}}$, expects will fall into two categories: established or extinct infections. Figure 5 shows the frequency and complementary cumulative frequency distributions, or $1 - (\text{cumulative frequency})$, for the number of cells consumed by 10^6 SM simulated infections. The SM infection outcomes neatly separate into 2 distinct categories: infections that consumed none or very few cells (close to 0%, on the left); and those that consumed many (around 80%, on the right), which we will assume correspond to infections said to have gone extinct or to have established, respectively, as indicated by red arrows. The presumed extinct infections make up 85% (848,227) of the 10^6 SM simulations, which matches the theoretically predicted probability of infection extinction, $\mathcal{P}_{V \rightarrow \text{Extinction}}$, for the infection parameters used (see Methods). Of these, 69% (59%/85%) result in no cells consumed, 93% ((59 + 16 + 4)%/85%) in fewer than 3 cells consumed, and 95% in fewer than 5 cells consumed, or less than 0.01% of all susceptible cells. Of the presumed established infections, all consumed at least 30,600 cells or about 10,000 \times more than infections considered extinct. The presumed extinct infections are unlikely to engage a measurable immune response nor trigger symptoms, whereas the presumed established infections, consuming 80% of cells, certainly would. In this case, the difference between extinction and establishment is statistically and biologically significant. While 85% of patients would essentially avoid infection altogether, the other 15%, (around 1 in 6) would experience substantial infections.

We show in the Methods (see Eq. (40)) that the median fraction of cells consumed by established infections in the SM monotonically decreases as one increases the quantity T^*/N_{cells} , where

$$\frac{T^*}{N_{\text{cells}}} = \frac{c/(\beta N_{\text{cells}}/s)}{\gamma p \tau_I - 1}. \quad (10)$$

For the presumed established infections in Fig. 5, Eq. (40) based on T^*/N_{cells} predicts that $\sim 80\%$ of cells (31,873 out of 40,000) will be consumed, indicated by a vertical solid line in the figure. Biologically, the quantity T^*/N_{cells} is not particularly meaningful, but mathematically it shows how the median fraction of cells consumed by established infection depends on and is sensitive to each parameter.

Figure 6 shows the frequency and complementary cumulative frequency distributions of the number of cells consumed by an infection, using parameter sets where T^*/N_{cells} is increasingly closer to 1. It is achieved here by increasing the efficacy of an antiviral reducing the virus production rate, $p \rightarrow (1 - \varepsilon)p$, thus increasing T^*/N_{cells} .

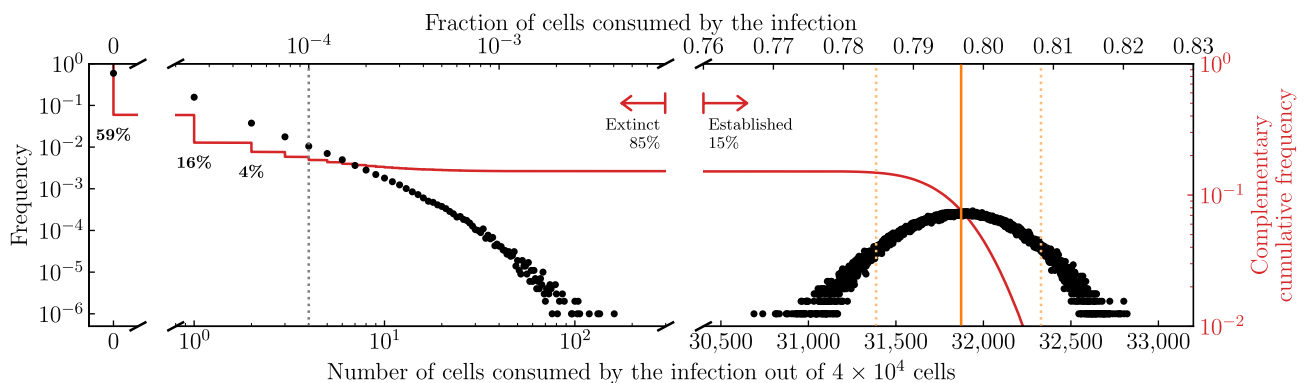


Figure 5. Infection extinction versus establishment. The distribution for the frequency (black dots) or complementary cumulative frequency (solid red curve), or $1 - (\text{cumulative frequency})$, of the number of cells consumed by the infection, $N_{\text{cells}} - T(\infty)$, out of $N_{\text{cells}} = 4 \times 10^4$ cells, based on 10^6 SM simulations. For the frequency distribution, each black dot represents the fraction of the SM simulations where exactly this discrete number of cells were consumed by the infection (e.g. 0, 1, 2, ...). The 95th-percentile upper and lower bounds (95% bounds) for the number of cells consumed by extinct infections is between zero and the vertical (grey) dotted line, and that by established infections is indicated by a pair of vertical (orange) dashed lines flanking the median (vertical solid line). Parameters are the same as in Fig. 4 for an antiviral reducing the virus production rate, $p \rightarrow (1 - \varepsilon)p$, at efficacy $\varepsilon = 0.79$ such that $T^*/N_{\text{cells}} = 0.5$ (see Methods).

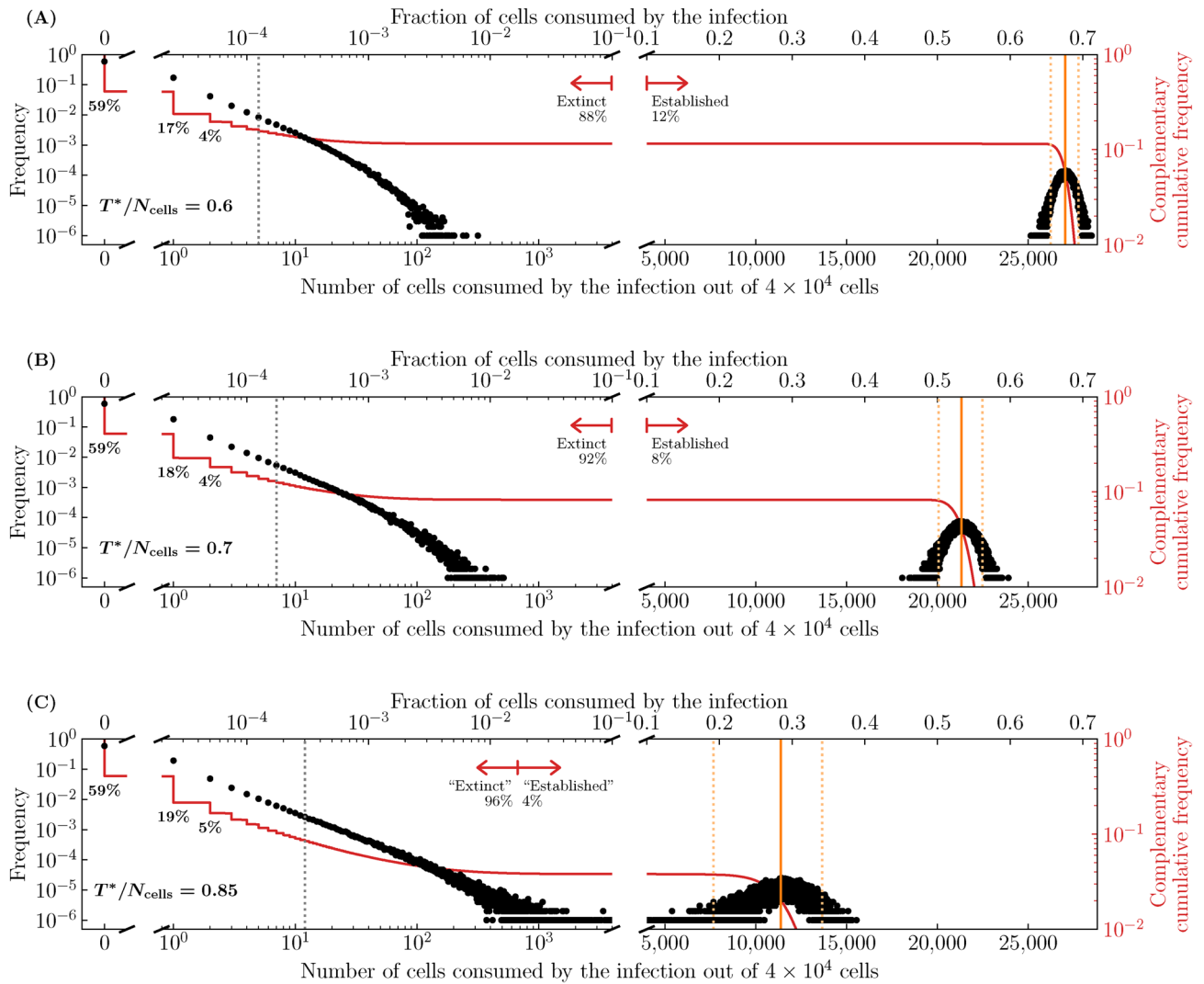


Figure 6. Effect of infection parameters on extinction and establishment. The frequency or complementary cumulative frequency of the fraction or number of cells consumed by the infection out of $N_{\text{cells}} = 4 \times 10^4$ cells where the efficacy (ε) of an antiviral reducing the virus production rate, i.e. $p \rightarrow (1 - \varepsilon)p$, was varied such that (A) $T^*/N_{\text{cells}} = 0.6$; (B) 0.7 or (C) 0.85. Everything else is generated, computed, and represented visually as described in the caption of Fig. 5.

As T^*/N_{cells} is increased from 0.6 to 0.85, the median number of cells consumed by infections presumed extinct increases slightly, while that by established infections decreases from $\sim 68\%$ down to $\sim 28\%$, and the distinction blurs as the two distributions begin to merge. In Fig. 6C, the red arrow demarcating so-called extinct and established infections is located so that $\sim 96\%$ of the 10^6 SM simulated infections are to the left. Mathematically, this location corresponds to the expected fraction of extinct infections, $\mathcal{P}_{V \rightarrow \text{Extinction}}$, but the difference in the fraction of cells consumed by established and extinct infections is becoming less biologically meaningful. Therefore, when an antiviral is applied and the establishment probability is non-zero ($R_0 > 1$), varying the antiviral efficacy affects not only the establishment probability but also the fraction of cells consumed by both established and extinct infections, and likely also the intensity of the immune response and symptoms. For higher antiviral efficacies (higher T^*/N_{cells}), the biological distinction between extinct and established infections becomes increasingly meaningless.

The parameter values used in our SM are taken from Czuppon et al.²³ where the number of cells susceptible to SARS-CoV-2 (parameter N_{cells}) is assumed to be 40,000 cells, or around just 0.01% of the estimated 4×10^8 cells in the human upper respiratory tract. Yet in earlier work, some of the same authors estimate that a more plausible 1% of these cells (4 million cells) are susceptible to SARS-CoV-2⁴⁴. Mathematically, assuming a larger N_{cells} with a proportionally smaller β yields the same probability of infection extinction ($\mathcal{P}_{V \rightarrow \text{Extinction}}$) and the same fraction but not the same number of cells infected by established infections, because it leaves T^*/N_{cells} but not T^* unchanged. Figure 7 explores the fraction and number of cells consumed by infections as the size of the cell population, N_{cells} , is increased while keeping βN_{cells} fixed, and thus also the establishment probability and the fraction of cells consumed by established infections. In Fig. 7A, with the smallest cell population considered, the distributions for the number of cells consumed by extinct and established infections are merged, making the two

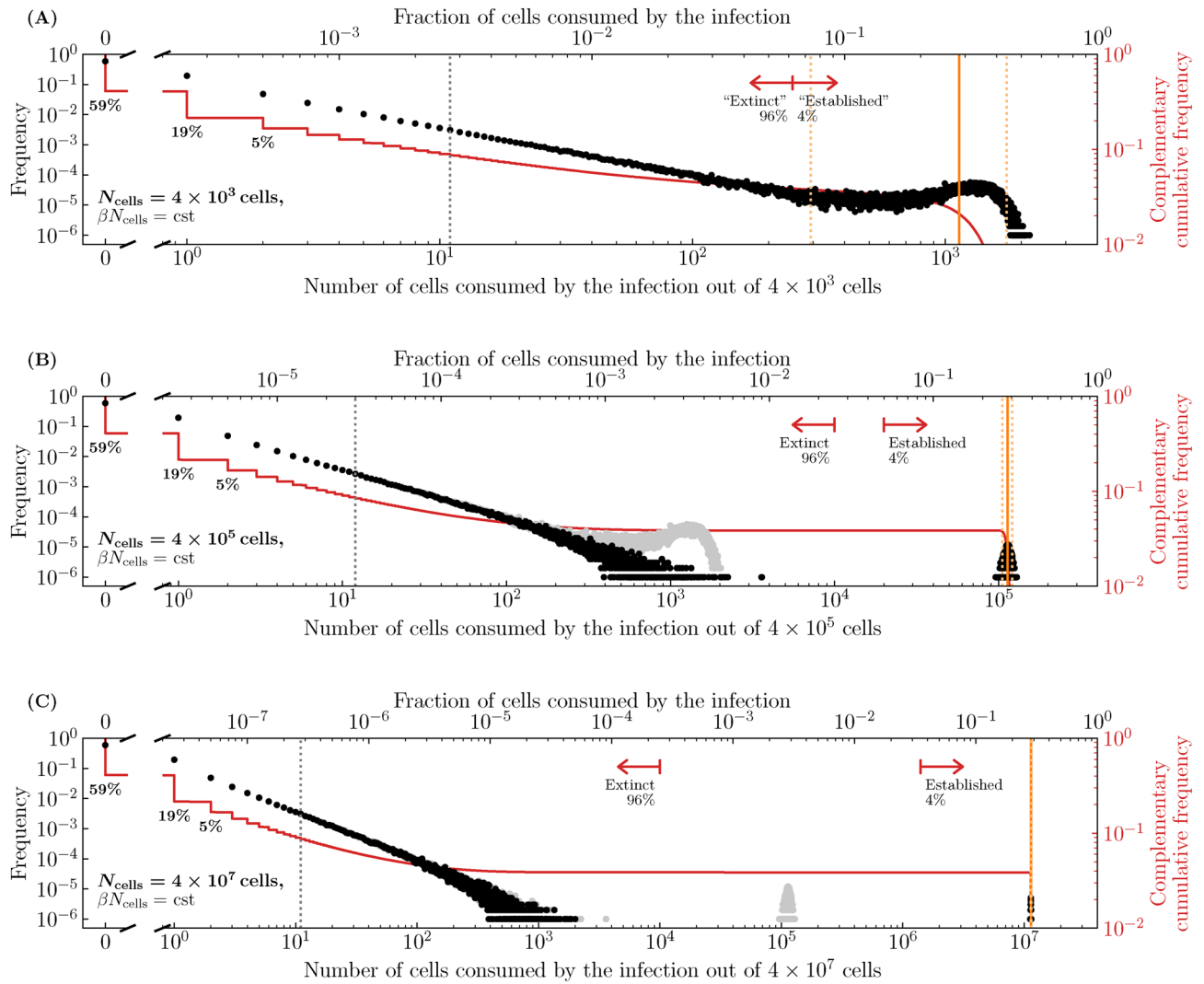


Figure 7. Effect of the susceptible cell population size on extinction and establishment. The frequency and complementary cumulative frequency of the fraction or number of cells consumed by the infection out of N_{cells} , as the latter is varied, while βN_{cells} , and therefore also the establishment probability and T^*/N_{cells} , are fixed; **(A)** $N_{\text{cells}} = 4 \times 10^3$ cells; **(B)** 4×10^5 cells or **(C)** 4×10^7 cells. The frequency distribution in **(A)** is shown in **(B)** in grey, and that in **(B)** is shown in **(C)** in grey. As in Fig. 6(C), the efficacy (ε) of an antiviral reducing the virus production rate, $p \rightarrow (1 - \varepsilon)p$, was such that $T^*/N_{\text{cells}} = 0.85$. Everything else is generated, computed, and represented visually as described in the caption of Fig. 5.

outcomes biologically indistinguishable. In Fig. 7B,C, as the size of the cell population (N_{cells}) is increased, the fraction of cells consumed by established infections remains centred at $\sim 28\%$ (because T^*/N_{cells} is unchanged), while the number of cells consumed (28% of N_{cells}) increases as N_{cells} increases, from 4,000 cells up to 40 million cells. Underestimating the number of host cells susceptible to infection, N_{cells} , can thus have profound implications for the expected severity and extent of immune engagement expected from so-called established infections.

Figure 8 explores more generally how the median and 95% bounds of the distribution for the number and fraction of cells consumed by established infections vary as a function of T^* and N_{cells} . Increasing T^*/N_{cells} from 0.5 to 0.94 decreases the median fraction of cells consumed from 80% down to 12%. Extrapolating from Fig. 8B (see Methods), this corresponds to about zero to 7,600 cells ($[0, 0.19] \times N_{\text{cells}}$) consumed by established infections for $N_{\text{cells}} = 4 \times 10^4$, sufficiently few that all extinct and established infections could possibly be asymptomatic. In contrast, for a susceptible cell population of $N_{\text{cells}} = 4 \times 10^6$, established infections would consume $\sim [0.11, 0.13] \times N_{\text{cells}}$, or between $\sim 440,000$ to 520,000 cells, which are unlikely to be asymptomatic infections. For larger (possibly more realistic) susceptible cell population sizes, the distinction between the fraction of cells consumed by extinct and established infections is greater, and a higher antiviral efficacy (higher T^*/N_{cells}) is required for the distinction to vanish.

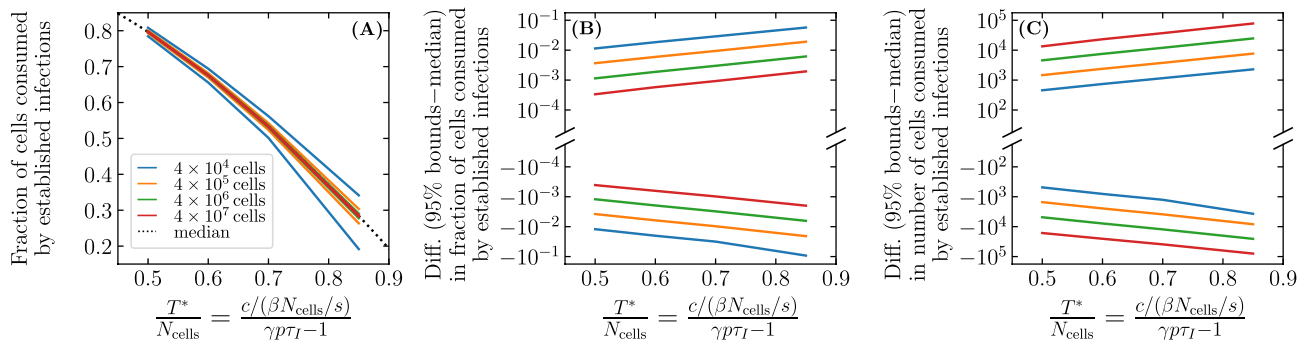


Figure 8. Effect of the susceptible cell population size on cells consumed by infection. The 95% bounds about the median (Eq. (40)) for the fraction and/or number of cells consumed by established infections, as a function of T^*/N_{cells} , for different susceptible cell population sizes (N_{cells}) while holding βN_{cells} fixed. The 95% bounds are based on 10^6 SM simulations of which no less than about 40,000 were established infections ($\mathcal{P}_{V \rightarrow \text{Establishment}}=4\%$ at $T^*/N_{\text{cells}} = 0.85$). Here, $T^*/N_{\text{cells}} = [c/(\beta N_{\text{cells}}/s)]/[\gamma p \tau_I - 1]$ is increased by increasing the efficacy of an antiviral reducing the virus production rate, $p \rightarrow (1 - \varepsilon)p$.

Reduction in the number of cells consumed by infections under antiviral therapy

As shown above, prophylactic antiviral therapy not only reduces an infection’s establishment probability, but also decreases the fraction of cells consumed by established infections, which decreases as $T^*/N_{\text{cells}} = [c/(\beta N_{\text{cells}}/s)]/[\gamma p \tau_I - 1]$ increases (see Methods, Eq. (40)). Figure 9A shows the fraction of cells consumed by established infection, given an initial inoculum of only one infectious virion, as a function of antiviral efficacy (ε) for antivirals reducing either β , p , or γ . At a given efficacy, ε , an antiviral reducing p or γ will increase T^*/N_{cells} identically because both appear in the same way in T^*/N_{cells} , whereas one reducing β will cause a smaller increase in T^*/N_{cells} . As such, an antiviral reducing p or γ will reduce the fraction of cells consumed by established infections more than one reducing β . Here again, β and c appear as a ratio in T^*/N_{cells} , as they did in the infection establishment probability, Eq. (9), such that an antiviral increasing c and/or decreasing β will have the same impact on reducing the fraction of cells consumed by established infections. Figure 9B illustrates how these differences, and therefore their biological relevance, depends on the antiviral efficacy under consideration. The difference is most pronounced as one approaches the line labelled ① from either side, which corresponds to the efficacy ($\varepsilon \approx 0.87$) at which $R_0 = 1$ for an antiviral reducing p or γ while still $R_0 > 1$ for one reducing β . The shaded areas on segments of the curves in Fig. 9 correspond to the 95% bounds, and can only be computed from the SM simulations over portions of the curve where established and extinct infections are clearly distinct. It shows that the SM variability is small relative to the difference in the fraction of cells infected.

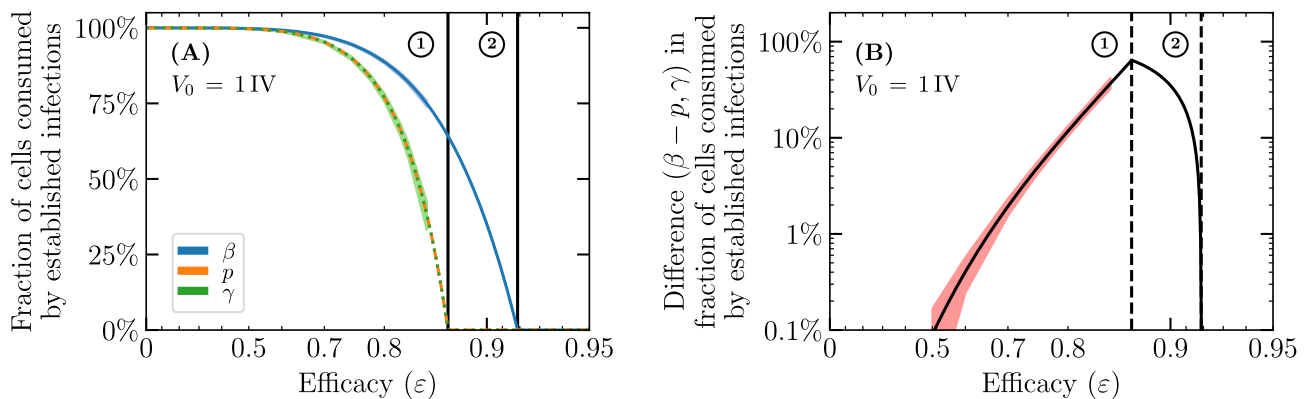


Figure 9. Effect of different antivirals on the fraction of cells consumed by established infections. (A) Fraction of cells consumed by established infections initiated with only one infectious virion as a function of efficacy (ε) for antivirals acting prophylactically either to reduce the virus entry rate, $\beta \rightarrow (1 - \varepsilon)\beta$, the virus production rate, $p \rightarrow (1 - \varepsilon)p$, or the probability of a successful cell infection post virus entry, $\gamma \rightarrow (1 - \varepsilon)\gamma$. (B) Difference in the fraction of cells consumed by established infections shown in (A) for an antiviral reducing β minus that for one reducing either p or γ . The vertical lines are the same as in Fig. 4 and indicate antiviral efficacies at which the establishment probability goes to zero ($R_0 = 1$) for an antiviral reducing p or γ (line ①); or β (line ②). The thin shaded areas visible over portions of the curves correspond to the 95% bounds about the median over (A) 10,000 SM simulations; or (B) 10,000 differences between pairs of SM simulations for an antiviral reducing either β or γ , drawn at random with replacement. Infection parameters are the same as in Fig. 4.

Figure 10 shows the number of cells consumed by extinct and established infections, given an initial inoculum of a single infectious virion, for antivirals reducing β , p or γ at an antiviral efficacy of $\varepsilon = 0.86$. This efficacy was chosen so that, given the infection parameters, the establishment probability is the same for an antiviral reducing β or p ($\mathcal{P}_{V \rightarrow \text{Establishment}} = 4\%$), yet it corresponds to a lower establishment probability for an antiviral reducing γ ($\mathcal{P}_{V \rightarrow \text{Establishment}} = 0.5\%$). This can be seen from Fig. 4: it corresponds to the efficacy where the orange line (antiviral reducing p) crosses the blue line (antiviral reducing β). At this efficacy, and for the chosen infection parameters, an antiviral reducing β compared to one reducing p or γ will result nearly $\sim 3\times$ more cells consumed by established infections, and likely engage a more intense immune response and symptoms. There could be clinical merit to weighing a reduction in the probability of infection establishment against a reduction in the severity of those infections that will become established.

We investigated a second parameter set also explored in Czuppon et al.²³, which is characterized by a 10-fold decrease in the number of susceptible cells N_{cells} , a 10-fold increase in the virus production rate p (and in the average burst size, $\mathcal{B} = p\tau_I$), and a corresponding ~ 10 -fold decrease in βN_{cells} (see Supplementary Material, Section S3). For this second parameter set, we found that an antiviral reducing γ was comparable to one reducing β , but better than one reducing p , for reducing the establishment probability, but all 3 modes of action reduced the fraction of cells consumed by established infections by the same amount at the same drug efficacy, ε . Since the degree to which an antiviral reducing γ is better than one reducing p or β depends on infection parameters, it is critical that such parameters be well-determined if such investigations are to yield useful predictions. More generally, we can show that, for infections initiated with one or a few infectious virions, an antiviral *reducing* γ

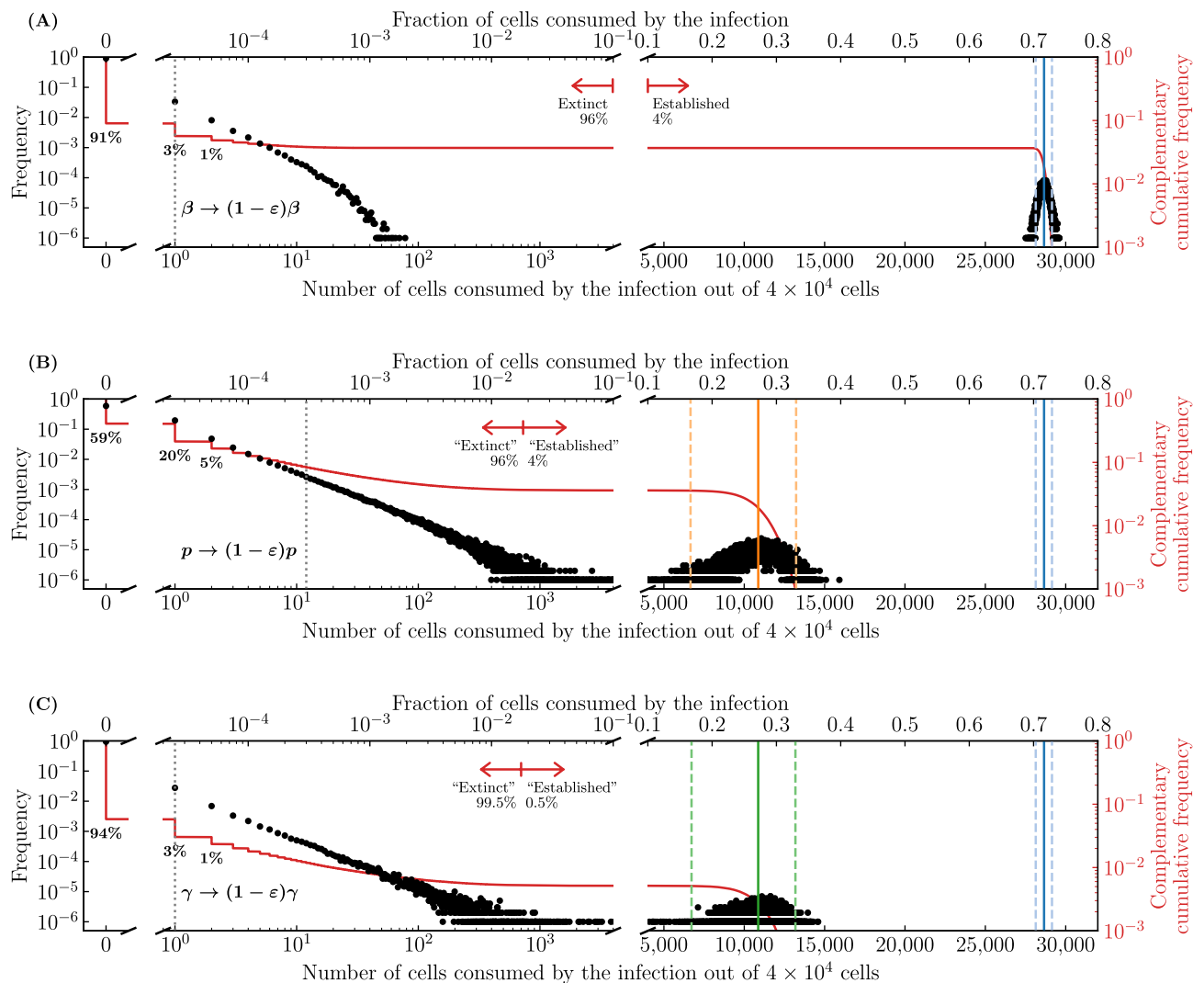


Figure 10. Antiviral with different modes of action cause different fraction of cells to be consumed by infections. The number of cells consumed by the infection out of $N_{\text{cells}} = 4 \times 10^4$ cells for antivirals with efficacy $\varepsilon = 0.86$ acting either to reduce (A) the virus entry rate, $\beta \rightarrow (1 - \varepsilon)\beta$, (B) the virus production rate, $p \rightarrow (1 - \varepsilon)p$, or (C) the probability of a successful cell infection post virus entry, $\gamma \rightarrow (1 - \varepsilon)\gamma$. Antiviral modes of action are represented by different colours: reducing β (blue), p (orange) or γ (green). Everything else is generated, computed, and represented visually as described in the caption of Fig. 5.

will always be better or at least as good than one reducing p or β at reducing the establishment probability and the fraction of cells consumed by established infections *irrespective of the parameters* (see [Methods](#)).

Lastly, as in Czuppon et al.²³, we also considered antiviral therapy for an infection initiated with only one infectious cell which could be representative of post-exposure antiviral therapy (see Supplementary Material, Section S4). In this case, we found that an antiviral reducing p or γ has the same effectiveness, greater than one reducing β , in reducing the establishment probability, which differs from the results reported above for an infection initiated with a single infectious virion. In contrast, we found that an antiviral reducing p or γ reduced the fraction of cells consumed by established infections more so than one reducing β at the same efficacy, which exactly matched the results for an infection initiated with only one infectious virion.

Overall, these results over different infection parameters and antiviral efficacy indicates that, under certain conditions, the fraction of cells consumed by both or either extinct and established infections can be an important consideration when evaluating and comparing antivirals with different modes of action. It also highlights the importance of properly identifying biologically relevant base parameter values in order to provide meaningful comparisons.

Continuous versus burst release of virus and the impact of the virus burst size distribution

In the MFM and SM used herein, virus is released continuously at a fixed rate of p infectious virion per hour by infectious cells. In the SM, this fixed rate maps to a Poisson distributed, discrete, stochastic number of infectious virions produced per time step by each infected cell over the duration of its infectious lifespan. The duration of this infectious phase is represented by an Erlang distribution characterized by shape parameter n_I and average duration τ_I . As shown in Eq. (6), this means that the SM's stochastic virus burst size follows a negative binomial or Pascal distribution with a mean corresponding to the MFM's virus burst size, $\mathcal{B} = p\tau_I$, where n_I now corresponds to the distribution's integer-valued, stopping-time parameter. Figure 11A illustrates how the SM's burst size distribution varies as a function of n_I , for a fixed average burst size.

Up to this point, as with most SMs, the duration of the infectious phase has been assumed to be exponentially distributed ($n_I = 1$ in Eq. (1)) and, hence, the burst size to be geometrically distributed (special case of the negative binomial distribution). With higher values of n_I , the burst size is represented by the more general case of the negative binomial distribution, and as $n_I \rightarrow \infty$, it approaches a Poisson distribution. The choice of n_I does not affect the fraction of cells consumed by established infections, since n_I does not appear in Eq. (40), but it does affect the extinction probability, as per Eq. (8). Previously, Yan et al.⁵ have shown that increasing n_I leads to a decrease in the extinction probability when infection is initiated with a small initial number of infectious virions. Figure 11B shows this for the infection parameters used so far for an infection initiated with a single infectious virion.

Recall that the extinction probability is an implicit expression with 2 main terms,

$$\mathcal{P}_{V \rightarrow \text{Extinction}} = (1 - \mathcal{P}_{V \rightarrow I}) + \underbrace{\mathcal{P}_{V \rightarrow I} \cdot \sum_{m=0}^{\infty} \text{NB} \left(m \mid r_{\text{fail}} = n_I, p_{\text{success}} = \frac{\mathcal{B}}{n_I + \mathcal{B}} \right)}_{\text{burst size distribution}} \cdot (\mathcal{P}_{V \rightarrow \text{Extinction}})^m.$$

where the first term, $1 - \mathcal{P}_{V \rightarrow I}$, is the probability that the initial virion inoculum fails to productively infect a single cell, and the second is the likelihood that it does actually cause a cell infection which then itself fails to establish an infection. Figure 11C shows how the contribution from the second term, i.e. the extinction

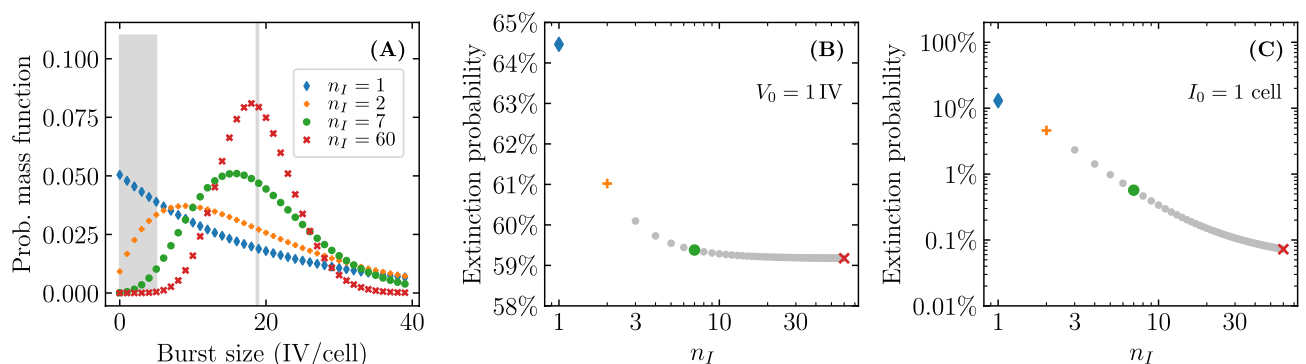


Figure 11. Effect of the infectious phase duration's shape parameter, n_I , on the extinction probability. (A) Negative binomial distributed virus burst size as a function of the shape parameter (n_I) of the Erlang distributed infectious cell lifespan. The shaded region highlight a region where the burst size is small (≤ 5 infectious virions) and therefore more likely to result in infection extinction. The vertical grey line corresponds to the average burst size value ($\mathcal{B} = p\tau_I = 18.8 \text{ IV/cell}$), which is constant as n_I is varied. (B,C) The probability of infection extinction as a function of n_I , given an infection initiated with only one infectious virion (B); or one infectious cell (C). Note that the extinction probability is shown on a linear scale in (B) but a logarithmic scale in (C) to better capture the relationships. Unless otherwise specified, the parameters are the same as in Fig. 4.

probability for an infection initiated with a single infected cell, decreases as n_I increases. What is observed in Fig. 11B, therefore, is the shrinking contribution to $\mathcal{P}_{V \rightarrow \text{Extinction}}$ by the second term as n_I increases.

This second term depends critically on the burst size distribution which, in our SM, is a result of the Erlang distributed infectious phase duration. As the latter distribution goes from exponential ($n_I = 1$) to fat-tailed ($n_I \in [2, 7]$), to normal-like ($n_I = 60$), tending ultimately to a Dirac delta distribution ($n_I \rightarrow \infty$), it becomes less probable that an infectious cell will have a very small burst size, as shown in Fig. 11A. For example, the probability that an infectious cell will have a burst size less than or equal to 5 infectious virions (shaded region in Fig. 11A) is 27% when $n_I = 1$ compared to 0.07% when $n_I = 60$ ($> 380 \times$ less likely). This is why, as n_I increases, the probability of infection extinction once one cell has been infected decreases, and this effect is more pronounced for larger virus burst sizes.

Work by Pearson et al.², expanded in Conway et al.³, concluded that the continuous release of infectious virions throughout an infected cell's lifespan leads to a higher probability of infection extinction than when virions are released as a single burst at the end of a cell's lifespan^{45–48}. In the authors' continuous production model, infected cells release virions at a constant rate over the duration of their infectious lifespan (Fig. 12A), while in their burst model, virions are released all at once, as a single burst, upon an infected cell's death (Fig. 12B). The total number of virions thus produced (burst size) is geometrically distributed in their continuous production model (Fig. 12C), and Kronecker delta distributed (a fixed value) in their burst model (Fig. 12D). As such, two attributes were simultaneously varied and conflated: the timing of virus production by infected cells (continuously versus as a burst); and the total number of virions released (burst size). In the end, the difference in extinction probabilities was the result of the different burst size distributions (geometric vs. delta distributed), rather than the timing of the release (continuous vs. burst). As can be seen from Eq. (8), the timing of release does not affect the extinction probability since each infection event, i.e. whether or not each infectious virion results in an infection, can be treated independently of other infectious virions.

Biologically, the mode or timing of the release does not have a one-to-one relationship with the total number of virions that will be released (burst size). In fact, Pearson et al.² considered a variation of their burst model wherein the burst size, released all at once upon the cell's death, followed a Poisson distribution (Fig. 12E), rather than a Kronecker delta distribution. Figure 12F–G illustrates the differences in the extinction probability for the geometric, Kronecker delta, and Poisson distributed burst sizes explored in Pearson et al.². Echoing earlier results (Fig. 11B), distributions with a higher probability of low burst sizes (geometric followed by Poisson then delta) result in higher extinction probabilities. The effect is most pronounced when the average burst size, \mathcal{B} , is small.

Having explored how the shape parameter of the infectious phase duration distribution (n_I) affects the extinction probability, we can now consider its impact in evaluating and comparing antivirals. For a higher value of n_I , since the extinction probability is lower, the establishment probability, or $1 - (\text{extinction probability})$, is higher. This means that although the effectiveness of antivirals in reducing the establishment probability does not change qualitatively (the better ones remain better), it does change quantitatively (see Supplementary Material Fig. S6). For example, the choice of n_I affects ε_{50} , the efficacy at which the establishment probability is 50% of its value without antivirals, i.e. at $\varepsilon = 0$ (see Table 2). Czuppon et al.²³ state that for an infection initiated with 10 infectious virions, ε_{50} is 81% for an antiviral reducing β , and 85% for an antiviral reducing p , where the authors have assumed $n_I = 1$. When $n_I = 60$, however, we find that ε_{50} is comparable for an antiviral reducing β vs. p (85% vs. 86%), and far less (77%), and therefore potentially easier to achieve, for one reducing γ . This again highlights the importance of properly estimating parameters before making quantitative comparisons of antiviral regimens.

Discussion

Stochastic models (SMs) of virus infection kinetics usually do not represent failure of a virion to cause an infection post cell entry. Yet biologically, there are many ways in which a virion, post cell entry, will fail to cause a cell infection that will yield infectious progeny. Herein, we constructed a SM of virus infection kinetics with an explicit parameter (γ) to represent the probability that a virion will cause a productive infection post cell entry. The SM was first used to estimate the extinction probability of an infection, i.e. the probability that an infection will fail to take hold or spread significantly.

Previously, Czuppon et al.²³ compared prophylactic antivirals reducing virus entry (β) or production (p) for their ability to reduce the establishment probability, or $1 - (\text{extinction probability})$, of a SARS-CoV-2 infection in patients. Extending this work, we investigated how an antiviral reducing γ would compare against antivirals reducing β or p . We found that a prophylactic antiviral reducing γ was best at reducing the establishment probability when infection is initiated with a small number of infectious virions. When instead an infection is initiated with an initial number of infectious cells, possibly representative of post-exposure antiviral therapy, we found that an antiviral reducing γ or p caused the same reduction in the establishment probability, better than that for an antiviral reducing β . While the degree to which an antiviral with a particular mode of action is better than another depends on the chosen parameters, under prophylactic therapy our SM predicts that an antiviral reducing γ will invariably reduce the probability of establishment *at least as much or more so*, than one reducing β or p with the same efficacy, irrespective of the choice of parameters.

In HIV antiviral therapy, reverse transcriptase inhibitors (RTIs) prevent the transcription of viral DNA from viral RNA, a step that occurs after virus entry and is necessary for replication. Conway et al.³ have reported that, under pre-exposure antiviral therapy, RTIs which they represented as reducing their SM's cell infection rate (β) are more effective than protease inhibitors (reducing the virus production rate, p), at reducing the risk of HIV infection. By adding γ in their expression for the extinction probability, we can show that an antiviral reducing γ , which better captures the mode of action of a RTI, rather than β is even more effective at reducing the risk of infection for pre-exposure antiviral therapy, for the infection parameter sets explored. Irrespective of the parameters considered, *reducing γ will always be as effective or more so* in this respect. These findings (detailed

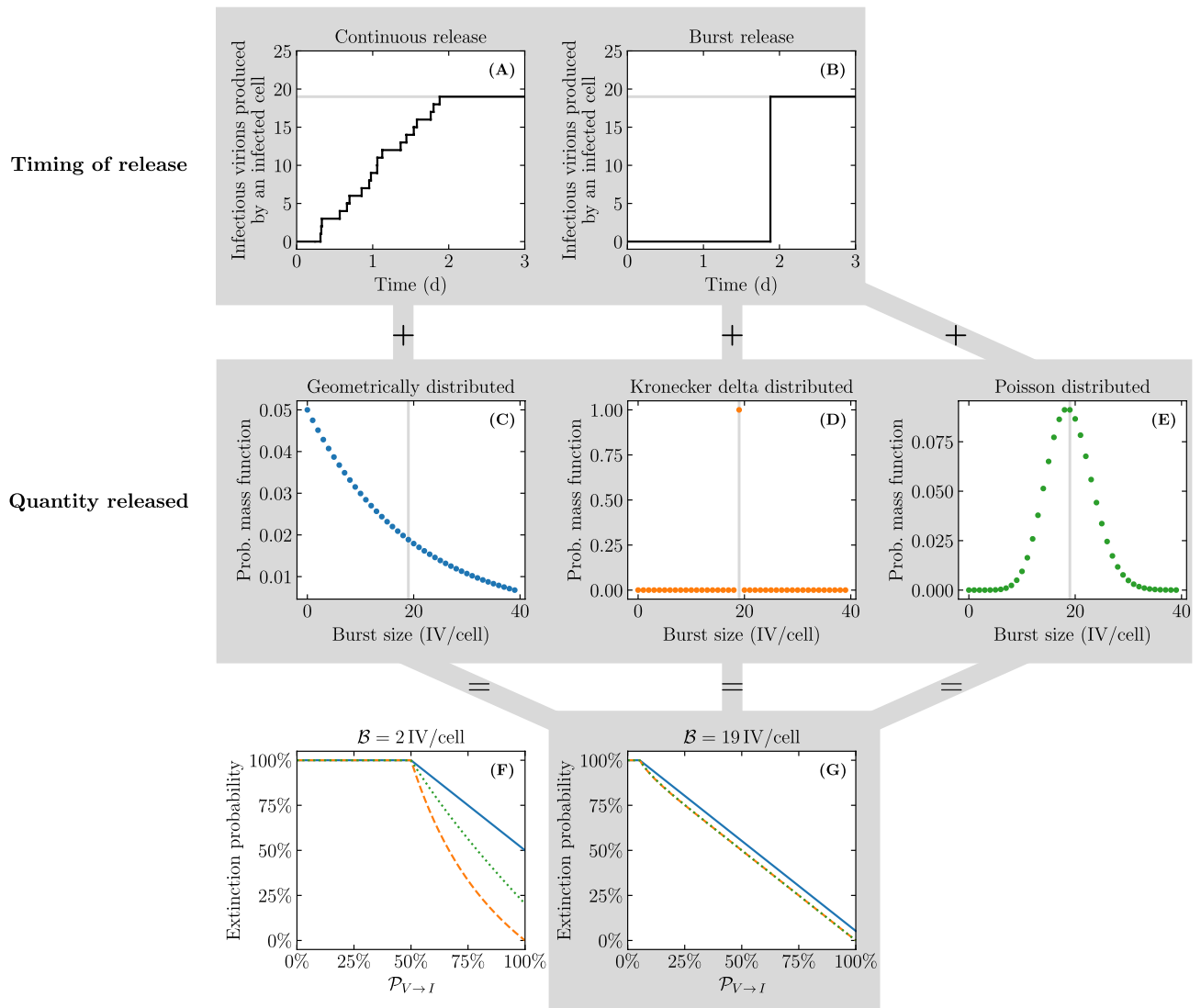


Figure 12. Visual representation for the different modes of virus production explored in Pearson et al.². Two attributes are simultaneously varied for the SMs explored in Pearson et al.²: the timing of infectious virion released by infected cells and the total quantity of infectious virions released. Cumulative number of infectious virions released by an infectious cell (A) at a constant rate over its lifespan; or (B) as a single burst upon its death. Distributions explored for the total number of infectious virions released (burst size, \mathcal{B}) by a population of infected cells: (C) geometric; (D) Kronecker delta; or (E) Poisson. Pearson et al.² defined continuous production to mean continuous release + geometrically distributed \mathcal{B} (A+C); burst production to mean burst release + delta distributed \mathcal{B} (B+D); and explored an alternative random burst model with burst release + Poisson distributed \mathcal{B} (B+E). The vertical grey line represents the average burst size value $\mathcal{B} = 19$ IV/cell. The infection extinction probability for the 3 burst size distributions (C,D,E), as a function of the probability that a single infectious virion productively infects at least one cell ($\mathcal{P}_{V \rightarrow I}$), given an average burst size of (F) 2 IV/cell; or (G) 19 IV/cell.

in Supplementary Material, Section S5) echo those reported herein for SARS-CoV-2 parameters, as to the equal or better performance of antivirals reducing γ , compared to p or β . The introduction of parameter γ to capture productive cell infection failure after an infectious virion has successfully entered a cell is an important consideration when comparing antivirals based on their mode of action.

When the SM was used to simulate infections under various conditions and therapeutic efficacies, the infection outcomes at low antiviral efficacies separated into two clear groups: infection of none or very few cells; and infection of many or most cells. The proportion of infections resulting in nearly no cells consumed, presumed to be the result of infection extinction, indeed matched the infection extinction probability derived for our SM. For infections where many or most cells were consumed, presumed to be established infections, we derived a theoretical expression for the median fraction of cells consumed, and its dependence on the SM parameters. Antiviral therapy acting on a particular infection parameter reduced both the infection's establishment probability and the number of cells consumed by presumed established infections. This finding had 2 important implications.

| Mode of action | Actual ε_{50} (relative to γ) | |
|--|---|----------------|
| | for $n_I = 1$ | for $n_I = 60$ |
| for an initial inoculum of $V_0 = 1$ IV | | |
| Reducing γ | 43% (0%) | 49% (0%) |
| Reducing β | 57% (14%) | 62% (13%) |
| Reducing p | 77% (34%) | 82% (33%) |
| for an initial inoculum of $V_0 = 10$ IV | | |
| Reducing γ | 71% (0%) | 77% (0%) |
| Reducing β | 81% (10%) | 85% (8%) |
| Reducing p | 85% (14%) | 86% (9%) |

Table 2. Efficacy at which the establishment probability is 50% of its value without antivirals (ε_{50}).

Firstly, since the infection's establishment probability and the number of cells consumed by established infections depend differently on the SM's parameters, different antiviral modes of action decreased one quantity more effectively than the other. Secondly, and perhaps more importantly, as antiviral efficacy increases, the fraction of cells consumed by established infections decreases, eventually reaching a point where the fraction consumed by established and extinct infections becomes indistinguishable. For example, at a drug efficacy that yields an equal probability of infection extinction, we found that an antiviral reducing β , compared to one reducing p , resulted in $\sim 3\times$ more cells consumed by so-called established infections. At that same efficacy, an antiviral reducing γ yielded both a lower probability of infection establishment, and resulted in fewer cells consumed overall. Since even an infection extinction probability of 99% will lead to infection establishment in one person out of a hundred, reducing the severity of this infection should be an important clinical consideration when comparing antivirals.

Infection severity often correlates, or is assumed to correlate, with peak or total viral load^{49–52}, this is not always the case^{53,54}, and different measures such as the degree of lung involvement can be better indicators^{55,56}. For example, Myers et al.⁵⁵ stained whole-lung sections of IAV infected mice for influenza nucleoprotein to quantify the % active (antigen-positive) and inactive (lesioned and minimally antigen-positive) area. This quantity is analogous to the fraction of cells consumed by infections tracked herein. Myers et al.⁵⁵ showed that the measured % total lesioned (active + inactive) lung area nonlinearly correlates with the % weight loss, an indicator of disease severity. Looking at the probability that infections will consume more than some biologically critical number of cells, e.g. that identified as sufficient to trigger an immune response and its associated symptoms or minimally sufficient for transmission, could provide a more appropriate target against which to select optimal antiviral candidates.

To facilitate direct comparison of our work herein to past work by others, most of our results were based on assuming that the lifespan of infectious cells, the period during which an infected cell is releasing virus, is exponentially distributed (Erlang distributed with shape parameter $n_I = 1$). But the lifespan of infectious cells for many different viruses has been shown in vitro to be inconsistent with an exponential distribution, and instead follow a more normal-like distribution ($n_I > \sim 10$). Yan et al.⁵ have shown previously that increasing n_I leads to a decrease in the extinction probability. Herein we established that this happens because the probability that an infectious cell will randomly produce a very small total number of virions (burst size) becomes vanishingly small as n_I increases. We showed that while n_I does not affect the fraction of cells consumed by established infections, increasing n_I decreases the probability of infection extinction by decreasing the likelihood of very small burst sizes, thus requiring higher antiviral efficacy to achieve the same risk reduction.

Previous work by Pearson et al.², expanded in Conway et al.³, states that continuous virus production and release over a infected cell's lifespan leads to a higher risk of infection extinction than when virus is released all at once as a burst upon an infected cell's death^{45–48}. Herein we demonstrated that it is the burst size distribution (total number of virions released by infected cells), and not the mode of release (continuous vs all at once), that determines the probability of infection extinction. We explored why the mode or timing of the release (continuous vs burst) does not imply a specific burst size distribution.

The SM introduced herein is general and, like the mean-field mathematical model on which it is based, it could be applicable to a wide range of different viruses. It does, however, make a number of simplifying assumptions that might limit the scope of its applicability. We list some of those below, and discuss how they might impact the findings reported herein.

Immune response and cell regeneration Our results, along with nearly all previous work mentioned herein which used SMs to study virus infection establishment versus extinction focused on in vivo rather than in vitro infections. More accurately, the parameters of these SMs were chosen so as to recapitulate in vivo infection kinetics, but did not include cell regeneration nor an explicit immune response (except for Yan et al.⁵). For established infections, which consume a much higher number of cells, the immune response likely plays an important role in clearing the infection, and towards the end of the infection cell regeneration could begin to replenish the target cell population. Both could likely translate into fewer cells consumed by established infection such that the distinction in the fraction of cells consumed by extinct versus established infections could vanish at even lower antiviral efficacies.

Well-mixed assumption versus spatial inhomogeneity Our SM assumes that the virus and susceptible cells constitute a well-mixed system: all cells are exposed to the same virus concentration. The assumption is reasonable for infection in a liquid medium, for viruses whose primary mode of dissemination is through cell-free dissemination to susceptible cells²⁷. But a number of viruses disseminate exclusively or primarily via direct, cell-to-cell contact, possibly via or as the cause of syncytia formation^{57–60}. Compared to cell-free infection, cell-to-cell infection can be more rapid, more likely to be successful due to transfer of multiple virus genomes, and can offer protection against extracellular defenses (e.g. neutralizing antibodies)⁶¹. The well-mixed assumption might also hold for a mixed population of cells with differing susceptibility for the virus, where our SM would represent the average susceptibility across all cell types. But the specific spatial arrangement of susceptible cells, or large inhomogeneity in the concentration of virus over the space covered by susceptible cells, could result in significant deviations of the infection kinetics from the idealized, well-mixed assumption. For example, in certain solid organs, tissues, or in vitro organoids where cells are organized as layers, as the infection progresses, susceptible cells from the upper layers die and detach, revealing fresh susceptible cells from lower layers, or causing differentiation of other cells into susceptible cells. Such complex interactions and geometries would require tailored modification of the SM presented herein, and detailed, often unavailable knowledge of the parameters required to describe these more complicated arrangements. For IAV infection within the human respiratory tract, we previously used a mathematical model to show how the rapid mucus escalator-induced advection of virions released from the lower respiratory tract upwards towards the nose and mouth, has a strong protective effect. By sweeping virus away from cells in lower sites towards those in upper sites, advection partly helped infection to establish more easily and proceed more rapidly in upper sites, while providing a high degree of protection to lower sites⁶². It would be interesting to re-visit this past work in a stochastic context to compare it against the results presented herein.

Defective interfering and semi-infectious particles Our SM explicitly represents virion that successfully enter a cell but fail to cause an infection, but assumes this process leaves the cell in the same state as if entry had not occurred. Biologically, repeated virus failure post cell entry would likely trigger antiviral pathways within the cell, leaving it in a different state than its naive, susceptible state. Additionally, the accumulation of failed virions within a cell over a sufficiently short time could eventually add up to one functional productive infection (e.g., semi-infectious particles with an incomplete set of segmented genomes rescuing one another^{40,41}) or lead to defective interfering particle production⁴².

Virion pleomorphism Our SM does not consider the impact of virus pleomorphism^{63–66}. For RSV, the pattern observed in vitro for the rate of loss of virion infectivity could be explained by the combined slow versus rapid loss of infectivity by two distinct virus populations, possibly filamentous versus spherical particles⁶⁷. For IAV and IAV pseudotyped with Ebola virus surface proteins, filamentous particles were observed in vitro to be less sensitive than spherical particles to neutralizing antibodies or fusion inhibitors⁶⁸. It is possible that representing the diverse population of individual virions as more or less identical virions whose average properties represent the average over all the subpopulations it contains provides a suitable approximation of this diversity. If two very distinct subpopulations exist, it is also possible that the most infectious population largely determines the infection kinetics which would then also be well represented based on a single homogeneous virus population.

Generally, it is unclear how important these various factors are to shaping the course of an infection, and thus the extent to which they need to be represented. Very simple models in the past have proven to be very effective at not only reproducing, but also predicting the course and outcome of infections both in vitro and in vivo^{33,62,69,70}. This suggests that under physiologically relevant conditions, at least some of these factors can be safely neglected or implicitly represented. With regards to the specific results presented herein, the probability that the infection will become extinct depends on the first few infection events, and if the infection does become extinct very few cells will get consumed. Hence, it is reasonable to expect that the immune response and cell regeneration would not significantly affect the extinction probability or the fraction of cells consumed by extinct infections. Similarly, with so few virions produced by extinct infections, cell co-infection or re-infection is improbable, such that semi-infectious and defective interfering particles would have a negligible impact, the dominant virus subpopulation would determine the outcome, only cells from upper layers of a tissue or organ would see the virus, and spatial inhomogeneity might not yet have developed. For established infections, undoubtedly at least some of these processes would significantly impact the overall fraction of cells consumed. If, however, they did so equally over all antiviral modes of action, these processes might only alter the precise fraction of cells consumed by each mode, and not the more general findings herein as to which mode of action constitutes the most suitable target.

Methods

Parameter values used to generate all figures

The following are the parameters used to generate each figure found within the manuscript.

Figures 4, 13, 14, 15 $p = (11.2/24) \text{ IV}/(\text{cell} \cdot \text{h})$, $N_{\text{cells}} = 4 \times 10^4 \text{ cells}$, $R_0 = 7.69$, $\tau_E = 24/5 \text{ h}$, $\tau_I = (24/0.595) \text{ h}$, $c = (10/24) \text{ h}^{-1}$, $n_E = 1$, $n_I = 1$ and $\beta = [cR_0/\tau_I]/[N_{\text{cells}}(p - R_0/\tau_I)]^{23}$. Also, $\gamma = 1 \text{ cell}/\text{IV}$ and $s = 1 \text{ mL}$.

Figure 5 As in Fig. 4 but with $p \rightarrow (1 - \varepsilon)p$, where $\varepsilon = 0.79$.

Figure 6 As in Fig. 5 but with $\varepsilon = 0.82, 0.84, 0.86$.

Figure 7 As in Fig. 5 but with $\varepsilon = 0.86$, $N_{\text{cells}} = 4 \times [10^3, 10^5, 10^7] \text{ cells}$ and βN_{cells} fixed.

Figure 8 As in Fig. 5 but with $\varepsilon = 0.79, 0.82, 0.84, 0.86$, $N_{\text{cells}} = 4 \times [10^4, 10^5, 10^6, 10^7] \text{ cells}$ and βN_{cells} fixed. Extrapolation of 95% bounds discussed in the text was done by performing linear regression of the \log_{10} of the difference (95% bounds—median) in fraction of cells consumed presented in Fig 8B.

Figure 9 As in Fig. 4 but SM simulations with $\varepsilon = 0.50, 0.60, 0.70, 0.75, 0.78, 0.80, 0.82, 0.83, 0.85$.

Figure 10 As in Fig. 4 but with an antiviral acting either to reduce β , p or γ , where $\varepsilon = 0.86$.

Figure 11 As in Fig. 4 but with $n_I = 1, 2, 7, 60$.

Figure 12 $\mathcal{B} = 19$ IV/cell.

Figure 17A As in Fig. 4 but with $\beta \rightarrow (1 - \varepsilon)\beta$, where $\varepsilon = 0.81$.

Figures 16 and 17B,C are generated using Eqs. (42) and (48), neither of which depend directly on infection parameters, but instead depend on T^*/N_{cells} and/or R_0 whose value is specified thereon.

Details of the SM implementation

The SM is given by Eq. (1), where the changes in the SM variables are random whole numbers drawn at each time step, as follows:

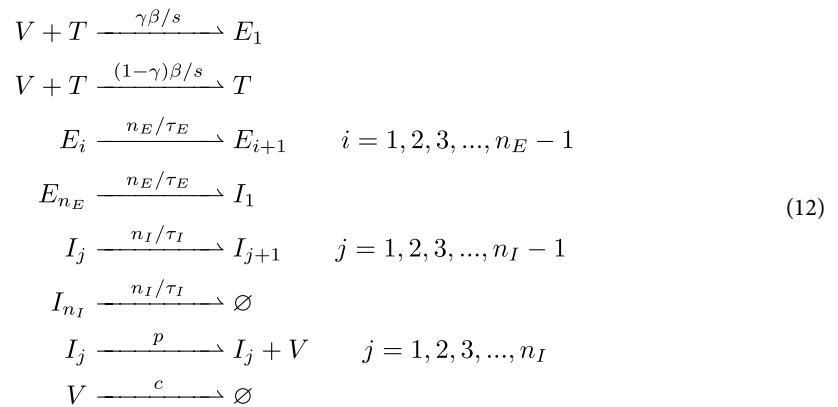
- $E_i^{\text{out}} = \text{Binomial}(n = E_i^t, p_E = \Delta t \cdot n_E/\tau_E)$ corresponds to the number of cells in the i^{th} eclipse compartment (E_i) that will transition to the $(i + 1)^{\text{th}}$ compartment over a time interval Δt , given that there are E_i^t cells in the i^{th} compartment at time step t . Each of the E_i^t cells in eclipse compartment i undergo an independent Bernoulli trial with two possible outcomes: either the cell transitions to the $(i + 1)^{\text{th}}$ compartment with success probability $p_E = \Delta t \cdot n_E/\tau_E$, or otherwise remains in the i^{th} compartment. The time step, Δt is chosen to be sufficiently small to ensure $p_E < 1$, as discussed below.
- $I_j^{\text{out}} = \text{Binomial}(n = I_j^t, p_I = \Delta t \cdot n_I/\tau_I)$ is the number of cells in the j^{th} infectious compartment (I_j) that transition to the $(j + 1)^{\text{th}}$ compartment over a time interval Δt , given that there are I_j^t cells in the j^{th} compartment at time step t , and the probability of transition $p_I = \Delta t \cdot n_I/\tau_I$.
- $V^{\text{prod}} = \text{Poisson}(\lambda = \Delta t \cdot p \sum_{j=1}^{n_I} I_j^t)$ is the number of infectious virions newly produced into the supernatant over a time Δt , given that there are $\sum_{j=1}^{n_I} I_j^t$ infectious cells at time step t , each producing infectious virions at a rate of p infectious virions per hour. The Poisson distributed random variable, V^{prod} is the number of events that occurred, given the expected number of occurrences over a time Δt , namely $\lambda = \Delta t \cdot p \sum_{j=1}^{n_I} I_j^t$.
- $V^{\text{decay}}, V^{\text{enter}}, \text{otherwise} = \text{Trinomial}(n = V^t, p_1 = \Delta t \cdot c, p_2 = \Delta t \cdot \beta T^t/s, p_3 = 1 - p_1 - p_2)$ corresponds to the number of infectious virions in the supernatant at time t , V^t , that end up in each of 3 possible fates. V^{decay} is the number of infectious virions that lose infectivity with probability $p_1 = \Delta t \cdot c$, V^{enter} are lost from the supernatant as they enter a target cell with probability $p_2 = \Delta t \cdot \beta T^t/s$, and otherwise infectious virions do neither and remain in the supernatant with probability $p_3 = 1 - p_1 - p_2$. Each probability (p_1, p_2, p_3) can be considered constant over time interval Δt , provided a sufficiently small time step is chosen, as discussed below.
- $N^{\text{inf}} = |\{x_i | x_i \in \mathcal{U}\{a = 1, b = T^t\}, 1 \leq i \leq V^{\text{suc}}\}|$ is the number of target cells that become infected ($T \rightarrow E_1$) over time Δt , given that V^{enter} infectious virions enter into target cells, out of which $V^{\text{suc}} = \text{Binomial}(n = V^{\text{enter}}, p_V = \gamma)$ infectious virions ultimately lead to successful cell infection, given probability $p_V = \gamma$. Random variable N^{inf} does not correspond to any named probability mass function. The N^{inf} expression simulates randomly placing V^{suc} infectious virions into T^t cells chosen at random with replacement, $\{x_i | x_i \in \mathcal{U}\{a = 1, b = T^t\}, 1 \leq i \leq V^{\text{suc}}\}$, and counts the number of distinct cells that received one or more infectious virions, $|\{\dots\}|$. Computationally, it is implemented in python as `len(numpy.unique(numpy.random.choice(a=T^t, size=V^suc, replace=True)))`.

The duration of the SM's discrete time steps, Δt , which sets the probability of event occurrences or the number of such events, is computed at each iteration step t as

$$\Delta t = \frac{\mathcal{P}_{\text{events}}}{\max \left\{ \frac{\beta T^t}{s}, c, \frac{n_E}{\tau_E}, \frac{n_I}{\tau_I} \right\}} \quad (11)$$

where $\mathcal{P}_{\text{events}}$ is the probability of occurrence of the most likely event, namely of virion loss due to cell entry ($\beta T^t/s$) or loss of infectivity (c), or transition of cells from one infected state to another (n_E/τ_E and n_I/τ_I). A value of $\mathcal{P}_{\text{events}} = 0.05$ (or 5%) is used, as it was found to be sufficiently small to provide accurate solutions (see "Numerical accuracy of the SM's solutions" in Methods).

Some readers might prefer to see the SM expressed as transitions^{2,5}, namely



One limitation of this representation is that it indicates that one virion that enters a cell and that is successful at causing an infection with rate $\gamma\beta/s$ will cause the infection of one cell ($V + T \rightarrow E_1$). In reality, in our SM, one successful virion does not necessarily cause the infection of one cell (see N^{inf} in Table 1) because one cell could receive two successful infectious virions which would only result in one rather than two new successful cell infections.

Our stochastic modelling approach is similar to standard applications of the τ -leaping method⁷¹. One important difference is that the τ -leaping method represents the number of events that will occur for each reaction as a Poisson random variable. Because there is no upper bound on a Poisson distributed random variable, this assumption can lead to negative populations of reactants when the number of events that cause the loss of a reactant are larger than the number of reactants available^{72–74}. To address this issue, variations of the τ -leaping method make use of finite range binomial^{72,73} or multinomial⁷⁴ random variables, as was done herein.

Our SM expression for the number of target cells that become infected, N^{inf} , as a function of the number of virion entries that successfully lead to cell infection (V^{suc}) is novel. It differs from previous work that assumed as many infected cells as successful virions entries ($N^{\text{inf}} = V^{\text{suc}}$) by instead correctly accounting for cases when more than one successful virions enter the same cell leading to a single cell infection ($N^{\text{inf}} \leq V^{\text{suc}}$). In contrast, our theoretical expressions for the extinction probability, Eq. (2), and the median fraction of cells consumed by established infections, Eq. (40), assume $N^{\text{inf}} = V^{\text{suc}}$, as in previous work. Yet in cases where infection extinction and establishment are clearly distinct, the proportion of SM infections presumed extinct agrees with the theoretically predicted extinction probability, and the median fraction of cells consumed by SM infections presumed established agrees well with our theoretical expression. This is likely because the number of infectious virions V^{suc} considered herein is sufficiently small compared to the number of cells T^t that the entry of more than one V^{suc} into the same T^t cell occurs rarely and $N^{\text{inf}} = V^{\text{suc}}$ largely holds for the parameters explored herein. Nevertheless, our SM's expression for N^{inf} more accurately represents the underlying process, and its computational cost is negligible, and thus worth keeping.

The MFM was developed independently from the SM. As such, the MFM was not intended to be exactly the same as the mean of the SM equations. If one wanted to derive the mean of the SM equations, one would have to consider higher order moments of the SM variables^{75–77}. Using a simplified version of our SM, we derived second order moment equations, approximated third order moments, and found no significant disagreement between the MFM solution with and without second order moment terms (see Supplementary Material, Section S1).

Numerical accuracy of the SM's solutions

SM simulation output presented in Results, for both extinct and established infections, are those associated with Figs. 5, 6, 7, 10, i.e. those showing the frequency distribution for the number of cells consumed by 10^6 SM simulated infections. The most important quantities discussed about these figures are the simulated fraction of extinct infections and median fraction of cells consumed by established infections. These can be compared to the theoretical extinction probability, $\mathcal{P}_{V \rightarrow \text{Extinction}}$, and median fraction of cells consumed by established infections, $1 - T(\infty)/N_{\text{cells}}$. When extinction and establishment are clearly distinct, the absolute error of $\mathcal{P}_{V \rightarrow \text{Extinction}}$ and $1 - T(\infty)/N_{\text{cells}}$ is of the order of 10^{-4} or 0.01% (see Table 3 and 4). Further, the absolute error of $\mathcal{P}_{V \rightarrow \text{Extinction}}$ is always within one standard deviation of the expected distribution for the fraction of extinct infections after 10^6 SM simulations, $1/10^6 \cdot \text{Binomial}(n = 10^6, p = \mathcal{P}_{V \rightarrow \text{Extinction}})$ (see Table 3). This indicates that the time step criteria (Eq. (11)) and the number of SM simulations were sufficient to provide good accuracy for our key results.

The numerical solution of the SM was further validated by replacing the SM's random variables, e.g. E_i^{out} , by their expected value, e.g. $E_i^{\text{out}} = E_i^t \cdot \Delta t \cdot n_E/\tau_E$, and comparing the solution against that obtained for the MFM in Eq. (1) with a standard numerical ODE solver. The solutions were found to be in agreement over the wide range of parameter values explored (not shown).

MFM using the expected value of the SM's random variables

One way of defining a MFM is by replacing the random variables of the SM by their expected values, listed in Table 5. This is trivial for the random variables that are drawn from well-defined probability distributions. But it is not for N^{inf} , which is determined by simulating the process of randomly placing V^{suc} infectious virions into T^t cells chosen at random with replacement, $\{x_i | x_i \in \mathcal{U}\{a = 1, b = T^t\}, 1 \leq i \leq V^{\text{suc}}\}$, and counting the number of distinct cells that received one or more infectious virions, $\{...\}$.

Consider then e_n , the expected number of empty cells after placing n infectious virions into T^t cells. After placing $n - 1$ infectious virions into T^t cells, there is on average e_{n-1} empty cells by definition. There is then a e_{n-1}/T^t probability to place the last infectious virion into an empty cell. It follows that

$$\begin{aligned}
 e_n &= \frac{e_{n-1}}{T^t} (e_{n-1} - 1) + \left(1 - \frac{e_{n-1}}{T^t}\right) e_{n-1} \\
 e_n &= \frac{e_{n-1}^2}{T^t} - \frac{e_{n-1}}{T^t} + e_{n-1} - \frac{e_{n-1}^2}{T^t} \\
 e_n &= e_{n-1} \left(\frac{T^t - 1}{T^t}\right)
 \end{aligned}
 \tag{13}$$

By recursion and since trivially $e_0 = T^t$,

$$e_n = T^t \left(\frac{T^t - 1}{T^t}\right)^n
 \tag{14}$$

The expected number of N^{inf} cells after placing V^{suc} infectious virions into T^t cells is therefore represented as

$$T^t \left[1 - \left(\frac{T^t - 1}{T^t}\right)^{V^{\text{suc}}}\right]
 \tag{15}$$

When there are far fewer successful virion entry events than there are target cells ($V^{\text{suc}} \ll T^t$), the above expression for the expected value of N^{inf} equals V^{suc} .

Probability that an infectious virion succeeds at causing a productive cell infection

Let us derive the probability that an infectious virion is successful at causing a productive cell infection, $\mathcal{P}_{V \rightarrow I}$, in a population of fully susceptible, uninfected cells ($T = N_{\text{cells}}$).

In a time step Δt , the probability that an infectious virion neither loses infectivity nor enters a cell is $(1 - \Delta t c - \Delta t \beta N_{\text{cells}}/s)$, the probability that an infectious virion enters a cell is $\Delta t \beta N_{\text{cells}}/s$ and the probability that an infectious virion post cell entry will be successful at causing a cell infection is γ . The probability that an infectious virion is successful at causing a productive cell infection in time $t = k\Delta t$ is then given by the following expression,

$$(1 - \Delta t c - \Delta t \beta N_{\text{cells}}/s)^{k-1} (\Delta t \beta N_{\text{cells}}/s) \gamma
 \tag{16}$$

More generally, the probability that this happens in any time $t = k\Delta t$ where $0 \leq t \leq t_{\text{incub}}$ and $t_{\text{incub}} = n\Delta t$ is the incubation time, is then expressed as

| Figures | $\mathcal{P}_{V \rightarrow \text{Extinction}}^{\text{Theoretical}}$ | $\mathcal{P}_{V \rightarrow \text{Extinction}}^{\text{Simulated}}$ | Standard deviation [†] | Absolute error |
|---------|--|--|---------------------------------|----------------|
| 5 | 0.848170 | 0.848227 | 0.000359 | 0.000057 |
| 6(A) | 0.884467 | 0.884348 | 0.000320 | 0.000119 |
| 6(B) | 0.917389 | 0.917172 | 0.000275 | 0.000213 |
| 7(B) | 0.961389 | 0.961308 | 0.000193 | 0.000081 |
| 7(C) | 0.961389 | 0.961223 | 0.000193 | 0.000166 |
| 10(A) | 0.963306 | 0.963198 | 0.000188 | 0.000108 |

Table 3. Absolute errors for the extinction probability. [†] Standard deviation =

$$1/10^6 \cdot \sqrt{10^6 \cdot \mathcal{P}_{V \rightarrow \text{Extinction}}^{\text{Theoretical}} \cdot (1 - \mathcal{P}_{V \rightarrow \text{Extinction}}^{\text{Theoretical}})}$$

| Figures | $[1 - T(\infty)/N_{\text{cells}}]^{\text{Theoretical}}$ Established, median | $[1 - T(\infty)/N_{\text{cells}}]^{\text{Simulated}}$ Established, median | Absolute error |
|---------|--|--|----------------|
| 5 | 0.796818 | 0.796850 | 0.000032 |
| 6(A) | 0.675769 | 0.675850 | 0.000081 |
| 6(B) | 0.533028 | 0.533325 | 0.000297 |
| 7(B) | 0.284271 | 0.284442 | 0.000171 |
| 7(C) | 0.284271 | 0.284187 | 0.000084 |
| 10(A) | 0.716323 | 0.716375 | 0.000052 |

Table 4. Absolute errors for the median fraction of cells consumed by established infections.

| Random variable | Expected value |
|--|---|
| E_i^{out} | $E_i^t \cdot \Delta t n_E / \tau_E$ where $i = 1, 2, \dots, n_E$ |
| I_j^{out} | $I_j^t \cdot \Delta t n_I / \tau_I$ where $j = 1, 2, \dots, n_I$ |
| V^{prod} | $\Delta t \cdot p \sum_{j=1}^{n_I} I_j^t$ |
| $V^{decay}, V^{enter}, \text{otherwise}$ | $V^t \cdot \Delta t c, V^t \cdot \Delta t \beta T^t / s, V^t \cdot (1 - \Delta t c - \Delta t \beta T^t / s)$ |
| N^{inf} | $T^t [1 - [(T^t - 1) / T^t]^{V^{suc}}]$ for $T^t \geq 1$ and 0 for $T^t < 1$ where the expected value of V^{suc} is $V^{enter} \cdot \gamma$ |

Table 5. Expected value of the random variables used by the SM.

$$\begin{aligned}
 & \sum_{k=1}^n [(1 - \Delta t c - \Delta t \beta N_{cells}/s)^{k-1} (\Delta t \beta N_{cells}/s) \gamma] \\
 &= [\Delta t \gamma \beta N_{cells}/s] \sum_{k=1}^n [1 - \Delta t (c + \beta N_{cells}/s)]^{k-1} \\
 &= [\Delta t \gamma \beta N_{cells}/s] \underbrace{\sum_{k=0}^{n-1} [1 - \Delta t (c + \beta N_{cells}/s)]^k}_{\text{partial sum of geometric series}} \\
 &= [\Delta t \gamma \beta N_{cells}/s] \left[\frac{1 - [1 - \Delta t (c + \beta N_{cells}/s)]^n}{1 - [1 - \Delta t (c + \beta N_{cells}/s)]} \right] \\
 &= [\Delta t \gamma \beta N_{cells}/s] \left[\frac{1 - [1 - \Delta t (c + \beta N_{cells}/s)]^n}{\Delta t (c + \beta N_{cells}/s)} \right] \\
 &= \left[\frac{\gamma \beta N_{cells}/s}{c + \beta N_{cells}/s} \right] \left[1 - \left(1 - \frac{t_{incub} (c + \beta N_{cells}/s)}{n} \right)^n \right]
 \end{aligned} \tag{17}$$

In our SM, Δt is set according to Eq. (11). Let us assume that Δt is sufficiently small that we can take the limit of Eq. (17) as $n \rightarrow \infty$ (i.e. $\Delta t \rightarrow 0$), namely,

$$\begin{aligned}
 & \left[\frac{\gamma \beta N_{cells}/s}{c + \beta N_{cells}/s} \right] \left[1 - \lim_{n \rightarrow \infty} \left(1 - \frac{t_{incub} (c + \beta N_{cells}/s)}{n} \right)^n \right] \\
 &= \left[\frac{\gamma \beta N_{cells}/s}{c + \beta N_{cells}/s} \right] \left[1 - e^{-t_{incub} (c + \beta N_{cells}/s)} \right]
 \end{aligned} \tag{18}$$

As $t_{incub} \rightarrow \infty$, Eq. (18) is then given by the following expression,

$$\mathcal{P}_{V \rightarrow I} = \frac{\gamma \beta N_{cells}/s}{c + \beta N_{cells}/s} = \frac{\gamma}{1 + c / (\beta N_{cells}/s)} \tag{19}$$

Fig. 13 shows that the frequency of success of an infectious virion to cause a productive cell infection in a population of uninfected cells, generated from 10^5 SM simulations, is in agreement with Eq. (19), over a wide range of infection parameters. This means that our earlier oassumption was reasonable that Δt is sufficiently small that we can take the limit as $\Delta t \rightarrow 0$.

Probability that an infectious cell produces m infectious virions

Let us derive the probability that an infectious cell produces m infectious virions over its lifespan ($\mathcal{P}_{I \rightarrow mV}$).

Assuming that successive infectious virion productions are independent events that happen under constant rate p , then the probability that an infectious cell produces m infectious virions over its lifespan t is given by Poisson($m | \lambda = t p$).

The negative binomial (NB) distribution represents the probability that there are k successes in a sequence of independent and identically distributed Bernoulli trials with probability of success p before a specified number of failures r occurs, namely,

$$\text{NB}(k|r, p) = \binom{k+r-1}{k} (1-p)^r p^k \tag{20}$$

The probability that an infectious cell transition to the next compartment is given by Binomial($1 | n = 1, p_I = \Delta t n_I / \tau_I$). Therefore the probability that an infectious cell transition to the next compartment for n_I time steps and does not for $n - n_I$ time steps is then given by NB($n_I | r = n - n_I, p_I = \Delta t n_I / \tau_I$). It is possible to show that the continuous analogue of this discrete probability distribution is the Erlang distribution

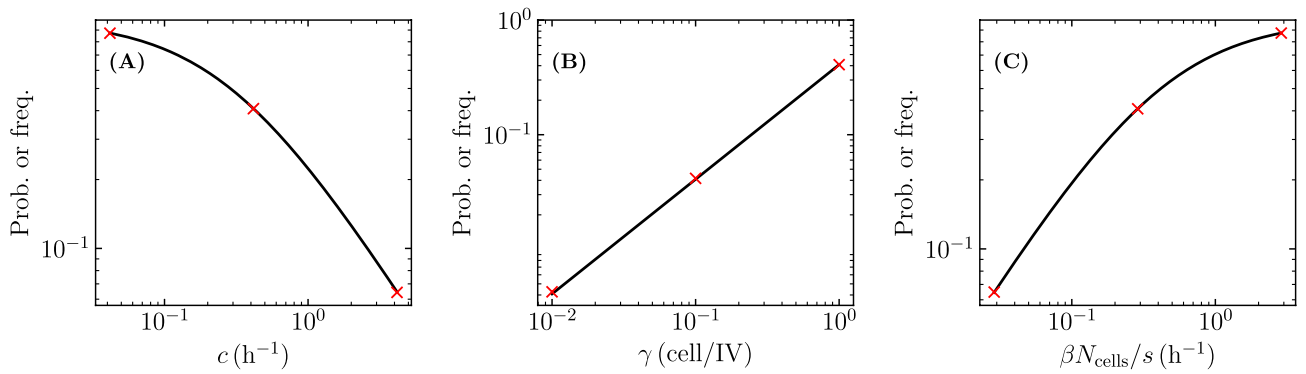


Figure 13. Probability that an infectious virion is successful at causing a productive cell infection. Frequency of success of an infectious virion to cause a productive cell infection in a population of uninfected cells for 10^5 SM simulations (red scatter) compared to $\mathcal{P}_{V \rightarrow I}$ the derived expression for the probability of success of an infectious virion to cause a productive cell infection in a population of uninfected cells (black curve, Eq. (19)) while either varying (A) c , (B) γ or (C) $\beta N_{\text{cells}}/s$. Unless otherwise specified, the parameters were the same as in Fig. 4.

(Erlang($t|k = n_I, \lambda = n_I/\tau_I$)). As a result, we make use of this probability density function to express the probability that the infectious lifespan has duration t .

The probability that an infectious cell produces m infectious virions over its lifespan, i.e. that the burst size of an infected cell is m , can then be derived by integrating over time t , the joint probability distribution that an infectious cell produces m infectious virions given a lifespan of duration t and that an infectious cell has a lifespan of duration t , namely,

$$\mathcal{P}_{I \rightarrow mV} = \int_0^\infty \text{Poisson}(m|\lambda = pt) \cdot \text{Erlang}(t|k = n_I, \lambda = n_I/\tau_I) dt = \text{NB}(m|r = n_I, p_B = \mathcal{B}/(n_I + \mathcal{B})) \tag{21}$$

where $\mathcal{B} = p\tau_I$ is the average burst size. When $n_I = 1$, $\mathcal{P}_{I \rightarrow mV} = \text{NB}(m|r = 1, p_B = \mathcal{B}/(1 + \mathcal{B})) = \text{Geom}(m|p = 1/(1 + \mathcal{B}))$ where Geom represents the geometric distribution. When $n_I \rightarrow \infty$, $\mathcal{P}_{I \rightarrow mV} = \text{NB}(m|r = n_I, p_B = \mathcal{B}/(n_I + \mathcal{B})) = \text{Poisson}(m|\lambda = \mathcal{B})$.

Figure 14 shows that the normalized histogram of the burst size, generated from 10^5 SM simulations, is in agreement with Eq. (21), over a wide range of infection parameters.

Extinction probability

Following Pearson et al.², let us derive the extinction probability of an infection. The extinction probability of an infection with initially one infectious virion ($\mathcal{P}_{V \rightarrow \text{Extinction}}$) can be written as

$$\mathcal{P}_{V \rightarrow \text{Extinction}} = (1 - \mathcal{P}_{V \rightarrow I}) + \mathcal{P}_{V \rightarrow I} \cdot \mathcal{P}_{I \rightarrow \text{Extinction}} \tag{22}$$

since the infection can become extinct if the initial infectious virion fails to cause a productive cell infection with probability $1 - \mathcal{P}_{V \rightarrow I}$ or if it does cause a productive cell infection with probability $\mathcal{P}_{V \rightarrow I}$ but that cell infection then leads to extinction with probability $\mathcal{P}_{I \rightarrow \text{Extinction}}$.

Likewise, the extinction probability of an infection with initially one infectious cell ($\mathcal{P}_{I \rightarrow \text{Extinction}}$) can be expressed as

$$\mathcal{P}_{I \rightarrow \text{Extinction}} = \sum_{m=0}^\infty \mathcal{P}_{I \rightarrow mV} \cdot (\mathcal{P}_{V \rightarrow \text{Extinction}})^m \tag{23}$$

since the initial infectious cell will produce some number m of infectious virions over its lifespan with probability $\mathcal{P}_{I \rightarrow mV}$ and each one of these produced infectious virions can be seen as an independent infection event that can lead to extinction with probability $\mathcal{P}_{V \rightarrow \text{Extinction}}$.

Substituting our derived expression for $\mathcal{P}_{I \rightarrow mV}$ (Eq. 21) into Eq. (23) yields

$$\begin{aligned} \mathcal{P}_{I \rightarrow \text{Extinction}} &= \sum_{m=0}^\infty \text{NB}(m|r = n_I, p_B = \mathcal{B}/(n_I + \mathcal{B})) \cdot (\mathcal{P}_{V \rightarrow \text{Extinction}})^m \\ \mathcal{P}_{I \rightarrow \text{Extinction}} &= \sum_{m=0}^\infty \binom{m + n_I - 1}{m} \left[1 - \frac{\mathcal{B}}{n_I + \mathcal{B}}\right]^{n_I} \left[\frac{\mathcal{B}}{n_I + \mathcal{B}}\right]^m \cdot (\mathcal{P}_{V \rightarrow \text{Extinction}})^m \\ \mathcal{P}_{I \rightarrow \text{Extinction}} &= \left[\frac{n_I}{n_I + \mathcal{B}}\right]^{n_I} \sum_{m=0}^\infty \binom{m + n_I - 1}{m} \left[\frac{\mathcal{B}}{n_I + \mathcal{B}} \cdot \mathcal{P}_{V \rightarrow \text{Extinction}}\right]^m \end{aligned} \tag{24}$$

The following binomial series can be used to simplify Eq. (24),

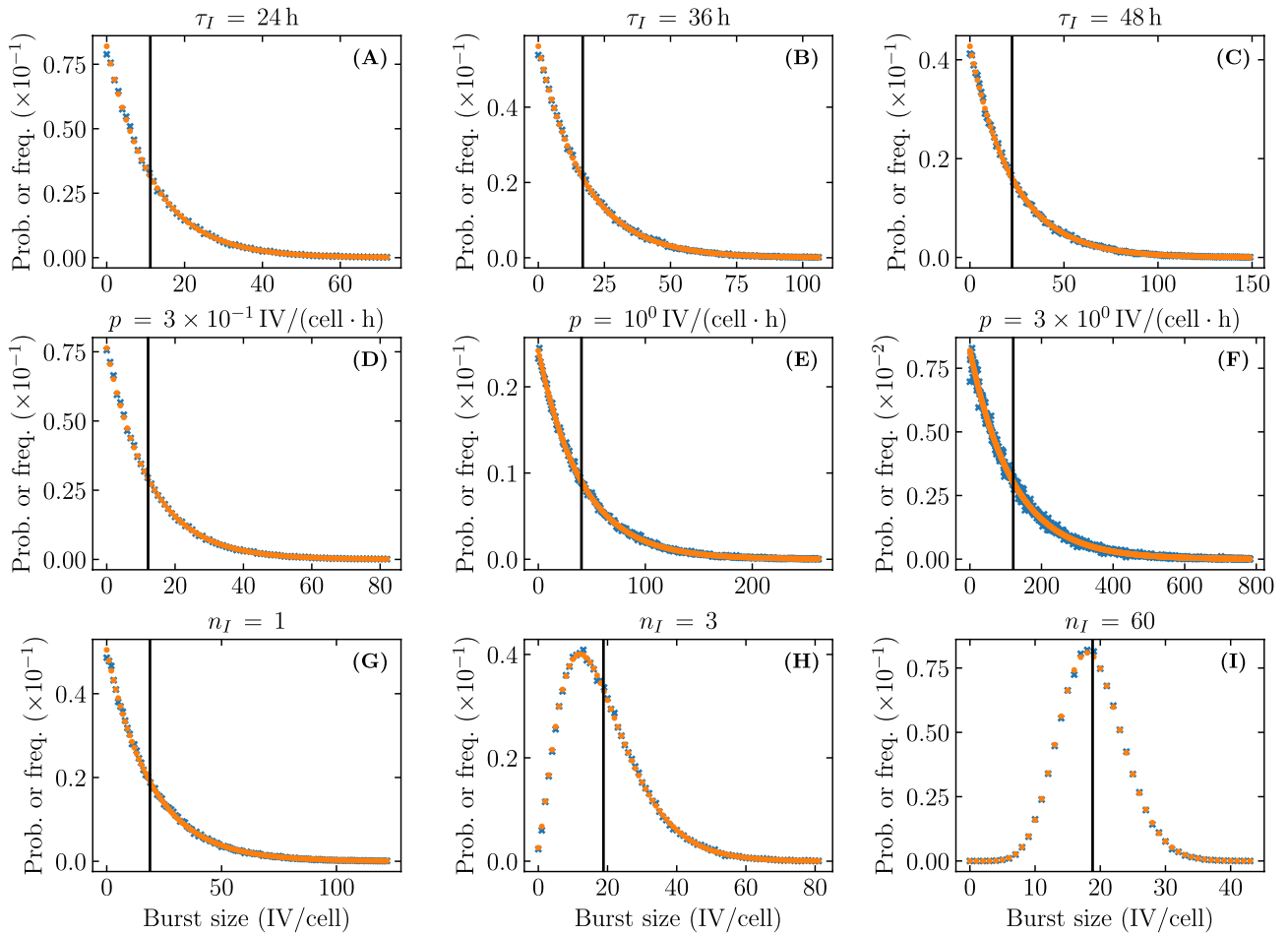


Figure 14. Probability that an infectious cell will produce m infectious virions. Normalized histogram of the burst size, i.e. the total number of infectious virions produced by one infectious cell over its lifespan, for 10^5 SM simulations (blue x), compared to the derived expression for the probability mass function of the burst size distribution (orange circle, Eq. (21)) while either varying (A–C) τ_I , (D–F) p or (G–I) n_I . The black lines show the average burst size values $\mathcal{B} = p\tau_I$. Unless otherwise specified, the parameters were the same as in Fig. 4.

$$\frac{1}{(1-x)^s} = \sum_k \binom{s+k-1}{k} x^k \tag{25}$$

Let us derive some related expressions,

$$x := \frac{\mathcal{B}}{n_I + \mathcal{B}} \cdot \mathcal{P}_{V \rightarrow \text{Extinction}} \tag{26}$$

$$\begin{aligned} \implies 1-x &= \frac{n_I + \mathcal{B}(1 - \mathcal{P}_{V \rightarrow \text{Extinction}})}{n_I + \mathcal{B}} \\ \implies \frac{1}{(1-x)^{n_I}} &= \left[\frac{n_I + \mathcal{B}}{n_I + \mathcal{B}(1 - \mathcal{P}_{V \rightarrow \text{Extinction}})} \right]^{n_I} \end{aligned} \tag{27}$$

Eq. (24) can then be simplified to

$$\begin{aligned} \mathcal{P}_{I \rightarrow \text{Extinction}} &= \left[\frac{n_I}{n_I + \mathcal{B}} \right]^{n_I} \left[\frac{n_I + \mathcal{B}}{n_I + \mathcal{B}(1 - \mathcal{P}_{V \rightarrow \text{Extinction}})} \right]^{n_I} \\ &= \left[\frac{n_I}{n_I + \mathcal{B}(1 - \mathcal{P}_{V \rightarrow \text{Extinction}})} \right]^{n_I} \\ &= \left[\frac{\mathcal{B}(1 - \mathcal{P}_{V \rightarrow \text{Extinction}})}{n_I} + 1 \right]^{-n_I} \end{aligned} \tag{28}$$

Substituting Eq. (28), our derived expression for $\mathcal{P}_{V \rightarrow I}$ (Eq. (19)) into Eq. (22) yields

$$\mathcal{P}_{V \rightarrow \text{Extinction}} = (1 - \mathcal{P}_{V \rightarrow I}) + \mathcal{P}_{V \rightarrow I} \left[\frac{\mathcal{B}(1 - \mathcal{P}_{V \rightarrow \text{Extinction}})}{n_I} + 1 \right]^{-n_I} \tag{29}$$

$$\begin{aligned} \mathcal{P}_{V \rightarrow \text{Extinction}} &= \left[1 - \frac{\gamma \beta N_{\text{cells}}/s}{c + \beta N_{\text{cells}}/s} \right] + \frac{\gamma \beta N_{\text{cells}}/s}{c + \beta N_{\text{cells}}/s} \left[\frac{\mathcal{B}(1 - \mathcal{P}_{V \rightarrow \text{Extinction}})}{n_I} + 1 \right]^{-n_I} \\ \mathcal{P}_{V \rightarrow \text{Extinction}} &= 1 - \frac{\gamma}{c/(\beta N_{\text{cells}}/s) + 1} + \frac{\gamma}{c/(\beta N_{\text{cells}}/s) + 1} \left[\frac{\mathcal{B}(1 - \mathcal{P}_{V \rightarrow \text{Extinction}})}{n_I} + 1 \right]^{-n_I} \\ 0 &= 1 - \frac{[1 + c/(\beta N_{\text{cells}}/s)] [1 - \mathcal{P}_{V \rightarrow \text{Extinction}}]}{\gamma} - \left[\frac{\mathcal{B}(1 - \mathcal{P}_{V \rightarrow \text{Extinction}})}{n_I} + 1 \right]^{-n_I} \end{aligned} \tag{30}$$

This expression does not seem to have an analytical solution for $\mathcal{P}_{V \rightarrow \text{Extinction}}$ but a solution can be found numerically by finding the roots of the expression on the right-hand side of Eq. (30) using `scipy.optimize.fsolve`. Figure 15 shows that the frequency of extinction given an infection initiated with only one infectious virion, generated from 10^5 SM simulations, is in agreement with the probability of extinction given an infection initiated with only one infectious virion, over a wide range of infection parameters. The extinction probability given any number of infectious virions V_0 and infectious cells I_0 is given by $(\mathcal{P}_{V \rightarrow \text{Extinction}})^{V_0} \cdot (\mathcal{P}_{I \rightarrow \text{Extinction}})^{I_0}$. It is important to note that, as others^{2,5}, we also make the assumption in this derivation that $T = N_{\text{cells}}$ is constant. When $n_I = 1$, Eq. (29) simplifies to

$$\mathcal{P}_{V \rightarrow \text{Extinction}} = (1 - \mathcal{P}_{V \rightarrow I}) + \mathcal{P}_{V \rightarrow I} \left[\frac{1}{1 + \mathcal{B}(1 - \mathcal{P}_{V \rightarrow \text{Extinction}})} \right] \tag{31}$$

In this case, $\mathcal{P}_{V \rightarrow \text{Extinction}} = (1 - \mathcal{P}_{V \rightarrow I}) + 1/\mathcal{B}$ is a solution to the above expression as,

$$\begin{aligned} (1 - \mathcal{P}_{V \rightarrow I}) + 1/\mathcal{B} &= (1 - \mathcal{P}_{V \rightarrow I}) + \mathcal{P}_{V \rightarrow I} \left[\frac{1}{1 + \mathcal{B}[1 - (1 - \mathcal{P}_{V \rightarrow I} + 1/\mathcal{B})]} \right] \\ (1 - \mathcal{P}_{V \rightarrow I}) + 1/\mathcal{B} &= (1 - \mathcal{P}_{V \rightarrow I}) + \mathcal{P}_{V \rightarrow I} \left[\frac{1}{\mathcal{B} \cdot \mathcal{P}_{V \rightarrow I}} \right] \\ (1 - \mathcal{P}_{V \rightarrow I}) + 1/\mathcal{B} &= (1 - \mathcal{P}_{V \rightarrow I}) + 1/\mathcal{B} \\ \therefore \text{LHS} &= \text{RHS} \end{aligned}$$

$\mathcal{P}_{V \rightarrow \text{Extinction}} = (1 - \mathcal{P}_{V \rightarrow I}) + 1/\mathcal{B}$ is equivalent to Eq. (15) in Pearson et al.². $\mathcal{P}_{V \rightarrow \text{Extinction}} = (1 - \mathcal{P}_{V \rightarrow I}) + 1/\mathcal{B}$ can also be written in terms of the basic reproductive number $R_0 = \mathcal{B} \cdot \mathcal{P}_{V \rightarrow I}$ (see Eq. (7)), i.e. $\mathcal{P}_{V \rightarrow \text{Extinction}} = 1 - (R_0 - 1)/\mathcal{B}$. Also, $\mathcal{P}_{V \rightarrow \text{Establishment}}$ or $1 - \mathcal{P}_{V \rightarrow \text{Extinction}}$ is then given by $\mathcal{P}_{V \rightarrow I} - 1/\mathcal{B}$ or $(R_0 - 1)/\mathcal{B}$.

When $n_I \rightarrow \infty$, $\mathcal{P}_{I \rightarrow mV} = \text{Poisson}(m|\lambda = \mathcal{B})$, therefore, we obtain,

$$\mathcal{P}_{V \rightarrow \text{Extinction}} = (1 - \mathcal{P}_{V \rightarrow I}) + \mathcal{P}_{V \rightarrow I} \cdot \sum_{m=0}^{\infty} \text{Poisson}(m|\lambda = \mathcal{B}) \cdot (\mathcal{P}_{V \rightarrow \text{Extinction}})^m \tag{32}$$

which is equivalent to Eq. (26) in Pearson et al.².

Expression for the fraction of cells consumed by infections

As it happens, the median fraction of cells consumed by established infections in the SM approximately corresponds to the fraction of cells consumed in the deterministic MFM for a given parameter set. We can use the latter to obtain an analytical expression for the median fraction of cells infected by our SM to understand how each SM parameter affects this quantity.

So let us derive the fraction of cells consumed in the MFM for an infection starting with a number of infectious virions $V(0)$. Following⁷⁸, we have the following relations from Eq. (1),

$$\dot{T} + \sum_{i=1}^{n_E} \dot{E}_i + \sum_{j=1}^k \dot{I}_j = -n_I I_k / \tau_I \implies I_k = -(\tau_I / n_I) \left[\dot{T} + \sum_{i=1}^{n_E} \dot{E}_i + \sum_{j=1}^k \dot{I}_j \right] \tag{33}$$

$$\frac{d}{dt} \ln(T) = \frac{\dot{T}}{T} = -\gamma \beta V / s \implies \frac{1}{\gamma \beta / s} \frac{d}{dt} \ln(T) = -V \tag{34}$$

Substituting these relations into Eq. (1) for \dot{V} , we obtain

$$\dot{V} = -(p\tau_I / n_I) \left[n_I \left(\dot{T} + \sum_{i=1}^{n_E} \dot{E}_i \right) + \sum_{k=1}^{n_I} \sum_{j=1}^k \dot{I}_j \right] + \frac{c}{\gamma \beta / s} \frac{d}{dt} \ln(T) + \dot{T} / \gamma \tag{35}$$

Integrating Eq. (35) from 0 to $+\infty$ and using the fact that $E_i(0) = E_i(\infty) = I_j(0) = I_j(\infty) = V(\infty) = 0$ for $i = 1, 2, \dots, n_E$ and $j = 1, 2, \dots, n_I$, it follows that,

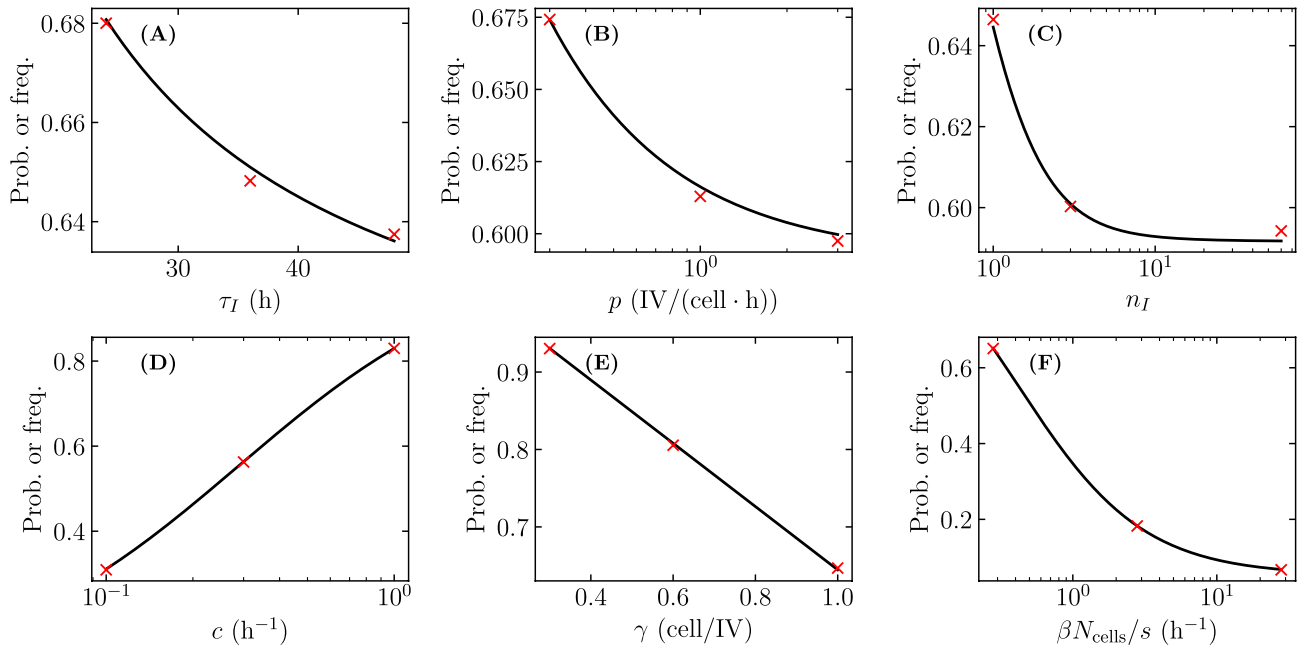


Figure 15. Extinction probability given an infection initiated with only one infectious virion. Frequency of failure of an infection initiated with only one infectious virion to cause more than 0.01% of cells to be infected by the end for 10^5 SM simulations (red scatter) compared to $\mathcal{P}_{V \rightarrow \text{Extinction}}$ the derived expression for the extinction probability given an infection initiated with only one infectious virion (black curve, Eq. (30)) while either varying (A) τ_I , (B) p , (C) n_I , (D) c , (E) γ or (F) $\beta N_{\text{cells}}/s$. Unless otherwise specified, the parameters were the same as in Fig. 4.

$$-V(0) = -p\tau_I\{T(\infty) - T(0)\} + \frac{c}{\gamma\beta/s}\{\ln[T(\infty)] - \ln[T(0)]\} + (1/\gamma)\{T(\infty) - T(0)\} \quad (36)$$

Dividing by $T(0) = N_{\text{cells}}$ we have

$$\begin{aligned} -V(0)/N_{\text{cells}} &= (1/\gamma - p\tau_I)[T(\infty)/N_{\text{cells}} - 1] + \frac{c}{\gamma\beta N_{\text{cells}}/s} \ln(T(\infty)/N_{\text{cells}}) \\ -V(0)/N_{\text{cells}} + 1/\gamma - p\tau_I &= (1/\gamma - p\tau_I)T(\infty)/N_{\text{cells}} + \frac{c}{\gamma\beta N_{\text{cells}}/s} \ln(T(\infty)/N_{\text{cells}}) \\ -\gamma V(0)/N_{\text{cells}} + 1 - \gamma p\tau_I &= (1 - \gamma p\tau_I)T(\infty)/N_{\text{cells}} + \frac{c}{\beta N_{\text{cells}}/s} \ln(T(\infty)/N_{\text{cells}}) \\ -\frac{\gamma\beta/s \cdot V(0)}{c} - \frac{(\beta N_{\text{cells}}/s) \cdot (\gamma p\tau_I - 1)}{c} &= -\frac{(\beta N_{\text{cells}}/s) \cdot (\gamma p\tau_I - 1)}{c} T(\infty)/N_{\text{cells}} + \ln(T(\infty)/N_{\text{cells}}) \end{aligned} \quad (37)$$

Introducing a new quantity we call the critical fraction of cells uninfected, $T^*/N_{\text{cells}} = [c/(\beta N_{\text{cells}}/s)]/[\gamma p\tau_I - 1]$, into Eq. (37) yields

$$\begin{aligned} -\frac{\gamma\beta/s \cdot V(0)}{c} - \frac{1}{T^*/N_{\text{cells}}} &= -\frac{T(\infty)/N_{\text{cells}}}{T^*/N_{\text{cells}}} + \ln(T(\infty)/N_{\text{cells}}) \\ e^{-[\gamma\beta/s \cdot V(0)]/c} e^{-1/(T^*/N_{\text{cells}})} &= T(\infty)/N_{\text{cells}} \cdot e^{-(T(\infty)/N_{\text{cells}})/(T^*/N_{\text{cells}})} \\ -\frac{e^{-1/(T^*/N_{\text{cells}})}}{T^*/N_{\text{cells}}} e^{-[\gamma\beta/s \cdot V(0)]/c} &= -\frac{T(\infty)/N_{\text{cells}}}{T^*/N_{\text{cells}}} \cdot e^{-(T(\infty)/N_{\text{cells}})/(T^*/N_{\text{cells}})} \end{aligned} \quad (38)$$

Eq. (38) is of the form $z = w \exp(w)$. The Lambert W function is the function that gives the inverse relation, i.e. $W(z) = w$. Therefore, we have the following,

$$\begin{aligned} -\frac{T(\infty)/N_{\text{cells}}}{T^*/N_{\text{cells}}} &= W_0\left(-\frac{e^{-1/(T^*/N_{\text{cells}})}}{T^*/N_{\text{cells}}} e^{-[\gamma\beta/s \cdot V(0)]/c}\right) \\ T(\infty)/N_{\text{cells}} &= -T^*/N_{\text{cells}} \cdot W_0\left(-\frac{e^{-1/(T^*/N_{\text{cells}})}}{T^*/N_{\text{cells}}} e^{-[\gamma\beta/s \cdot V(0)]/c}\right) \end{aligned} \quad (39)$$

where here we make use of W_0 the upper branch of the Lambert W function. The upper branch is used because $T(\infty)/T^* < 1$ since $T(\infty)$ is always reached after T^* and therefore is always smaller.

The fraction of cells consumed by the infection, $1 - T(\infty)/N_{\text{cells}}$, is then given by

$$1 - T(\infty)/N_{\text{cells}} = 1 + T^*/N_{\text{cells}} \cdot W_0\left(-\frac{e^{-1/(T^*/N_{\text{cells}})}}{T^*/N_{\text{cells}}} e^{-[\gamma\beta/s \cdot V(0)]/c}\right) \tag{40}$$

If instead there was initially a number of infectious cells but no infectious virions ($I_1(0) \neq 0, V(0) = 0$) then Eq. (36) would be

$$0 = -p\tau_I\{T(\infty) - T(0) - I_1(0)\} + \frac{c}{\gamma\beta/s}\{\ln[T(\infty)] - \ln[T(0)]\} + (1/\gamma)\{T(\infty) - T(0)\} \tag{41}$$

which is equivalent to Eq. (36) where $V(0)$ is replaced with $p\tau_I I_1(0)$. This means that Eq. (40) would be the same but with $V(0)$ replaced with $p\tau_I I_1(0)$.

For $V(0)$ such that $[\gamma\beta/s \cdot V(0)]/c \approx 0$, Eq. (40) simplifies to

$$1 - T(\infty)/N_{\text{cells}} = 1 + T^*/N_{\text{cells}} \cdot W_0\left(-\frac{e^{-1/(T^*/N_{\text{cells}})}}{T^*/N_{\text{cells}}}\right) \tag{42}$$

The Maclaurin series of $W_0(x)$ is $W_0(x) = x - x^2 + 3x^3/2 - \dots \approx x$. It follows that, when $T^*/N_{\text{cells}} \ll 1$, Eq. (42) simplifies to

$$1 - T(\infty)/N_{\text{cells}} \approx 1 + T^*/N_{\text{cells}} \cdot -\frac{e^{-1/(T^*/N_{\text{cells}})}}{T^*/N_{\text{cells}}} \tag{43}$$

$$1 - T(\infty)/N_{\text{cells}} \approx 1 - e^{-1/(T^*/N_{\text{cells}})}$$

Figure 16 shows that the fraction of cells consumed by the infection, $1 - T(\infty)/N_{\text{cells}}$, is simply a strictly decreasing function of T^*/N_{cells} and that, when $T^*/N_{\text{cells}} \ll 1, 1 - T(\infty)/N_{\text{cells}} \approx 1 - e^{-1/(T^*/N_{\text{cells}})}$. But what does T^*/N_{cells} represent? At the start of an infection, all cells are uninfected, $T(0) = N_{\text{cells}}$, and the reproductive number, $R(0)$, i.e. the number of successful secondary infections caused by an infected cell over its lifespan, corresponds to the basic reproductive number (R_0), given by Eq. (7), where $R_0 > 1$ for an infection that has a non-zero establishment probability. As the infection progresses and fewer uninfected cells remain, each infectious virion has fewer opportunities to cause a successful cell infection and $R(t)$ decreases.

Let us derive an expression for the reproductive number $R(t)$ as a function of the fraction of cells uninfected $T(t)/N_{\text{cells}}$ for a given basic reproductive number R_0 (Eq. (7)). To begin, using Eq. (7) and the left-hand side of Eq. (10), we have the following relation,

$$\frac{1}{R_0} \frac{p\tau_I \cdot \gamma\beta N_{\text{cells}}/s}{c + \beta N_{\text{cells}}/s} = \frac{p\tau_I \cdot \gamma\beta T^*/s}{c + \beta T^*/s}$$

$$\frac{1}{R_0} \frac{N_{\text{cells}}}{c/(\beta/s) + N_{\text{cells}}} = \frac{T^*}{c/(\beta/s) + T^*}$$

$$N_{\text{cells}}[c/(\beta/s)] + N_{\text{cells}}T^* = R_0[c/(\beta/s)]T^* + R_0N_{\text{cells}}T^* \tag{44}$$

$$c/(\beta/s) \cdot (N_{\text{cells}} - R_0T^*) = N_{\text{cells}}T^*(R_0 - 1)$$

$$c/(\beta/s) = \frac{N_{\text{cells}}T^*(R_0 - 1)}{N_{\text{cells}} - R_0T^*}$$

The left-hand side of Eq. (10) can be rearranged to

$$\gamma p\tau_I = \frac{c + \beta T^*/s}{\beta T^*/s} \tag{45}$$

$$\gamma p\tau_I = \frac{c/(\beta/s) + T^*}{T^*}$$

Substituting Eq. (44) into Eq. (45), we obtain

$$\gamma p\tau_I = \frac{1}{T^*} \left[\frac{N_{\text{cells}}T^*(R_0 - 1)}{N_{\text{cells}} - R_0T^*} + T^* \right]$$

$$\gamma p\tau_I = \frac{N_{\text{cells}}(R_0 - 1)}{N_{\text{cells}} - R_0T^*} + 1 \tag{46}$$

$$\gamma p\tau_I = \frac{R_0(N_{\text{cells}} - T^*)}{N_{\text{cells}} - R_0T^*}$$

The reproductive number can also be written as

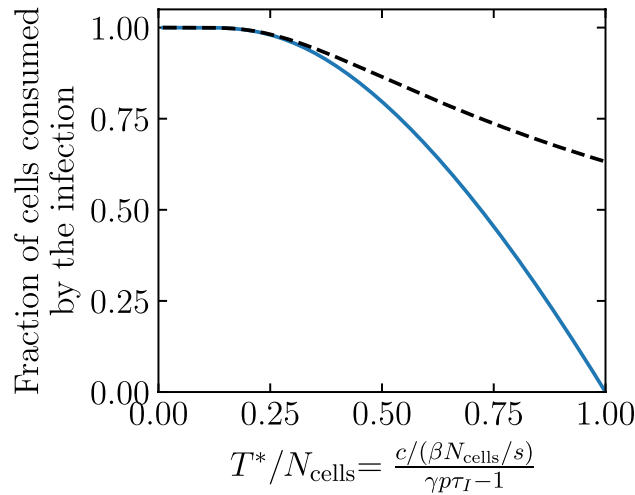


Figure 16. Fraction of cells consumed by infections in the MFM. The fraction of cells consumed by infection in the MFM $(1 - T(\infty)/N_{\text{cells}}$, Eq. (42), blue solid line) as a function of the fraction of cells uninfected when the reproductive number $R(t) = 1 (T^*/N_{\text{cells}})$. The black dashed line is $1 - \exp[-1/(T^*/N_{\text{cells}})]$.

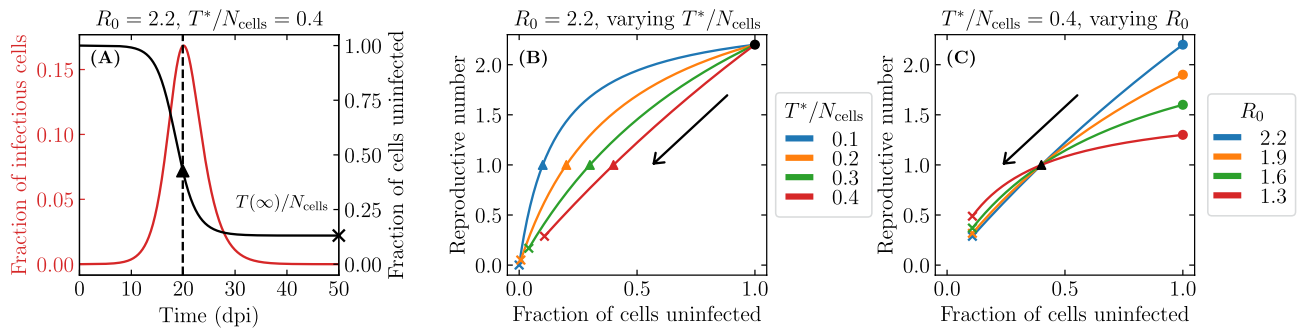


Figure 17. Exploring the critical fraction of uninfected cells, T^*/N_{cells} . (A) MFM time courses for the fraction of uninfected (black) and infectious (red) cells, for an infection initiated with 10 infectious virions under antiviral therapy reducing the virus entry rate, $\beta \rightarrow (1 - \varepsilon)\beta$, at efficacy $\varepsilon = 0.81$, which yields an establishment probability ($\mathcal{P}_{V \rightarrow \text{Establishment}} = 48\%$) that is $\sim 50\%$ of its value without antivirals²³. The point $(t, T^*/N_{\text{cells}})$ is represented by a triangle; the time when $T = T^*$ by a dashed line; and $T = T(\infty)$ by an \times on the right vertical axis. The parameters are provided in Methods, but notably $n_I = 1$. (B,C) The reproductive number, $R(t)$, as a function of the fraction of cells that remain uninfected, $T(t)$, over the course of an infection (time is implicit), based on Eq. (48) when either $R_0 = 2.2$ and T^*/N_{cells} is varied (B); or $T^*/N_{\text{cells}} = 0.4$ and R_0 is varied (C). The start of the infection is represented by a circle; the critical point where $T = T^*$ and $R = 1$ by a triangle; and the end of the infection where $T = T(\infty)$ as given by Eq. (42) by an \times .

$$R(t) = \frac{p\tau_I \cdot \gamma\beta T(t)/s}{c + \beta T(t)/s} \tag{47}$$

$$R(t) = \frac{\gamma p\tau_I \cdot T(t)}{c/(\beta/s) + T(t)}$$

Substituting Eqs. (44) and (45) into Eq. (47), we then have

$$R(t) = \frac{\frac{R_0(N_{\text{cells}} - T^*)}{N_{\text{cells}} - R_0 T^*} \cdot T(t)}{\frac{N_{\text{cells}} T^* (R_0 - 1)}{N_{\text{cells}} - R_0 T^*} + T(t)}$$

$$R(t) = \frac{R_0 T(t) \cdot (N_{\text{cells}} - T^*)}{N_{\text{cells}} T^* (R_0 - 1) + T(t) \cdot (N_{\text{cells}} - R_0 T^*)} \tag{48}$$

$$R(t) = \frac{R_0 \cdot T(t)/N_{\text{cells}} \cdot (1 - T^*/N_{\text{cells}})}{T^*/N_{\text{cells}} \cdot (R_0 - 1) + T(t)/N_{\text{cells}} \cdot (1 - R_0 \cdot T^*/N_{\text{cells}})}$$

As an infection progresses from $T(0) = N_{\text{cells}}$ where $R(t = 0) = R_0$, eventually the number of uninfected cells reaches a critical value, $T(t) = T^*$, such that

$$R(t) = 1 \equiv \frac{p\tau_I \cdot \gamma}{1 + c/(\beta T^*/s)} \cdot \frac{T^*}{N_{\text{cells}}} = \frac{c/(\beta N_{\text{cells}}/s)}{\gamma p\tau_I - 1} \tag{49}$$

When the number of uninfected cells remaining $T(t)$ equals T^* , then by definition $R(t) = 1$ marking the infectious cell population peak, as shown in Fig. 17(A). Thereafter, the infectious cell population declines and the number of cells that remain uninfected approaches its final value, $T(\infty)$. When the initial inoculum consists of one or a few infectious virions, the fraction of cells consumed by infections in the MFM or by established infections in the SM, $1 - T(\infty)/N_{\text{cells}}$, can be represented by a strictly decreasing function of T^*/N_{cells} , Eq. (42). This makes sense because if the infection parameters are such that fewer cells need to be consumed by the infection to reach $R(t) = 1$ (larger T^*), then fewer cells will have been consumed by the time the infection ends (larger $T(\infty)$), as shown in Fig. 17(B). This theoretical expression for the fraction of cells consumed is not, however, simply a function of R_0 . Figure 17(C) shows how an infection can start with a higher R_0 but $R(t)$ can decrease more rapidly as the infection progresses such that $R(t) = 1$ with the same fraction of uninfected cells (T^*/N_{cells}), ultimately resulting in the same fraction of cells consumed, $1 - T(\infty)/N_{\text{cells}}$.

Generalizing impact of antiviral reducing γ versus one reducing β or p

Impact on the establishment probability

γ vs β The establishment probability, or $1 -$ (extinction probability) (Eq. (8)), for an infection initiated with a single infectious virion is given by

$$\mathcal{P}_{V \rightarrow \text{Establishment}} = \mathcal{P}_{V \rightarrow I} - \mathcal{P}_{V \rightarrow I} \left[\frac{\mathcal{B} \cdot \mathcal{P}_{V \rightarrow \text{Establishment}}}{n_I} + 1 \right]^{-n_I} \tag{50}$$

While an antiviral reducing β decreases both the numerator and denominator of $\mathcal{P}_{V \rightarrow I}$ (Eq. (5)), one reducing γ decreases only the numerator. Therefore, at equal efficacy, an antiviral reducing γ decreases $\mathcal{P}_{V \rightarrow I}$ more than one reducing β . If $\mathcal{P}_{V \rightarrow I}$ decreases monotonically as $\mathcal{P}_{V \rightarrow \text{Establishment}}$ decreases, an antiviral causing a greater decrease in $\mathcal{P}_{V \rightarrow I}$ will cause an equal or greater decrease in $\mathcal{P}_{V \rightarrow \text{Establishment}}$.

Let us show that $\mathcal{P}_{V \rightarrow I}$ is a monotonically increasing function of $\mathcal{P}_{V \rightarrow \text{Establishment}}$. Eq. (50) can be rewritten as

$$\mathcal{P}_{V \rightarrow I} = \frac{\mathcal{P}_{V \rightarrow \text{Establishment}}}{1 - \left[\frac{\mathcal{B} \cdot \mathcal{P}_{V \rightarrow \text{Establishment}}}{n_I} + 1 \right]^{-n_I}} = \frac{x}{1 - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I}} \tag{51}$$

where we define $x \equiv \mathcal{P}_{V \rightarrow \text{Establishment}}$ for convenience. Taking the derivative of the above with respect to x yields

$$\begin{aligned} \frac{\partial \mathcal{P}_{V \rightarrow I}}{\partial x} &= \frac{\partial}{\partial x} \left(\frac{x}{1 - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I}} \right) \\ &= \frac{\left(1 - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I} \right) \cdot \frac{\partial}{\partial x}(x) - x \cdot \frac{\partial}{\partial x} \left(1 - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I} \right)}{\left(1 - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I} \right)^2} \\ &= \frac{1 - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I} - \mathcal{B} \cdot x \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I-1}}{\left(1 - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I} \right)^2} \\ &= \frac{\left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I-1} \left(\left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{n_I+1} - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right] - \mathcal{B} \cdot x \right)}{\left(1 - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I} \right)^2} \\ &= \frac{\left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{n_I+1} - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right] - \mathcal{B} \cdot x}{\left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{n_I+1} \left(1 - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I} \right)^2} \\ \frac{\partial \mathcal{P}_{V \rightarrow I}}{\partial x} &= \frac{\left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right] \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{n_I} - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right] - \mathcal{B} \cdot x}{\left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{n_I+1} \left(1 - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I} \right)^2} \end{aligned} \tag{52}$$

Using the binomial theorem, Eq. (52) can be written as

$$\frac{\partial \mathcal{P}_{V \rightarrow I}}{\partial x} = \frac{\left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right] \sum_{k=0}^{n_I} \binom{n_I}{k} \left[\frac{\mathcal{B} \cdot x}{n_I} \right]^k - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right] - \mathcal{B} \cdot x}{\left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{n_I+1} \left(1 - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I} \right)^2}$$

Since $n_I \geq 1$, we can expand the sum at least for $k = 0$ and $k = 1$, namely,

$$\begin{aligned} \frac{\partial \mathcal{P}_{V \rightarrow I}}{\partial x} &= \frac{\left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right] \binom{n_I}{0} \left[\frac{\mathcal{B} \cdot x}{n_I} \right]^0 + \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right] \binom{n_I}{1} \left[\frac{\mathcal{B} \cdot x}{n_I} \right]^1 + \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right] \sum_{k=2}^{n_I} \binom{n_I}{k} \left[\frac{\mathcal{B} \cdot x}{n_I} \right]^k - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right] - \mathcal{B} \cdot x}{\left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{n_I+1} \left(1 - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I} \right)^2} \\ &= \frac{\left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right] + \mathcal{B} \cdot x \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right] + \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right] \sum_{k=2}^{n_I} \binom{n_I}{k} \left[\frac{\mathcal{B} \cdot x}{n_I} \right]^k - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right] - \mathcal{B} \cdot x}{\left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{n_I+1} \left(1 - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I} \right)^2} \\ \frac{\partial \mathcal{P}_{V \rightarrow I}}{\partial x} &= \frac{\frac{(\mathcal{B} \cdot x)^2}{n_I} + \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right] \sum_{k=2}^{n_I} \binom{n_I}{k} \left[\frac{\mathcal{B} \cdot x}{n_I} \right]^k}{\left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{n_I+1} \left(1 - \left[\frac{\mathcal{B} \cdot x}{n_I} + 1 \right]^{-n_I} \right)^2}. \end{aligned} \tag{53}$$

Since $(\mathcal{B} \cdot x/n_I) + 1 \in [1, \mathcal{B}]$, $\frac{\partial \mathcal{P}_{V \rightarrow I}}{\partial x} = \frac{\partial \mathcal{P}_{V \rightarrow I}}{\partial \mathcal{P}_{V \rightarrow \text{Establishment}}} \geq 0 \forall (x, \mathcal{B}, n_I)$. This means that $\mathcal{P}_{V \rightarrow I}$ is a monotonically increasing function of $\mathcal{P}_{V \rightarrow \text{Establishment}}$. Since an antiviral reducing γ causes a greater decrease in $\mathcal{P}_{V \rightarrow I}$ than one reducing β , at equal efficacy an antiviral reducing γ will decrease $\mathcal{P}_{V \rightarrow \text{Establishment}}$ *at least as much or more so* than one reducing β .

γ vs p The establishment probability, or $1 -$ (extinction probability) (Eq. (8)), for an infection initiated with a single infectious virion in the presence of an antiviral reducing γ is given by

$$\mathcal{P}_{V \rightarrow \text{Establishment}}^{(\gamma)} = (1 - \varepsilon) \mathcal{P}_{V \rightarrow I} - (1 - \varepsilon) \mathcal{P}_{V \rightarrow I} \left[\frac{\mathcal{B} \cdot \mathcal{P}_{V \rightarrow \text{Establishment}}^{(\gamma)}}{n_I} + 1 \right]^{-n_I}, \tag{54}$$

because γ is in the numerator of $\mathcal{P}_{V \rightarrow I}$, whereas an antiviral reducing p is given by

$$\mathcal{P}_{V \rightarrow \text{Establishment}}^{(p)} = \mathcal{P}_{V \rightarrow I} - \mathcal{P}_{V \rightarrow I} \left[\frac{(1 - \varepsilon) \mathcal{B} \cdot \mathcal{P}_{V \rightarrow \text{Establishment}}^{(p)}}{n_I} + 1 \right]^{-n_I}, \tag{55}$$

because p is in $\mathcal{B} = p\tau_I$ but does not appear in $\mathcal{P}_{V \rightarrow I}$, such that

$$\mathcal{P}_{V \rightarrow \text{Establishment}}^{(\gamma)} = (1 - \varepsilon) \mathcal{P}_{V \rightarrow \text{Establishment}}^{(p)}. \tag{56}$$

Since $\varepsilon \in (0, 1]$, an antiviral reducing γ will always decrease the establishment probability ($\mathcal{P}_{V \rightarrow \text{Establishment}}$) *at least as much or more so* than an antiviral reducing p .

Impact on the fraction of cells consumed by established infections

The median fraction of cells consumed by established infections, Eq. (40), decreases monotonically as you decrease the argument of W_0 by reducing either the term $[\gamma(\beta/s)/c \cdot V(0)]$ or the term $N_{\text{cells}}/T^* = (\beta/s)/c \cdot [\gamma p\tau_I - 1]$ or both.

γ versus β At a given efficacy, an antiviral reducing γ or β decreases the first term by the same amount, but one reducing γ will decrease N_{cells}/T^* more than one reducing β . Therefore, an antiviral reducing γ will always decrease the fraction of cells consumed *at least as much or more so* than an antiviral reducing β .

γ versus p At a given efficacy, an antiviral reducing γ or p decreases N_{cells}/T^* by the same amount, but one reducing γ will also reduce the first term whereas one reducing p has no effect on the first term. Therefore, an antiviral reducing γ will always decrease the fraction of cells consumed *at least as much or more so* than an antiviral reducing p .

Data availability

There are no primary data in the paper. The code for the mean-field model (MFM) and the stochastic model (SM) is freely available on GitHub (<https://github.com/cquir/vir-inf-stoch>).

Received: 24 August 2022; Accepted: 4 October 2023

Published online: 11 October 2023

References

1. Conway, J. M. & Coombs, D. A stochastic model of latently infected cell reactivation and viral blip generation in treated HIV patients. *PLoS Comput. Biol.* 7(4), e1002033. <https://doi.org/10.1371/journal.pcbi.1002033> (2011).
2. Pearson, J. E., Krapivsky, P. & Perelson, A. S. Stochastic theory of early viral infection: Continuous versus burst production of virions. *PLoS Comput. Biol.* 7(2), e1001058. <https://doi.org/10.1371/journal.pcbi.1001058> (2011).

3. Conway, J. M., Konrad, B. P. & Coombs, D. Stochastic analysis of pre- and postexposure prophylaxis against HIV infection. *SIAM J. Appl. Math.* **73**(2), 904–928. <https://doi.org/10.1137/120876800> (2013).
4. Heldt, F. S., Kupke, S. Y., Dorl, S., Reichl, U. & Frensing, T. Single-cell analysis and stochastic modelling unveil large cell-to-cell variability in influenza A virus infection. *Nat. Commun.* **6**, 8938. <https://doi.org/10.1038/ncomms9938> (2015).
5. Yan, A. W., Cao, P. & McCaw, J. M. On the extinction probability in models of within-host infection: The role of latency and immunity. *J. Math. Biol.* **73**(4), 787–813. <https://doi.org/10.1007/s00285-015-0961-5> (2016).
6. Conway, J. M. & Perelson, A. S. Early HIV infection predictions: Role of viral replication errors. *SIAM J. Appl. Math.* **78**(4), 1863–1890. <https://doi.org/10.1137/17M1134019> (2018).
7. Sazonov, I., Grebennikov, D., Kelbert, M., Meyerhans, A. & Bocharov, G. Viral infection dynamics model based on a Markov process with time delay between cell infection and progeny production. *Mathematics* **8**(8), 1207. <https://doi.org/10.3390/math8081207> (2020).
8. Permanyer, M., Ballana, E. & Esté, J. A. Endocytosis of HIV: Anything goes. *Trends Microbiol.* **18**(12), 543–551. <https://doi.org/10.1016/j.tim.2010.09.003> (2010).
9. Fredericksen, B. L., Wei, B. L., Yao, J., Luo, T. & Garcia, J. V. Inhibition of endosomal/lysosomal degradation increases the infectivity of human immunodeficiency virus. *J. Virol.* **76**(22), 11440–11446. <https://doi.org/10.1128/jvi.76.22.11440-11446.2002> (2002).
10. Wei, B. L. *et al.* Inhibition of lysosome and proteasome function enhances human immunodeficiency virus type 1 infection. *J. Virol.* **79**(9), 5705–5712. <https://doi.org/10.1128/JVI.79.9.5705-5712.2005> (2005).
11. Burdick, R. C., Hu, W.-S. & Pathak, V. K. Nuclear import of APOBEC3F-labeled HIV-1 preintegration complexes. *Proc. Natl. Acad. Sci. U. S. A.* **110**(49), E4780–E4789. <https://doi.org/10.1073/pnas.1315996110> (2013).
12. Burdick, R. C. *et al.* Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes. *PLoS Pathog.* **13**(8), e1006570. <https://doi.org/10.1371/journal.ppat.1006570> (2017).
13. Ho, Y.-C. *et al.* Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell* **155**(3), 540–551. <https://doi.org/10.1016/j.cell.2013.09.020> (2013).
14. Colomer-Lluch, M., Ruiz, A., Moris, A. & Prado, J. G. Restriction factors: From intrinsic viral restriction to shaping cellular immunity against HIV-1. *Front. Immunol.* **9**, 2876. <https://doi.org/10.3389/fimmu.2018.02876> (2018).
15. Kieffer, T. L. *et al.* G → A hypermutation in protease and reverse transcriptase regions of human immunodeficiency virus type 1 residing in resting CD4+ T cells in vivo. *J. Virol.* **79**(3), 1975–1980. <https://doi.org/10.1128/JVI.79.3.1975-1980.2005> (2005).
16. Koff, W. C. & Knight, V. Inhibition of influenza virus uncoating by rimantadine hydrochloride. *J. Virol.* **31**(1), 261–263. <https://doi.org/10.1128/JVI.31.1.261-263.1979> (1979).
17. Qin, C. *et al.* Real-time dissection of dynamic uncoating of individual influenza viruses. *Proc. Natl. Acad. Sci. U. S. A.* **116**(7), 2577–2582. <https://doi.org/10.1073/pnas.1812632116> (2019).
18. Schelker, M. *et al.* Viral RNA degradation and diffusion act as a bottleneck for the influenza A virus infection efficiency. *PLoS Comput. Biol.* **12**(10), e1005075. <https://doi.org/10.1371/journal.pcbi.1005075> (2016).
19. Flatt, J. W. & Greber, U. F. G. Mismatch at the nuclear pore complex-stopping a virus dead in its tracks. *Cells* **4**(3), 277–296. <https://doi.org/10.3390/cells4030277> (2015).
20. Xiao, H., Killip, M. J., Staeheli, P., Randall, R. E. & Jackson, D. The human interferon-induced MxA protein inhibits early stages of influenza A virus infection by retaining the incoming viral genome in the cytoplasm. *J. Virol.* **87**(23), 13053–13058. <https://doi.org/10.1128/JVI.02220-13> (2013).
21. Desai, T. M. *et al.* IFITM3 restricts influenza A virus entry by blocking the formation of fusion pores following virus-endosome hemifusion. *PLoS Pathog.* **10**(4), e1004048. <https://doi.org/10.1371/journal.ppat.1004048> (2014).
22. Villalón-Letelier, F., Brooks, A. G., Saunders, P. M., Londrigan, S. L. & Reading, P. C. Host cell restriction factors that limit influenza A infection. *Viruses* **9**(12), 376. <https://doi.org/10.3390/v9120376> (2017).
23. Czuppon, P. *et al.* Success of prophylactic antiviral therapy for SARS-CoV-2: Predicted critical efficacies and impact of different drug-specific mechanisms of action. *PLoS Comput. Biol.* **17**(3), e1008752. <https://doi.org/10.1371/journal.pcbi.1008752> (2021).
24. Dobrovolny, H. M. & Beauchemin, C. A. A. Modelling the emergence of influenza drug resistance: The roles of surface proteins, the immune response and antiviral mechanisms. *PLoS ONE* **12**(7), e0180582. <https://doi.org/10.1371/journal.pone.0180582> (2017).
25. Beauchemin, C. A. A. *et al.* Modeling amantadine treatment of influenza A virus in vitro. *J. Theor. Biol.* **254**(2), 439–451. <https://doi.org/10.1016/j.jtbi.2008.05.031> (2008).
26. Liu, T., Luo, S., Libby, P. & Shi, G.-P. Cathepsin L-selective inhibitors: A potentially promising treatment for COVID-19 patients. *Pharmacol. Ther.* **213**, 107587. <https://doi.org/10.1016/j.pharmthera.2020.107587> (2020).
27. Holder, B. P. & Beauchemin, C. A. A. Exploring the effect of biological delays in kinetic models of influenza within a host or cell culture. *BMC Public Health* **11**(S1), S10. <https://doi.org/10.1186/1471-2458-11-S1-S10> (2011).
28. Beauchemin, C. A. A., Miura, T. & Iwami, S. Duration of SHIV production by infected cells is not exponentially distributed: Implications for estimates of infection parameters and antiviral efficacy. *Sci. Rep.* **7**, 42765. <https://doi.org/10.1038/srep42765> (2017).
29. Liao, L. E. *et al.* Quantification of Ebola virus replication kinetics in vitro. *PLOS Comput. Biol.* **16**(11), e1008375. <https://doi.org/10.1371/journal.pcbi.1008375> (2020).
30. Kakizoe, Y. *et al.* A method to determine the duration of the eclipse phase for in vitro infection with a highly pathogenic SHIV strain. *Sci. Rep.* **5**, 10371. <https://doi.org/10.1038/srep10371> (2015).
31. Petravic, J. *et al.* Intracellular dynamics of HIV infection. *J. Virol.* **88**(2), 1113–1124. <https://doi.org/10.1128/JVI.02038-13> (2014).
32. Paradis, E. G. *et al.* Impact of the H275Y and I223V mutations in the neuraminidase of the 2009 pandemic influenza virus in vitro and evaluating experimental reproducibility. *PLoS ONE* **10**(5), e0126115. <https://doi.org/10.1371/journal.pone.0126115> (2015).
33. Pinilla, L. T., Holder, B. P., Abed, Y., Boivin, G. & Beauchemin, C. A. A. The H275Y neuraminidase mutation of the pandemic A/H1N1 virus lengthens the eclipse phase and reduces viral output of infected cells, potentially compromising fitness in ferrets. *J. Virol.* **86**(19), 10651–10660. <https://doi.org/10.1128/JVI.07244-11> (2012).
34. Yan, A. W. *et al.* Quantifying mechanistic traits of influenza viral dynamics using in vitro data. *Epidemics* **33**, 100406. <https://doi.org/10.1016/j.epidem.2020.100406> (2020).
35. González-Parra, G. *et al.* A comparison of RSV and influenza in vitro kinetic parameters reveals differences in infecting time. *PLoS ONE* **13**(2), e0192645. <https://doi.org/10.1371/journal.pone.0192645> (2018).
36. González-Parra, G., Dobrovolny, H. M., Aranda, D. F., Chen-Charpentier, B. & Rojas, R. A. G. Quantifying rotavirus kinetics in the REH tumor cell line using in vitro data. *Virus Res.* **244**, 53–63. <https://doi.org/10.1016/j.virusres.2017.09.023> (2018).
37. Noecker, C. *et al.* Simple mathematical models do not accurately predict early SIV dynamics. *Viruses* **7**(3), 1189–1217. <https://doi.org/10.3390/v7031189> (2015).
38. Bai, F., Huff, K. E. & Allen, L. J. The effect of delay in viral production in within-host models during early infection. *J. Biol. Dyn.* **13**(sup1), 47–73. <https://doi.org/10.1080/17513758.2018.1498984> (2019).
39. Stegmann, T. *et al.* Evaluation of viral membrane fusion assays Comparison of the octadecylrhodamine dequenching assay with the pyrene excimer assay. *Biochemistry* **32**(42), 11330–11337. <https://doi.org/10.1021/bi00093a009> (1993).
40. Brooke, C. B. *et al.* Most influenza A viruses fail to express at least one essential viral protein. *J. Virol.* **87**(6), 3155–3162. <https://doi.org/10.1128/JVI.02284-12> (2013).
41. Diefenbacher, M., Sun, J. & Brooke, C. B. The parts are greater than the whole: The role of semi-infectious particles in influenza A virus biology. *Curr. Opin. Virol.* **33**, 42–46. <https://doi.org/10.1016/j.coviro.2018.07.002> (2018).

42. Liao, L. E., Iwami, S. & Beauchemin, C. A. A. (In)validating experimentally derived knowledge about influenza A defective interfering particles. *J. R. Soc. Interface* **13**(124), 20160412. <https://doi.org/10.1098/rsif.2016.0412> (2016).
43. Holford, N. H. G. & Sheiner, L. B. Understanding the dose-effect relationship: Clinical application of pharmacokinetic-pharmacodynamic models. *Clin. Pharmacokinet.* **6**(6), 429–453. <https://doi.org/10.2165/00003088-198106060-00002> (1981).
44. Gonçalves, A. *et al.* Timing of antiviral treatment initiation is critical to reduce SARS-CoV-2 viral load. *CPT Pharmacometrics Syst. Pharmacol.* **9**(9), 509–514. <https://doi.org/10.1002/psp4.12543> (2020).
45. Yuan, Y. & Allen, L. J. S. Stochastic models for virus and immune system dynamics. *Math. Biosci.* **234**(2), 84–94. <https://doi.org/10.1016/j.mbs.2011.08.007> (2011).
46. Shaw, G. M. & Hunter, E. HIV transmission. *Cold Spring Harb. Perspect. Med.* **2**(11), a006965. <https://doi.org/10.1101/cshperspect.a006965> (2012).
47. Xiao, Y., Miao, H., Tang, S. & Wu, H. Modeling antiretroviral drug responses for HIV-1 infected patients using differential equation models. *Adv. Drug Deliv. Rev.* **65**(7), 940–953. <https://doi.org/10.1016/j.addr.2013.04.005> (2013).
48. Sánchez-Taltavull, D. & Alarcón, T. Stochastic modelling of viral blips in HIV-1-infected patients: Effects of inhomogeneous density fluctuations. *J. Theor. Biol.* **371**, 79–89. <https://doi.org/10.1016/j.jtbi.2015.02.001> (2015).
49. Hayden, F. G. *et al.* Baloxavir Marboxil Investigators Group. Baloxavir Marboxil for uncomplicated influenza in adults and adolescents. *N. Engl. J. Med.* **379**(10), 913–923. <https://doi.org/10.1056/NEJMoal716197> (2018).
50. DeVincenzo, J. *et al.* Safety and antiviral effects of nebulized PC786 in a respiratory syncytial virus challenge study. *J. Infect. Dis.* **225**(12), 2087–2096. <https://doi.org/10.1093/infdis/jiaa716> (2022).
51. Hammond, J. *et al.* EPIC-HR Investigators. Oral nirmatrelvir for high-risk, nonhospitalized adults with Covid-19. *N. Engl. J. Med.* **386**(15), 1397–1408. <https://doi.org/10.1056/NEJMoa2118542> (2022).
52. DeVincenzo, J. P. *et al.* Viral load drives disease in humans experimentally infected with respiratory syncytial virus. *Am. J. Respir. Crit. Care Med.* **182**(10), 1305–1314. <https://doi.org/10.1164/rccm.201002-0221OC> (2010).
53. Duchamp, M. B. *et al.* Pandemic A (H1N1) 2009 influenza virus detection by real time RT-PCR: Is viral quantification useful?. *Clin. Microbiol. Infect.* **16**(4), 317–321. <https://doi.org/10.1111/j.1469-0691.2010.03169.x> (2010).
54. Granados, A., Peci, A., McGeer, A. & Gubbay, J. B. Influenza and rhinovirus viral load and disease severity in upper respiratory tract infections. *J. Clin. Virol.* **86**, 14–19. <https://doi.org/10.1016/j.jcv.2016.11.008> (2017).
55. Myers, M. A. *et al.* Dynamically linking influenza virus infection kinetics, lung injury, inflammation, and disease severity. *Elife* **10**, e68864. <https://doi.org/10.7554/eLife.68864> (2021).
56. Ai, T. *et al.* Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases. *Radiology* **296**(2), E32–E40. <https://doi.org/10.1148/radiol.2020200642> (2020).
57. Compton, A. A. & Schwartz, O. They might be giants: Does syncytium formation sink or spread HIV infection?. *PLoS Pathog.* **13**(2), e1006099. <https://doi.org/10.1371/journal.ppat.1006099> (2017).
58. Gamble, A. *et al.* Drivers and distribution of henipavirus-induced syncytia: What do we know?. *Viruses* **13**(9), 1755. <https://doi.org/10.3390/v13091755> (2021).
59. Rajah, M. M., Bernier, A., Buchrieser, J. & Schwartz, O. The mechanism and consequences of SARS-CoV-2 spike-mediated fusion and syncytia formation. *J. Mol. Biol.* **434**(6), 167280. <https://doi.org/10.1016/j.jmb.2021.167280> (2022).
60. Nardacci, R. *et al.* Syncytial apoptosis signaling network induced by the HIV-1 envelope glycoprotein complex: An overview. *Cell Death Dis.* **6**(8), e1846–e1846 (2015).
61. Cifuentes-Munoz, N., Najjar, F. E. & Dutch, R. E. Viral cell-to-cell spread: Conventional and non-conventional ways. *Adv. Virus Res.* **108**, 85–125. <https://doi.org/10.1016/bs.aivir.2020.09.002> (2020).
62. Quirouette, C., Younis, N. P., Reddy, M. B. & Beauchemin, C. A. A. A mathematical model describing the localization and spread of influenza A virus infection within the human respiratory tract. *PLoS Comput. Biol.* **16**(4), e1007705. <https://doi.org/10.1371/journal.pcbi.1007705> (2020).
63. Liljeroos, L., Krzyzaniak, M. A., Helenius, A. & Butcher, S. J. Architecture of respiratory syncytial virus revealed by electron cryotomography. *Proc. Natl. Acad. Sci. U. S. A.* **110**(27), 11133–11138. <https://doi.org/10.1073/pnas.1309070110> (2013).
64. Harris, A. *et al.* Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proc. Natl. Acad. U. S. A.* **103**(50), 19123–19127. <https://doi.org/10.1073/pnas.0607614103> (2006).
65. Bharat, T. A. M. *et al.* Structural dissection of Ebola virus and its assembly determinants using cryo-electron tomography. *Proc. Natl. Acad. Sci. U. S. A.* **109**(11), 4275–4280. <https://doi.org/10.1073/pnas.1120453109> (2012).
66. Hyatt, A. D., Zaki, S. R., Goldsmith, C. S., Wise, T. G. & Hengstberger, S. G. Ultrastructure of Hendra virus and Nipah virus within cultured cells and host animals. *Microbes Infect.* **3**(4), 297–306. [https://doi.org/10.1016/s1286-4579\(01\)01383-1](https://doi.org/10.1016/s1286-4579(01)01383-1) (2001).
67. Beauchemin, C. A. A., Kim, Y.-I., Yu, Q., Ciarrella, G. & DeVincenzo, J. P. Uncovering critical properties of the human respiratory syncytial virus by combining in vitro assays and in silico analyses. *PLoS ONE* **14**(4), e0214708. <https://doi.org/10.1371/journal.pone.0214708> (2019).
68. Li, T. *et al.* The shape of pleomorphic virions determines resistance to cell-entry pressure. *Nat. Microbiol.* **6**(5), 617–629. <https://doi.org/10.1038/s41564-021-00877-0> (2021).
69. Stafford, M. A. *et al.* Modeling plasma virus concentration during primary HIV infection. *J. Theor. Biol.* **203**(3), 285–301. <https://doi.org/10.1006/jtbi.2000.1076> (2000).
70. Baccam, P., Beauchemin, C., Macken, C. A., Hayden, F. G. & Perelson, A. S. Kinetics of influenza A virus infection in humans. *J. Virol.* **80**(15), 7590–7599. <https://doi.org/10.1128/JVI.01623-05> (2006).
71. Gillespie, D. T. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.* **115**(4), 1716–1733. <https://doi.org/10.1063/1.1378322> (2001).
72. Tian, T. & Burrage, K. Binomial leap methods for simulating stochastic chemical kinetics. *J. Chem. Phys.* **121**(21), 10356–10364. <https://doi.org/10.1063/1.1810475> (2004).
73. Chatterjee, A., Vlachos, D. G. & Katsoulakis, M. A. Binomial distribution based τ -leap accelerated stochastic simulation. *J. Chem. Phys.* **122**(2), 024112. <https://doi.org/10.1063/1.1833357> (2005).
74. Pettigrew, M. F. & Resat, H. Multinomial tau-leaping method for stochastic kinetic simulations. *J. Chem. Phys.* **126**(8), 084101. <https://doi.org/10.1063/1.2432326> (2007).
75. Heffernan, J. M. & Wahl, L. M. Natural variation in HIV infection: Monte Carlo estimates that include CD8 effector cells. *J. Theor. Biol.* **243**(2), 191–204. <https://doi.org/10.1016/j.jtbi.2006.05.032> (2006).
76. Allen, L. J. An introduction to stochastic processes with applications to biology. (2010).
77. Collinson, S., Khan, K. & Heffernan, J. M. The effects of media reports on disease spread and important public health measurements. *PLoS ONE* **10**(11), e0141423. <https://doi.org/10.1371/journal.pone.0141423> (2015).
78. Luo, S., Reed, M., Mattingly, J. C. & Koelle, K. The impact of host immune status on the within-host and population dynamics of antigenic immune escape. *J. R. Soc. Interface* **9**(75), 2603–2613. <https://doi.org/10.1098/rsif.2012.0180> (2012).

Acknowledgements

This work was supported in part by the Fondation Baxter & Alma Ricard (CQ), Discovery Grants 355837-2013 and 2022-03744 (CAAB) from the Natural Sciences and Engineering Research Council of Canada (www.nserc-crsng.gc.ca), KAKENHI grant 21H04595 from the Japan Society for the Promotion of Science (JSPS, www.jpsp).

go.jp), and by the Kato Sechi Female PI Incentive (www.riken.jp/en/careers/programs/kato_sechi) and Interdisciplinary Theoretical and Mathematical Sciences (iTHEMS, <http://ithems.riken.jp>) programmes at RIKEN (CAAB). This work was made possible by support provided by Calcul Québec (www.calculquebec.ca) and the Digital Research Alliance of Canada (www.alliancecan.ca).

Author contributions

C.A.A.B. oversaw the project's administration and funding acquisition. C.Q., D.C. and C.A.A.B. contributed to the conceptualization, data curation, software, research methodology, and to writing the original draft. All authors contributed to various parts of the investigation, formal analysis, and validation of the results, and to the review and editing of the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-44180-w>.

Correspondence and requests for materials should be addressed to C.A.A.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023