



Research



Cite this article: Williams B, Carruthers J, Gillard JJ, Lythe G, Perelson AS, Ribeiro RM, Molina-París C, López-García M. 2024

The reproduction number and its probability distribution for stochastic viral dynamics.

J. R. Soc. Interface **21**: 20230400.

<https://doi.org/10.1098/rsif.2023.0400>

Received: 12 July 2023

Accepted: 18 December 2023

Subject Category:

Life Sciences—Mathematics interface

Subject Areas:

biomathematics, computational biology

Keywords:

viral burst size, reproduction number, eclipse phase, infectious phase, Erlang distribution, stochastic model

Author for correspondence:

Bevelynn Williams

e-mail: bevelynnw@gmail.com

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.7007799>.

The reproduction number and its probability distribution for stochastic viral dynamics

Bevelynn Williams¹, Jonathan Carruthers², Joseph J. Gillard³, Grant Lythe¹, Alan S. Perelson⁴, Ruy M. Ribeiro⁴, Carmen Molina-París⁴ and Martín López-García¹

¹Department of Applied Mathematics, School of Mathematics, University of Leeds, Leeds, UK

²UK Health Security Agency, Salisbury, UK

³CBR Division, Defence Science and Technology Laboratory, Salisbury, UK

⁴T-6, Theoretical Biology and Biophysics, Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM, USA

BW, 0000-0002-9119-6568; JGG, 0000-0001-7486-8927; RMR, 0000-0002-3988-8241; CM-P, 0000-0001-9828-6737; ML-G, 0000-0003-3833-8595

We consider stochastic models of individual infected cells. The reproduction number, R , is understood as a random variable representing the number of new cells infected by one initial infected cell in an otherwise susceptible (target cell) population. Variability in R results partly from heterogeneity in the viral burst size (the number of viral progeny generated from an infected cell during its lifetime), which depends on the distribution of cellular lifetimes and on the mechanism of virion release. We analyse viral dynamics models with an eclipse phase: the period of time after a cell is infected but before it is capable of releasing virions. The duration of the eclipse, or the subsequent infectious, phase is non-exponential, but composed of stages. We derive the probability distribution of the reproduction number for these viral dynamics models, and show it is a negative binomial distribution in the case of constant viral release from infectious cells, and under the assumption of an excess of target cells. In a deterministic model, the ultimate in-host establishment or extinction of the viral infection depends entirely on whether the mean reproduction number is greater than, or less than, one, respectively. Here, the probability of extinction is determined by the probability distribution of R , not simply its mean value. In particular, we show that in some cases the probability of infection is not an increasing function of the mean reproduction number.

1. Introduction

Viruses, and some types of bacteria, infect host cells. When a viral particle is taken up by a host cell, the genome is replicated and used to produce viral proteins. New viral particles are then assembled inside the infected cell. For some viruses, progeny virions accumulate inside the host cell and eventually numerous viral particles exit at once in a burst, killing the cell. On the other hand, for most enveloped viruses, new viral particles are released throughout the lifetime of the infected cell via a process called budding. The total number of virions released by an infected cell during its lifetime is usually referred to as the ‘burst size’, even if the virus is released continuously by budding [1–4].

Stochasticity of the intracellular viral life cycle, coupled with heterogeneity in the lifetime of infected cells, can lead to variability in the number of virions released by individual cells. For instance, Hataye *et al.* [5] measured burst sizes from HIV-1-infected cells, and Bacsik *et al.* [6] measured production from individual influenza A virus-infected cells, with both references encountering high

variability. The progeny virions released by an infected cell may infect further host target cells, or may be degraded or cleared before they manage to do so. Hence, the number of secondary infected cells caused by a single infected cell will depend on its burst size and on the fraction of these progeny virions that are able to infect other cells.

The basic reproduction number was originally introduced in an epidemiological context to characterize the mean number of secondary infections caused by an initial infected individual in an otherwise susceptible population. Here we consider the cell-level equivalent in the context of within-host and *in vitro* viral dynamics, where the cell-level basic reproduction number will be understood as a random variable, R , representing the number of target cells infected by one initial infected cell in an otherwise susceptible (target) cell population [1,7–10]. The mean value of the basic reproduction number will be denoted by \bar{R} .

In deterministic models, the mean value of the basic reproduction number determines the outcome of infection: if $\bar{R} < 1$, then the virus-free steady state is stable and the infection will die out. In stochastic models, if the mean number of secondary infections caused by a single infected cell is smaller than one, the infection can grow, but will certainly die out eventually. Conversely, when $\bar{R} > 1$, deterministic models predict that an infection will be established, but in stochastic models there can be a non-zero probability that all virions and infected cells will be eliminated before the infection can become established [1,11]. This probability depends both on \bar{R} [12], and on the probability distribution of the reproduction number [1,11]. By exclusively focusing on the mean value of the reproduction number, one may fail to capture important dynamics of early infection events, which are inherently stochastic, particularly if an individual is exposed to a low infecting viral dose.

We study two Markov chain viral infection models, with non-exponential infectious period distributions. The models are based on one of the simplest deterministic models of viral infection, consisting of a system of ODEs for the populations of target cells, infected cells and virions [13]. One model considers viral release by budding and the other by bursting. In the model of viral release by budding, productively infected (i.e. infectious) cells are assumed to release virions at a constant rate. In the model of viral release by bursting, viral particles are produced intracellularly at a constant rate, and are eventually released in a burst when the infected cell dies. Pearson *et al.* [1] studied stochastic models similar to these, with a single infected cell compartment. We split the infected cell state into multiple compartments, leading to an Erlang-distributed, instead of an exponentially distributed, infectious period [2,7,14,15]. An Erlang distribution may be a more suitable characterization of the infectious period than an exponential one, and can lead to different estimates of model parameters [16]. The exponential case can still be recovered by setting the shape parameter of the Erlang distribution (given by the number of infected cell compartments) to one.

The deterministic model, and the starting point for our stochastic models, is successful in describing viral infections when the populations of infected cells and virions are large, and has been used to estimate viral kinetic parameters in the literature. On the other hand, the reproduction number probability distribution is particularly important to characterize the early stages of viral infection, when there may only be a few extracellular virions or infected cells. Even if the mean reproduction number is fairly large, the initial infection may

die out by chance, due to a non-zero probability that an infected cell will cause no secondary infections. This highlights the importance of looking beyond the mean of the reproduction number, and of considering other summary statistics from the probability distribution of this random variable. Previous studies that have focused on calculating the probability of viral extinction during the early stages of infection include Pearson *et al.* [1], who calculated the probability of extinction in models with either a geometric or Poisson distribution of the reproduction number. In this paper, we aim to show that deriving and quantifying the burst size and reproduction number distributions for different models and viruses can help to better understand how these distributions are affected by modelling choices and parameters, and how these distributions, in turn, affect the probability that the virus and infected cells will be eliminated before an infection becomes established.

Section 2 contains model descriptions, and derivations of the probability distributions of the burst size and reproduction number for these models. In §3, we obtain numerical results for the reproduction number probability distribution and the probability of viral extinction, in order to investigate the effects of various model assumptions and parameter values on these quantities of interest. In the electronic supplementary material, we show that similar results can be obtained for models with an age-dependent viral release rate.

2. Mathematical methods

We begin with models of viral dynamics described by ordinary differential equations (ODEs) in §2.1. We introduce a model in which infectious cells produce and release virions via budding at a constant rate, and we also consider a model in which virions are instead released in a burst. We then describe the Markov chain versions of these models in §2.2. For each model, we derive the average burst size and reproduction number (§2.3), the probability distributions of the burst size and reproduction number (§2.4 and §2.5) and the probability of viral extinction (§2.6). In the electronic supplementary material, we consider a model developed by Guedj *et al.* [17] for hepatitis C virus (HCV) infection. This model considers viral release by budding, but includes intracellular dynamics, resulting in a viral release rate from infected cells that is not constant but depends on the intracellular viral genome counts. We derive the probability distributions of the burst size and reproduction number for a stochastic version of this model, and make use of the parameter estimates from Guedj *et al.* [17] to obtain numerical results.

2.1. Deterministic models

2.1.1. Viral release by budding

Our ODE model of viral release by budding is based on those of Liao *et al.* [2] and Yan *et al.* [7], with a modification to represent removal of infected cells by the immune system [11,18]. Uninfected target cells are infected by viral particles with (per virion) rate β_c , then pass through an eclipse phase and an infectious phase. Thus, the eclipse phase is the period after viral entry but before the release of virions. The population of infected cells is partitioned into $n_E + n_I$ subsets: n_E subsets E_i , one for each eclipse-phase stage, and n_I subsets I_i , one for each infectious-phase stage. A cell in any of the

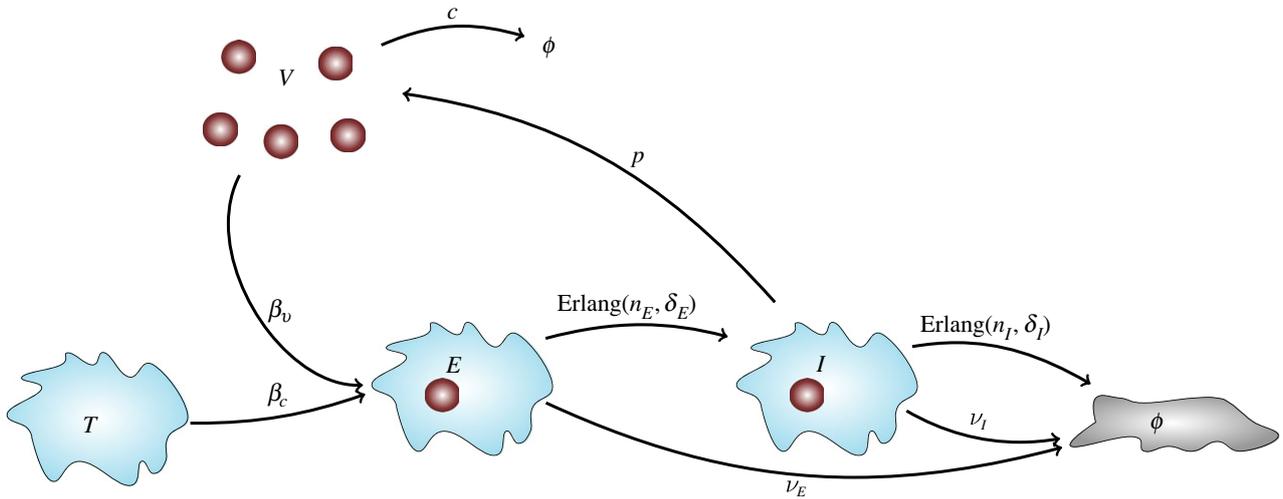


Figure 1. Model with constant viral release rate. Upon infection by a free (i.e. extracellular) infectious virion V , a target cell, T , enters the eclipse phase E followed by the infectious phase I . The arrows show the possible transitions and their corresponding rates, or the distribution of times, in the case of non-exponentially (Erlang) distributed transition times. Free virions are cleared with rate c , and eclipse-phase cells with rate v_E . Infectious cells are cleared by the immune system with rate v_I , or undergo virus-induced cell death. In (2.1), the eclipse and infectious phases consist of n_E and n_I stages, respectively.

infectious-phase stages releases virions at constant rate p . It dies upon exiting the last infectious stage. A diagram is provided in figure 1. The system of $2 + n_E + n_I$ ODEs is

$$\begin{aligned} \frac{dT}{dt} &= -\beta_c TV, \\ \frac{dE_1}{dt} &= \beta_c TV - \left(\frac{n_E}{\tau_E} + v_E\right) E_1, \\ \frac{dE_i}{dt} &= \frac{n_E}{\tau_E} (E_{i-1} - E_i) - v_E E_i, \quad i = 2, \dots, n_E, \\ \frac{dI_1}{dt} &= \frac{n_E}{\tau_E} E_{n_E} - \left(\frac{n_I}{\tau_I} + v_I\right) I_1, \\ \frac{dI_i}{dt} &= \frac{n_I}{\tau_I} (I_{i-1} - I_i) - v_I I_i, \quad i = 2, \dots, n_I, \\ \frac{dV}{dt} &= p \sum_{i=1}^{n_I} I_i - cV - \beta_v TV. \end{aligned} \quad (2.1)$$

The variables T and V denote the number of uninfected (or susceptible) target cells, and the number of free (i.e. extracellular) infectious virions. In the first ODE, some models also include production and death of uninfected target cells. However in many circumstances, these events are neglected when considering a short timescale [1]. Examples are found in models describing *in vitro* viral dynamics [2,7] or *in vivo* acute infection dynamics, where the viral load increases to a maximum, and then declines due to the depletion of target cells. Thus, we do not consider the generation of new target cells or death of uninfected target cells in our model.

The multi-stage representation of the eclipse and infectious phases in (2.1) is equivalent to the assumption that the time a cell spends in either phase follows an Erlang distribution. The Erlang distribution results from a sum of exponential random variables, and is a special case of the gamma distribution, with an integer-valued shape parameter. Here, the mean time spent in the eclipse phase (with n_E stages) is τ_E , so the mean time spent in each eclipse-phase stage is $1/\delta_E = \tau_E/n_E$. The mean time spent in the infectious phase (with n_I stages) is τ_I , so the mean time spent in each infectious-phase stage is $1/\delta_I = \tau_I/n_I$.

An additional mechanism of infected cell death due to the immune response, even before they release virions, has been included in the model by specifying that eclipse-phase cells are cleared with rate v_E , and infectious-phase cells with rate v_I [18]. Eclipse-phase cells have been found to express viral peptides, so $CD8^+$ T cells can recognize and kill them [19,20]. If eclipse-phase cells are killed at a slower rate than infectious-phase ones then $v_E < v_I$. The model without immune response can be recovered by setting $v_E = v_I = 0$, for instance, in order to represent *in vitro* dynamics [2,7]. When $v_E = v_I = 0$, the duration of the eclipse and infectious phases follow Erlang distributions. However, when $v_E > 0$ and $v_I > 0$, the distributions are altered by the possibility of cell death due to the immune response; that is, the contingency that the eclipse or infectious phase ‘completes early’ due to the cell being killed by the immune response.

Some of the virions released by infectious cells are non-infectious and cannot, therefore, infect new target cells. Here we assume that p denotes the release rate of *infectious* virions, and that the variable V represents the number of *infectious* viral particles. Free viral particles are cleared with rate c , which can be thought of as the sum of the rates of viral degradation and of loss due to antibody binding. In the last ODE, the term $\beta_v TV$ corresponding to loss of virions due to infection is explicitly included, although it is often neglected in the literature based on being small compared with the term cV [21]. Furthermore, if the number of target cells is assumed to remain constant, T_0 , both viral loss terms can be incorporated into one parameter, $c' = c + \beta_v T_0$ [22].

The parameters β_c and β_v in (2.1) have different dimensions: β_c has dimensions of per virion per time, whereas β_v has dimensions of per cell per time. Let us introduce β , the rate at which infection events occur, with units of per cell per virion per time. If n_c and n_v are the number of cells and virions involved in a general infection event, respectively, then $\beta_c = n_c \beta$ and $\beta_v = n_v \beta$. Here, since an infection event in the model consists of one virion infecting one target cell, we have $n_c = n_v = 1$. Therefore, numerically we have $\beta_c = \beta_v = \beta$, and from now on we will simply denote both rates by β .

An important quantity that can be calculated from (2.1), sometimes called the basic reproduction number, is the

mean number of infected cells produced by a single infected cell. When $v_E = v_I = 0$, the mean number of virions released by a single infected cell is $p\tau_I$. If the number of target cells present is T_0 , then each of the released virions infects a target cell with probability $\beta T_0 / (c + \beta T_0)$. Provided T_0 is large, the mean number of cells infected by virions released from a single infected cell is

$$\bar{R} = \frac{p\tau_I\beta T_0}{c + \beta T_0}. \quad (2.2)$$

For \bar{R} as defined in (2.2), in the limit $T_0 \rightarrow +\infty$, we have $\bar{R} \rightarrow p\tau_I$, which is the mean number of virions released from an infected cell. However, for models that neglect the ODE term corresponding to loss of virions due to infection [2], the mean reproduction number is

$$\bar{R} = \frac{p\tau_I\beta T_0}{c},$$

which becomes infinite as $T_0 \rightarrow +\infty$. This implies that there is no finite bound on the mean number of secondary infections produced by an infected cell, so that as the initial target cell population becomes larger, the mean number of secondary infections will eventually become larger than the mean number of virions released from a single infected cell. This is not reasonable since each virion can only infect at most one cell.

2.1.2. Viral release by bursting

In the model (2.1), viral particles are released from infectious cells by budding. However, for some viruses, progeny virions accumulate inside an infected host cell and are released in a burst upon cell lysis. Yuan and Allen considered a model of viral dynamics with bursting and included a mechanism to represent the immune response [11]. They made the assumption that if the death of an infected cell is virus-induced, then virions will be released in a burst, but if the infected cell is killed due to the action of an immune response, no virions will be released. Yuan and Allen [11] also assumed that each burst event leads to the same burst size. That is, if an infected cell bursts, its burst size would always be $N = p\tau_I$ and would not depend on the time when the cell bursts. This might be the case if the accumulation of virions inside the cell and the burst of the cell are coupled in such a way that the burst of the cell takes place when the number of intracellular virions reaches a given threshold. By contrast, here we consider a case in which the accumulation of virions and the burst of the cell are independent processes. In this case, cells which burst earlier may not produce as many viral particles as those that survive longer. This introduces variability in the burst size, related to the times of the burst events. We assume that intracellular virions are produced at a constant rate, p , during the ‘infectious’ phase, but are not released unless the cell bursts. We note that under this assumption, the eclipse phase represents a period in which the cell begins to synthesize viral proteins and to replicate the viral genome, but progeny virions have not yet started to be assembled.

In order to describe these events with an ODE model, we require additional variables to represent the total number of intracellular virions in the population of infectious cells. We let the variable I_i denote the number of infected cells in stage i , and the variable P_i denote the total number of intracellular infectious virions in all infected cells of stage i . In a

given infected cell, intracellular virions are assumed to be produced at a constant rate, p . Each P_i increases due to virion production at a rate proportional to the number of cells in infectious stage i and due to cells entering stage i . P_i decreases when cells exit stage i . The rate at which a cell in stage i will exit this stage (either by transitioning to stage $i + 1$ or due to immune clearance) is $(\delta_i + v_I)I_i$. In the deterministic model, the P_i virions are equally shared between all cells of stage i . Therefore, the number of virions that will be cleared or will transition to the next stage when such an event happens will be P_i/I_i . Thus, the overall rate at which virions transition from compartment P_i to P_{i+1} (or get released in a burst for $i = n_I$) is $\delta_i I_i \frac{P_i}{I_i} = \delta_i P_i$. The system of $2 + n_E + 2n_I$ ODEs is

$$\begin{aligned} \frac{dT}{dt} &= -\beta_c TV, \\ \frac{dE_1}{dt} &= \beta_c TV - (\delta_E + v_E) E_1, \\ \frac{dE_i}{dt} &= \delta_E (E_{i-1} - E_i) - v_E E_i, \quad i = 2, \dots, n_E, \\ \frac{dI_1}{dt} &= \delta_E E_{n_E} - (\delta_I + v_I) I_1, \\ \frac{dI_i}{dt} &= \delta_I (I_{i-1} - I_i) - v_I I_i, \quad i = 2, \dots, n_I, \\ \frac{dP_1}{dt} &= pI_1 - (\delta_I + v_I) P_1, \\ \frac{dP_i}{dt} &= pI_i + \delta_I (P_{i-1} - P_i) - v_I P_i, \quad i = 2, \dots, n_I, \\ \frac{dV}{dt} &= \delta_I P_{n_I} - cV - \beta_v TV. \end{aligned} \quad (2.3)$$

2.2. Stochastic models

We will consider stochastic versions of the models in (2.1) and (2.3), to then compute the associated probability distributions of the burst size and reproduction number.

2.2.1. Viral release by budding

The model in (2.1) can be formulated as a continuous-time Markov chain (CTMC), $\mathcal{X} = \{X(t) : t \in [0, +\infty)\}$, where

$$X(t) = (V(t), T(t), E_1(t), \dots, E_{n_E}(t), I_1(t), \dots, I_{n_I}(t)).$$

$V(t)$, $T(t)$, $E_i(t)$ and $I_j(t)$ (for $i \in \{1, \dots, n_E\}$ and $j \in \{1, \dots, n_I\}$) are discrete random variables with values in the set of non-negative integers, for $t \in [0, +\infty)$. The possible transitions allowed in the stochastic model are shown in table 1.

2.2.2. Viral release by bursting

The model in (2.3) can be considered as a Markov process in which each infected cell is independent, with its own number of intracellular virions, which increases at a constant rate until the cell dies. If the cell reaches stage n_I and the death of the cell is virus-induced (with rate δ_I), then the intracellular virions at the time of death will be released in a burst into the extracellular environment. On the other hand, if the infected cell is killed by the immune system (with rate v_E during the eclipse phase, or rate v_I during the infectious phase), it is assumed that no infectious virions will be released [11]. This can represent granzymes delivered into the infected cell which can damage the intracellular virions [23,24].

Table 1. Transitions and their corresponding rates in the Markov chain version of the model in (2.1). We abuse notation and denote by $(V(t), T(t), E_1(t), \dots, E_{n_E}(t), I_1(t), \dots, I_{n_I}(t))$ the random variables of the stochastic process, while $(V, T, E_1, \dots, E_{n_E}, I_1, \dots, I_{n_I})$ represents a state of the process at any given time. Only the variables that change in a given transition are included in the middle column.

event	state transition	rate
infection	$(V, T, E_1) \rightarrow (V-1, T+1, E_1+1)$	βTV
eclipse stage progression	$(E_i, E_{i+1}) \rightarrow (E_i-1, E_{i+1}+1)$	$\delta_E E_i$
immune system killing	$E_i \rightarrow E_i-1$	$v_E E_i$
eclipse to infectious phase	$(E_{n_E}, I_1) \rightarrow (E_{n_E}-1, I_1+1)$	$\delta_E E_{n_E}$
infectious stage progression	$(I_i, I_{i+1}) \rightarrow (I_i-1, I_{i+1}+1)$	$\delta_I I_i$
immune system killing	$I_i \rightarrow I_i-1$	$v_I I_i$
death of infectious cell	$I_{n_I} \rightarrow I_{n_I}-1$	$(\delta_I + v_I) I_{n_I}$
viral release by budding	$V \rightarrow V+1$	$p \sum_{i=1}^{n_I} I_i$
loss of extracellular virions	$V \rightarrow V-1$	cV

2.3. Average burst size and reproduction number

2.3.1. Viral release by budding

Here we compute the mean burst size and reproduction number for the stochastic model of viral release by budding. The burst size is defined as the total number of virions released by an infected cell during its lifetime. We note that a cell becomes infected at the moment of viral entry, so that a cell in either the eclipse or infectious phase will be an *infected* cell. There is no viral release during the eclipse phase and the rate of viral release during the infectious phase is p . Thus, the mean burst size is given by

$$0 \times \mathbb{E}[t_E] + p\mathbb{E}[t_I] = p\mathbb{E}[t_I],$$

where t_E and t_I are the random variables for the length of time that an infected cell spends in the eclipse and infectious phases, respectively (i.e. t_I is the infectious period). When $v_E = v_I = 0$, we have $\mathbb{E}[t_I] = \tau_I$.

We will now consider the case when the killing of infected cells due to the immune response is included ($v_E > 0, v_I > 0$). In this case, it is possible for cells to be killed by the immune system, either in the eclipse phase (before becoming infectious), or in the infectious phase before passing through all n_I stages. If a cell is killed during the eclipse phase, then it will never become infectious (i.e. $t_I = 0$) and therefore the burst size will be zero. Let K be the random variable representing the number of infectious stages that an infected cell passes through before it dies; that is, the state space of K is $k \in \{0, 1, \dots, n_I\}$, where $k = 0$ implies the cell dies in the eclipse phase, and $k \in \{1, \dots, n_I\}$ implies the cell dies in the infectious stage k . At each infectious stage, the rate of death is v_I , and the rate of progression to the next stage is $\delta_I = n_I/\tau_I$. Thus, the time taken to exit each stage will be the minimum of two competing exponential random variables with rates v_I and δ_I , respectively. This minimum is an exponential random variable with rate $\delta_I + v_I$, and is independent of whether the cell is killed by the immune response or progresses to the next stage. That is, the time spent in a given infectious stage will always be exponentially distributed with rate $\delta_I + v_I$, even if we condition on progression to the next stage. Thus, the infectious period, t_I , is equivalent to the sum of K exponential random variables, each with rate $\delta_I + v_I$, where K is itself a random variable. Therefore, if $K \in \{1, \dots, n_I\}$, t_I follows an Erlang ($K, \delta_I + v_I$) distribution. If $K = 0$, then we have $t_I = 0$. The overall distribution

of t_I will be a weighted sum of these distributions, with the weights given by the distribution of K , so that the infectious period, t_I , has the following probability density:

$$\frac{d}{dt} \mathbb{P}(t_I \leq t) = \sum_{k=1}^{n_I} \mathbb{P}(K = k) \frac{(\delta_I + v_I)^k t^{k-1} e^{-(\delta_I + v_I)t}}{(k-1)!}.$$

The probabilities $\mathbb{P}(K = k)$ giving the distribution of K can be found as follows. For a cell in any given eclipse stage, the exit rate of the cell out of this stage is $\delta_E + v_E$, and when the cell exits this stage it will either die with probability $v_E/(\delta_E + v_E)$, or move to the next stage with probability $\delta_E/(\delta_E + v_E)$. Hence the probability that the cell will survive all n_E eclipse stages and progress to the infectious phase is $(\delta_E/(\delta_E + v_E))^{n_E}$. Similarly, for a cell in infectious stage $i \in \{1, \dots, n_I - 1\}$, the exit rate of the cell out of this stage is $\delta_I + v_I$, and when the cell exits this stage it will either die with probability $v_I/(\delta_I + v_I)$, or move to the next infectious stage with probability $\delta_I/(\delta_I + v_I)$. Note that for the final infectious stage, n_I , the cell will always die when exiting this stage. Therefore, the probability that an infected cell passes through k infectious stages before it dies is given by

$$\mathbb{P}(K = k) = \begin{cases} 1 - r_E^{n_E}, & k = 0, \\ r_E^{n_E} r_I^{k-1} (1 - r_I), & k = 1, \dots, n_I - 1, \\ r_E^{n_E} r_I^{n_I-1}, & k = n_I, \end{cases} \quad (2.4)$$

where $r_E = \delta_E/(\delta_E + v_E)$, $r_I = \delta_I/(\delta_I + v_I)$. We note that when $v_E = v_I = 0$, one has $r_E = r_I = 1$, so that $K = n_I$ with probability one. The mean infectious period can then be calculated as

$$\begin{aligned} \mathbb{E}[t_I] &= \sum_{k=0}^{n_I} \mathbb{P}(K = k) \frac{k}{\delta_I + v_I} \\ &= \frac{r_E^{n_E}}{v_I} \left((1 - r_I)^2 \sum_{k=1}^{n_I-1} [k r_I^{k-1}] + n_I r_I^{n_I-1} (1 - r_I) \right) \\ &= \frac{1}{v_I} r_E^{n_E} (1 - r_I^{n_I}), \quad \text{for } v_I > 0. \end{aligned}$$

To obtain the mean reproduction number, we multiply the mean burst size, $p\mathbb{E}[t_I]$, by the mean fraction of these virions that will go on to infect new target cells,

$$\bar{R} = p\mathbb{E}[t_I] \frac{\beta T_0}{c + \beta T_0} = r_E^{n_E} (1 - r_I^{n_I}) \frac{p\beta T_0}{v_I(c + \beta T_0)}. \quad (2.5)$$

In the limit $v_E \rightarrow 0$ and $v_I \rightarrow 0$, we have $\mathbb{E}[t_I] \rightarrow \tau_I$ and we regain (2.2).

2.3.2. Viral release by bursting

For the model of viral release by bursting, the rate of intracellular virion production during the infectious phase is p , but virions will only be released in a burst if the cell is not killed by the immune system in any of the eclipse or infectious stages. The probability for a given infected cell to burst is $r_E^{n_E} r_I^{n_I}$, and given that the cell bursts, its mean infectious period will be $n_I / (\delta_I + v_I)$. Thus, the mean burst size is

$$r_E^{n_E} r_I^{n_I} \frac{pn_I}{\delta_I + v_I},$$

and the mean reproduction number is

$$\bar{R} = r_E^{n_E} r_I^{n_I+1} \frac{p\tau_I\beta T_0}{c + \beta T_0}. \quad (2.6)$$

In the limit $v_E \rightarrow 0$ and $v_I \rightarrow 0$, we have $r_E = r_I = 1$, and we regain (2.2).

2.4. Burst size probability distribution

2.4.1. Viral release by budding

Let B be the random variable representing the viral burst size, i.e. the total number of virions released by an infected cell during its lifetime. To find the distribution of B , we consider the dynamics of viral release from a single infected cell. As explained in the previous section, an infected cell will progress through a (random) number of infectious stages, $K \in \{0, 1, \dots, n_I\}$, before it dies. The probability distribution of K is given in (2.4). We will first consider the burst size distribution conditioned on the value of K , and then compute a weighted sum over the possible values of K . If $K = 0$, then the cell dies before entering the infectious phase and the burst size is zero. If $K = k$ with $k > 0$, then there are $k - 1$ progression events (from one infectious stage to the next) and b virion release events before the cell dies in the k th infectious stage. During the infectious phase, a cell is assumed to release virions at a constant rate, p , with an exponentially distributed time between the release of each new virion. Hence, each virion release event has probability $p / (p + \delta_I + v_I)$. We thus conclude that the total number of virions released by the cell before it dies is a negative binomial random variable describing the number of ‘successes’ before the k th ‘failure’, where a ‘success’ is the release of a virion and a ‘failure’ is leaving the stage. Therefore, given that the cell progresses through exactly $K = k > 0$ infectious stages, the burst size B follows a negative binomial distribution, with shape parameter k and success probability $p / (p + \delta_I + v_I)$. Once we sum over the possible values of K , the probability mass function (p.m.f.) for the number of virions released from an infected cell is

$$\begin{aligned} \mathbb{P}(B = b) &= \sum_{k=0}^{n_I} \mathbb{P}(B = b \mid K = k) \mathbb{P}(K = k) \\ &= (1 - r_E^{n_E})\delta_{b,0} + r_E^{n_E} \left(\sum_{k=1}^{n_I} f(b, k) \frac{\delta_I^{k-1} v_I}{(p + \delta_I + v_I)^k} \right. \\ &\quad \left. + f(b, n_I) \frac{\delta_I^{n_I-1} (\delta_I + v_I)}{(p + \delta_I + v_I)^{n_I}} \right), \end{aligned} \quad (2.7)$$

for $b \in \{0, 1, 2, \dots\}$, where $\delta_{b,0}$ represents the Kronecker delta, and

$$f(b, k) = \frac{(k + b - 1)!}{(k - 1)! b!} \left(\frac{p}{p + \delta_I + v_I} \right)^b. \quad (2.8)$$

If $v_E = v_I = 0$, the burst size distribution becomes a negative binomial with shape parameter n_I and success probability $p / (p + \delta_I)$, so that we can write

$$\mathbb{P}(B = b) = \binom{n_I + b - 1}{b} \left(\frac{p}{p + \delta_I} \right)^b \left(\frac{\delta_I}{p + \delta_I} \right)^{n_I}, \quad (2.9)$$

for $b \in \{0, 1, 2, \dots\}$. The expectation and variance of this negative binomial distribution are given by

$$\mathbb{E}[B] = \frac{pn_I}{\delta_I} = p\tau_I,$$

and

$$\text{Var}[B] = \mathbb{E}[B] \left(1 + \frac{\mathbb{E}[B]}{n_I} \right) = \frac{pn_I}{\delta_I} + \frac{p^2 n_I}{\delta_I^2}.$$

We note that if the infectious period is assumed to be exponentially distributed ($n_I = 1$), then the distribution in (2.9) becomes geometric.

2.4.2. Viral release by bursting

When $v_I = 0$, the probability distribution of the burst size is equal for the two models of budding and bursting, since the number of virions produced by an infected cell has the same distribution in both instances, but they are either released from the cell gradually in the case of budding, or remain in the cell until it dies in the case of bursting. However, when infectious cells can be killed by the immune system, the two strategies produce different burst size distributions. In the budding strategy, some virions may be released from an infectious cell before the cell is killed by the immune system, but in the bursting strategy, if an infectious cell is killed by the immune system before it bursts, then all intracellular virions are assumed to be eliminated in the process and the corresponding burst size is zero. Hence, in the case of bursting, virions will only be released from the infected cell if the cell is not killed by the immune system, and it exits the final infectious stage by bursting. Given that a burst does occur, the burst size will be negative binomially distributed with shape parameter n_I and success probability $p / (p + \delta_I + v_I)$. Thus, the burst size probability distribution is given by

$$\begin{aligned} \mathbb{P}(B = b) &= (1 - r_E^{n_E} r_I^{n_I})\delta_{b,0} \\ &\quad + r_E^{n_E} f(b, n_I) \left(\frac{\delta_I}{p + \delta_I + v_I} \right)^{n_I}, \end{aligned} \quad (2.10)$$

for $b \in \{0, 1, 2, \dots\}$, where $\delta_{b,0}$ represents the Kronecker delta, and $f(b, k)$ is defined in (2.8).

The burst size p.m.f.s and means for the budding and bursting models are summarized in table 2.

2.5. Reproduction number probability distribution

We now return to R , the random variable representing the cell-level reproduction number; that is, the number of secondary infections produced by a single infected cell in an otherwise susceptible cell population. Our aim in this

Table 2. Summary of the burst size and reproduction number distributions and means for the different model assumptions of viral release by budding and bursting, defined for $b, r \in \{0, 1, 2, \dots\}$. $\delta_{b,0}$ represents the Kronecker delta and we define $r_E = \delta_E/(\delta_E + v_E)$, $r_I = \delta_I/(\delta_I + v_I)$. $f(b, k)$ is defined in (2.8). Note that these reproduction number distributions were calculated for the case of a constant number of uninfected target cells (i.e. Case 1), where each virion has probability θ to infect a new cell. For Case 2, in which the depletion of uninfected target cells is considered, the approach in (2.12) should be followed, using the recursively defined probabilities, $p_r[b]$.

burst size distribution	
budding	$\mathbb{P}(B = b) = (1 - r_E^{n_E})\delta_{b,0} + r_E^{n_E} \left(\sum_{k=1}^{n_I-1} f(b, k) \frac{\delta_I^{k-1} v_I}{(p + \delta_I + v_I)^k} + f(b, n_I) \frac{\delta_I^{n_I-1} (\delta_I + v_I)}{(p + \delta_I + v_I)^{n_I}} \right)$
bursting	$\mathbb{P}(B = b) = (1 - r_E^{n_E} r_I^{n_I})\delta_{b,0} + r_E^{n_E} r_I^{n_I} f(b, n_I) \left(\frac{\delta_I}{p + \delta_I + v_I} \right)^{n_I}$
mean burst size	
budding	$\bar{B} = r_E^{n_E} (1 - r_I^{n_I}) \frac{p}{v_I}$
bursting	$\bar{B} = r_E^{n_E} r_I^{n_I} \frac{p n_I}{\delta_I + v_I}$
reproduction number distribution	
budding	$\mathbb{P}(R = r) = (1 - r_E^{n_E})\delta_{r,0} + r_E^{n_E} \left(\sum_{k=1}^{n_I-1} \binom{k+r-1}{r} \frac{(\theta p)^r \delta_I^{k-1} v_I}{(\theta p + \delta_I + v_I)^{r+k}} + \binom{n_I+r-1}{r} \frac{(\theta p)^r \delta_I^{n_I-1} (\delta_I + v_I)}{(\theta p + \delta_I + v_I)^{r+n_I}} \right)$
bursting	$\mathbb{P}(R = r) = (1 - r_E^{n_E} r_I^{n_I})\delta_{r,0} + r_E^{n_E} r_I^{n_I} \binom{n_I+r-1}{r} \frac{(\theta p)^r \delta_I^{n_I}}{(\theta p + \delta_I + v_I)^{r+n_I}}$
mean reproduction number	
budding	$\bar{R} = r_E^{n_E} (1 - r_I^{n_I}) \frac{p \theta}{v_I}$
bursting	$\bar{R} = r_E^{n_E} r_I^{n_I+1} p \tau_I \theta$

section is to find the p.m.f. of R , which may be written in terms of the burst size probabilities as

$$\mathbb{P}(R = r) = \sum_{b=0}^{+\infty} \mathbb{P}(R = r | B = b) \mathbb{P}(B = b),$$

with r a non-negative integer. Hence, we will find the probabilities $\mathbb{P}(R = r | B = b)$ (number of secondary infections produced, given that the initial infected cell releases $B = b$ virions during its lifetime), so that together with the probability distribution of B (2.7) for budding and (2.10) for bursting), we can compute the probability distribution of R .

The only possible outcomes for a released virion are (i) that it is lost or (ii) that it infects a target cell. Free virions are cleared with rate c , and infect new cells with rate βT , which depends on the number of available target cells, T . Usually, calculations of the reproduction number assume that the number of infectious viral particles is small enough so that the number of (uninfected) target cells is not significantly altered by infection events. Thus, the number of target cells is usually assumed constant [25], as in §2.5.1. However, if the number of target cells at the start of infection is small and the infection rate, β , is large, it will be important to consider the transition of target cells to infected cells, since the depletion of uninfected target cells will affect the ability of extracellular viral particles (released from infected cells) to infect new cells. The number of secondary infections caused by a single infected cell will depend on: (i) competition between virions produced by the initial infected cell for available target cells and (ii) competition for target cells from virions being produced during secondary infections. Here, in §2.5.2, we will address the former source of competition, while ignoring the second; that is, we consider depletion of target cells due to secondary infections caused by viral particles released from the initially infected cell, but do not

consider depletion of target cells due to viral particles released from subsequently infected cells.

2.5.1. Target cell number is constant (Case 1)

If the number of uninfected target cells is sufficiently large, their loss due to infection can be neglected, and their number assumed to be constant. In this case, extracellular virions are independent of each other, in the sense that they will have the same probability of successfully infecting a target cell, independently of the fate of other such virions. For each infectious virion released from an initially infected cell, the probability that it will infect a target cell is

$$\theta = \frac{\beta T_0}{c + \beta T_0},$$

where β is the rate at which extracellular virions infect target cells, T_0 is the constant number of target cells, and c is the rate of loss of extracellular virions. Therefore, the number of secondary infections will be the sum of b independent and identical Bernoulli random variables, and $R | B = b$, follows a binomial distribution with parameters (b, θ) . We then have, for r a non-negative integer,

$$\mathbb{P}(R = r) = \sum_{b=r}^{+\infty} \binom{b}{r} \theta^r (1 - \theta)^{b-r} \mathbb{P}(B = b).$$

Substituting in the p.m.f. of the burst size distribution from (2.7) (budding model) or (2.10) (bursting model), we find that the p.m.f. of the reproduction number has the same form as the burst size p.m.f., but with the parameter p replaced by θp , which is equivalent to assuming that secondary infections are produced at rate θp . This makes sense since each virion released from an infected cell is expected to infect a new cell with probability θ , and when considering only the number of secondary infections produced, we ignore the timescale of

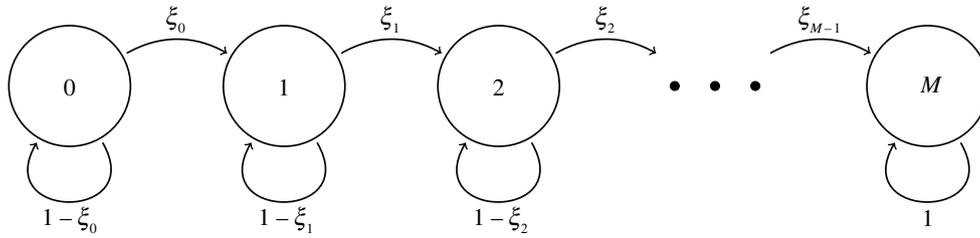


Figure 2. Counting secondary infections using a Markov chain. At each of b steps, one for each released virion, the process either moves one state to the right, corresponding to infection of a target cell, or stays put, corresponding to loss of a virion. $\xi_i = \beta(T_0 - i)/(\beta(T_0 - i) + c)$ is the probability of moving from state i to $i + 1$, where T_0 is the initial number of uninfected target cells available. The position after b steps is the number of secondary infections generated by b extracellular virions.

these events. The reproduction number p.m.f.s and means for the budding and bursting models are provided in table 2.

If $v_E = v_I = 0$, since the burst size follows a negative binomial distribution (see (2.9)), then the number of secondary infections also follows a negative binomial distribution, with shape parameter n_I and success probability $\theta p/(\theta p + \delta_I)$. The p.m.f. of the reproduction number is

$$\mathbb{P}(R = r) = \binom{n_I + r - 1}{r} \left(\frac{\theta p}{\theta p + \delta_I} \right)^r \left(\frac{\delta_I}{\theta p + \delta_I} \right)^{n_I}. \quad (2.11)$$

2.5.2. Target cell number is limiting (Case 2)

The probability of a given virion successfully infecting a cell before it is lost depends on the number of available target cells. Therefore, if the number of target cells is small, it is essential to take into account their reduction as they become infected. In this case, the virions cannot be treated as independent, but their probability of infecting a cell will depend on how many target cells are still available after previous infection events.

Let us assume that there are T_0 target cells in the population to begin with, and $B = b$ infectious virions have been released by one infected cell during its infectious phase. Since we are only interested in the number of secondary infections generated by a single infected cell, and not the times of those infection events, one can assume that at discrete time steps, a virion is chosen uniformly at random out of the $B = b$ virions, to either be lost or infect a target cell. In this way, we can then set up a Markov chain to calculate the probability distribution of the number of secondary infections generated by $B = b$ extracellular virions produced by a single infected cell.

Let us define the discrete-time Markov chain, $\mathcal{Y} = \{Y_n : n \in \{0, 1, 2, \dots, b\}\}$, where Y_n denotes the number of secondary infections that have occurred after n of the infectious extracellular virions have either been lost or infected an available target cell. The initial state of the process will be $Y_0 = 0$ and the state space of \mathcal{Y} will be given by $S_{\mathcal{Y}} = \{0, 1, 2, \dots, M\}$, where $M = \min\{b, T_0\}$ is the maximum number of secondary infections that can take place. The Markov chain \mathcal{Y} , shown in figure 2, can be defined by the following one-step transition probabilities, for states $i, j \in S_{\mathcal{Y}}$. For $i < M$,

$$p_{i,j} = \mathbb{P}(Y_{n+1} = j | Y_n = i) = \begin{cases} \xi_i, & \text{if } j = i + 1, \\ 1 - \xi_i, & \text{if } j = i, \\ 0, & \text{otherwise,} \end{cases}$$

where $\xi_i = \beta(T_0 - i)/(\beta(T_0 - i) + c)$. If the number of secondary infections reaches M , we have

$$p_{M,j} = \begin{cases} 1, & \text{if } j = M, \\ 0, & \text{otherwise.} \end{cases}$$

The probability that r secondary infections will occur (with r a non-negative integer), given that the cell released b infectious virions, is given by $\mathbb{P}(R = r | B = b) = \mathbb{P}(Y_b = r | Y_0 = 0)$, which is the probability that the process \mathcal{Y} is in state r after the b infectious extracellular virions have either been lost or infected target cells. For ease of notation, we write $p_r[b] = \mathbb{P}(Y_b = r | Y_0 = 0)$. Making use of first-step arguments, these probabilities can be calculated recursively, as follows. We have $p_0[0] = 1$. Then for $b \geq 1$,

$$p_r[b] = \begin{cases} p_0[b-1]p_{0,0}, & \text{if } r = 0, \\ p_{r-1}[b-1]p_{r-1,r} + p_r[b-1]p_{r,r}, & \text{if } 0 < r \leq M. \end{cases}$$

Thus, substituting the one-step transition probabilities from above, we have

$$p_r[b] = \begin{cases} (1 - \xi_0)^b, & \text{if } r = 0, \\ p_{r-1}[b-1]\xi_{r-1} + p_r[b-1](1 - \xi_r), & \text{if } 0 < r \leq \min(b, T_0), \\ 0, & \text{if } r > \min(b, T_0). \end{cases}$$

We can solve the previous system sequentially to obtain $p_r[b]$ for all required values of r and b . We get

$$\mathbb{P}(R = r) = \sum_{b=0}^{+\infty} p_r[b] \mathbb{P}(B = b). \quad (2.12)$$

If the initial number of target cells, T_0 , is large enough so that only a small fraction is going to become infected by the virions produced by a single infected cell, then it is reasonable to approximate $\xi_i = \theta$ for all i . This approximation leads to Case 1, where

$$p_r[b] = \binom{b}{r} \theta^r (1 - \theta)^{b-r}, \quad 0 \leq r \leq b.$$

2.6. Probability of viral extinction

It is useful to also calculate the probability of viral extinction during the early stages of infection. When the number of target cells is assumed constant, the initial dynamics of the models considered can be described by branching processes, in which each infected cell produces a number of new infected cells according to the reproduction number distribution. In such a branching process, the probability of extinction starting with one infected cell is equal to the

smallest fixed point of the probability generating function (p.g.f.) of the offspring distribution (see ch. 4 of Allen [26]). If the branching process is counting the number of infected cells at each generation, then the offspring distribution will be the reproduction number distribution.

In the case of budding, the p.g.f. of the reproduction number (when the number of target cells is constant) is given by

$$\pi_{\text{bud}}(s) = 1 - r_E^{n_I} \left[1 - \sum_{k=1}^{n_I} \frac{\tau_I v_I n_I^{k-1}}{(n_I + \tau_I v_I + (1-s)\tau_I \theta p)^k} - \left(\frac{n_I}{n_I + \tau_I v_I + (1-s)\tau_I \theta p} \right)^{n_I} \right]. \quad (2.13)$$

Further details to show how this p.g.f. was obtained can be found in Williams [27, p. 170].

In the case of bursting, the p.g.f. of the reproduction number (when the number of target cells is constant) is given by

$$\pi_{\text{burst}}(s) = 1 - r_E^{n_I} r_I^{n_I} \left[1 - \left(\frac{n_I + \tau_I v_I}{n_I + \tau_I v_I + (1-s)\tau_I \theta p} \right)^{n_I} \right]. \quad (2.14)$$

We note that when $v_I = 0$, π_{bud} and π_{burst} are identical, since the reproduction number distribution is the same for the budding and bursting cases.

When $n_I = 1$, corresponding to an exponentially distributed infectious period, the fixed points of the functions π_{bud} and π_{burst} can be found explicitly. For budding and $n_I = 1$, the solutions of $\pi_{\text{bud}}(s) = s$ are $s_1^{\text{bud}} = 1$ and

$$s_2^{\text{bud}} = 1 - r_E^{n_I} \left(1 - \frac{1}{\bar{R}_{\text{bud}}} \right),$$

where $\bar{R}_{\text{bud}} = r_E^{n_I} (\theta p / \delta_I + v_I)$ is the mean reproduction number in the budding case ($n_I = 1$). Similarly, for bursting and $n_I = 1$, the solutions to $\pi_{\text{burst}}(s) = s$ are $s_1^{\text{burst}} = 1$ and

$$s_2^{\text{burst}} = 1 - r_E^{n_I} r_I \left(1 - \frac{1}{\bar{R}_{\text{burst}}} \right),$$

where $\bar{R}_{\text{burst}} = r_E^{n_I} r_I (\theta p / \delta_I + v_I) = r_I \bar{R}_{\text{bud}}$. The smallest fixed point of the p.g.f. gives the probability of viral extinction starting with one infected cell, so that in each case extinction will be certain if $s_2 \geq 1$ (corresponding to $\bar{R} \leq 1$) and will occur with probability s_2 if $s_2 < 1$ (corresponding to $\bar{R} > 1$). Since $\bar{R}_{\text{burst}} \leq \bar{R}_{\text{bud}}$ and $s_2^{\text{burst}} \geq s_2^{\text{bud}}$, extinction will always be equally or more likely to occur in the bursting case than the budding one. In particular, when $1 < \bar{R}_{\text{bud}} \leq 1/r_I$, extinction is certain in the bursting case, but not in the budding one. For $n_I > 1$, the smallest fixed points of the p.g.f.s can be found numerically, giving the probability of viral extinction starting with one infected cell. Since each initial infected cell is independent, the probability of extinction starting with one infected cell can easily be used to find the probability of extinction starting with multiple infected cells [26].

3. Numerical results

In this section, we present the sensitivity of the reproduction number distribution and probability of viral extinction computed in §2 to different models of viral release and parameter values. We ignore any immune killing of eclipse-phase cells, so that $v_E = 0$ and every infected cell is assumed to survive the eclipse phase and enter the infectious phase.

3.1. Infection rate

We first consider the model without immune killing of infectious cells, corresponding to $v_I = 0$. In this case, the probability distribution of the number of virions released by an infected cell during its lifetime is the same for either model of viral release (budding or bursting). For both models, the burst size follows a negative binomial distribution with mean $p\tau_I$ and shape parameter n_I (see (2.9)). Once virions are released from an infected cell, the rate at which they infect new cells depends on β and on the number of target cells available to be infected, T . Section 2.5 described two cases to consider when calculating the distribution of the reproduction number, R . When the number of target cells is not limiting, since the population is very large or is constantly replenished (Case 1), the reproduction number follows the negative binomial distribution in (2.11) (for $v_E = v_I = 0$). The shape parameter of this distribution is n_I , and the mean number of secondary infections produced is $\bar{R} = \theta p \tau_I$, where $\theta = \beta T_0 / (c + \beta T_0)$ is the probability for a given released virion to infect a target cell. Alternatively, if the number of target cells is limiting (Case 2), the reproduction number distribution can be calculated using (2.12), in which the depletion of target cells due to infection is taken into account.

When the value of \bar{R} is much smaller than T_0 , and only a small fraction of the target cells are likely to become infected, the probability distributions of R for Case 1 and Case 2 are similar. The smaller the difference between the number of target cells and the expected number to become infected (i.e. $T_0 - \bar{R}$), the greater the difference of the distributions obtained from the two methods become.

Figure 3 shows the different probability distributions of R calculated with Case 1 and Case 2 methods, for some key values of the infection rate, β , and number of target cells, T_0 . All other parameters are fixed to the values in table 3, with $v_I = 0$. The Case 1 method assumes that there is a constant population of T_0 target cells which does not decrease as cells become infected. On the other hand, Case 2 assumes that there is a population of T_0 target cells to begin with, but each secondary infection reduces this number by one, so that it is not possible for more than T_0 secondary infections to occur. The two plots in the top row of figure 3 have parameter values of β and T_0 , such that $\theta = 0.014$ and the mean reproduction number is $\bar{R} = 17.6$ (for Case 1, from (2.2)). However for the plot on the top left, $T_0 = 10 < \bar{R}$, making the distribution of R very different, depending on whether it is calculated using the method for Case 1 or Case 2. When $T_0 = 10^3 \gg \bar{R}$, the two distributions become very similar. For the plots on the bottom row of figure 3, the values of β and T_0 give $\theta = 0.59$ and $\bar{R} = 735$. For the plot on the bottom left, where $T_0 = 10^3$ is only slightly larger than \bar{R} , there is a clear difference between the two distributions. When T_0 is increased to 10^5 in the plot on the bottom right, this difference becomes much smaller.

The heatmap in figure 4 illustrates the difference between the two distributions of Case 1 and Case 2 across a wider range of values of β and T_0 . The difference between the distributions is quantified by the Hellinger distance, defined by

$$H(P, Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_i (\sqrt{p_i} - \sqrt{q_i})^2},$$

for two discrete probability distributions, $P = \{p_0, p_1, p_2, \dots\}$ and $Q = \{q_0, q_1, q_2, \dots\}$. $H(P, Q)$ gives a value between 0

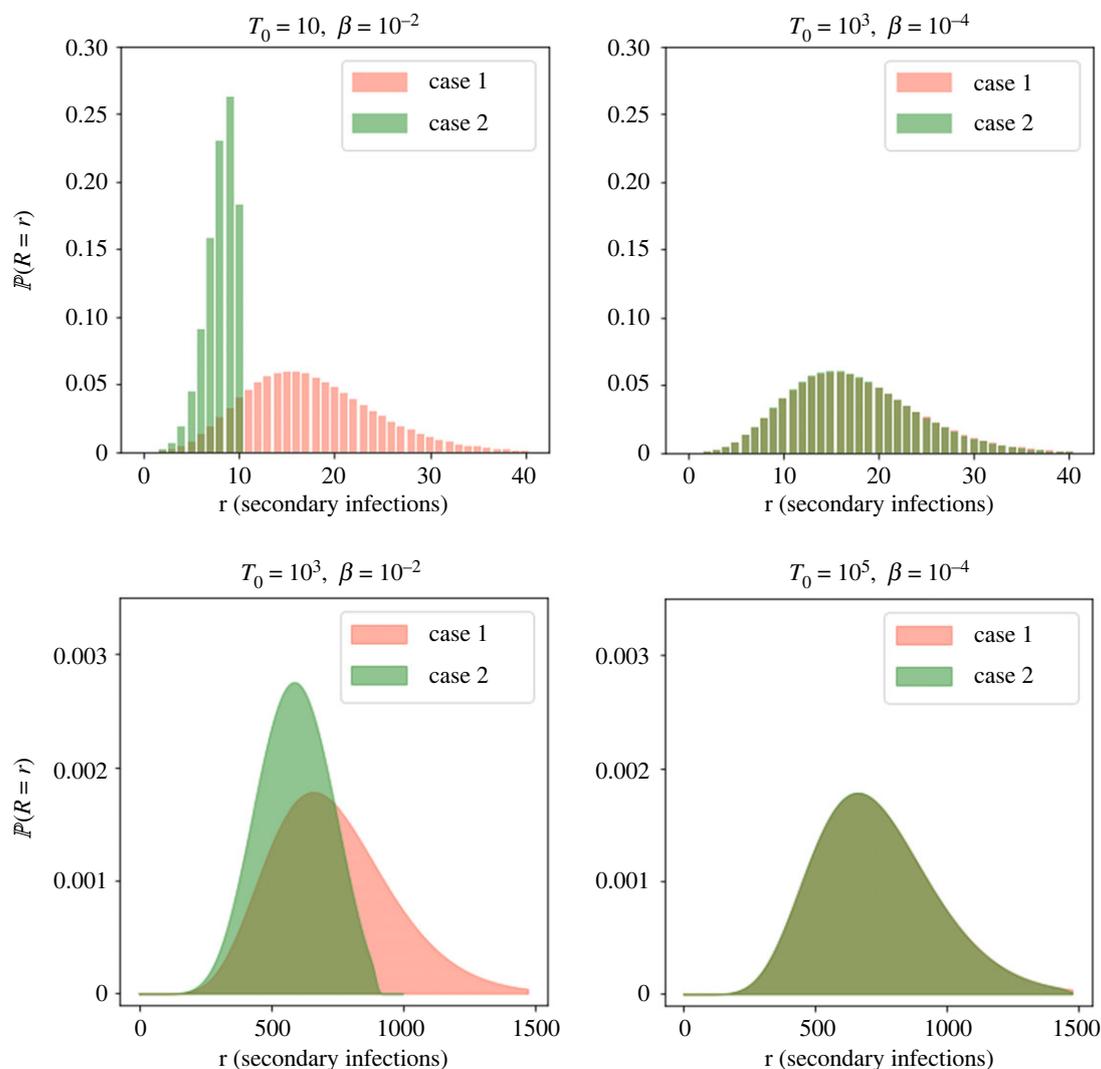


Figure 3. Histograms for the probability distribution of the reproduction number, R , for different values of β and T_0 . For each pair of parameter values, the distributions obtained from methods of Case 1 using (2.11) (number of target cells is constant) and Case 2 using (2.12) (number of target cells decreases as they become infected) are shown. All other parameter values are fixed to the values in table 3, with $v_I = 0$ (so the reproduction number distribution is equivalent under budding or bursting assumptions).

Table 3. Model parameters with their descriptions and units. We define $\delta_E = n_E/\tau_E$, $\delta_I = n_I/\tau_I$, and $\theta = \beta T_0/(c + \beta T_0)$. Values are given for some parameters, which have been used to obtain numerical results in §3. In all numerical results, we set $v_E = 0$, which means that the values of τ_E and n_E (determining the distribution of the eclipse phase) do not affect the burst size or reproduction number distribution.

parameter	description	units	value
p	viral release rate (budding) or intracellular production rate (bursting)	virions · (cell · day) ⁻¹	1000
τ_E	mean of Erlang-distributed eclipse phase in the absence of immune killing	days	n.a.
n_E	number of stages of the Erlang-distributed eclipse phase	—	n.a.
v_E	immune clearance rate of eclipse-phase cells	(cell · day) ⁻¹	0
τ_I	mean of Erlang-distributed infectious period in the absence of immune killing	days	1.25
n_I	number of stages of the Erlang-distributed infectious period	—	10
v_I	immune clearance rate of infectious cells	(cell · day) ⁻¹	varied
β	infection rate	(cell · virion · day) ⁻¹	10^{-4}
c	rate of loss of extracellular virions	(virion · day) ⁻¹	7
T_0	initial number of target cells	cells	10^5

and 1, where 0 indicates that the two distributions are identical. In general, for larger values of the infection rate, β , and smaller values of T_0 , the distributions become more

different, since these parameter values will mean that the virus is likely to infect a greater fraction of the available target cells.

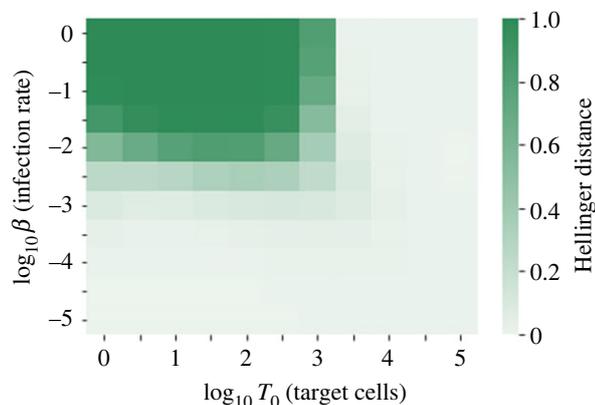


Figure 4. Heatmap showing the Hellinger distance between the two distributions of the reproduction number, R , calculated using Case 1 (2.11) and Case 2 (2.12), for different values of β and T_0 . All other parameter values are fixed to the values in table 3, with $v_I=0$ (so the reproduction number distribution is equivalent under budding or bursting assumptions).

For the value of $\beta = 10^{-4} (\text{cell} \cdot \text{virion} \cdot \text{day})^{-1}$ in table 3, the mean reproduction number will be much smaller than T_0 , for any $T_0 \in \mathbb{N}$. Therefore the distribution of R will be very similar using either method, and we will assume a constant number of target cells (Case 1) to calculate the reproduction number distributions in the following sections.

3.2. Immune killing

In vitro experiments have previously been used to obtain estimates for viral dynamics parameters. For example, Liao *et al.* [2] fitted a mathematical model similar to (2.1) with $v_E = v_I = 0$ to data of *in vitro* Ebola virus infection. In another recent study, Yan *et al.* [7] fitted a similar model to *in vitro* data and compared estimates of growth rate, reproduction number and generation time, for six influenza A strains. Studies like these are very useful for determining many key model parameters; however, they do not allow parameters that represent *in vivo* processes, such as immune system clearance of infected cells, to be estimated. Here we show that the reproduction number distribution is very sensitive to the rate at which the immune system is assumed to clear infected cells *in vivo*. Figures 5 and 6 illustrate the reproduction number distribution for different rates of immune killing of infectious cells, for the models of viral release by budding and bursting, respectively. In the model of budding, infectious cells are assumed to produce new infections at a constant rate, θp , until they die. An increased rate of immune clearance means that the infectious period distribution is changed, with more cells dying earlier. Therefore, it becomes more likely for low numbers of secondary infections to be produced.

In the bursting model, the reproduction number probability distribution is given by a zero-inflated negative binomial distribution; if an infected cell bursts and releases virions, then the number of secondary infections caused by the cell follows a negative binomial distribution with shape parameter n_I and probability $\theta p / (\theta p + \delta_I + v_I)$. The probability of zero secondary infections is inflated, since if the initial infected cell is killed by the immune system before it bursts, then it is assumed that zero virions will be released and zero secondary infections can be produced. Thus, for larger rates of immune clearance, the probability of zero

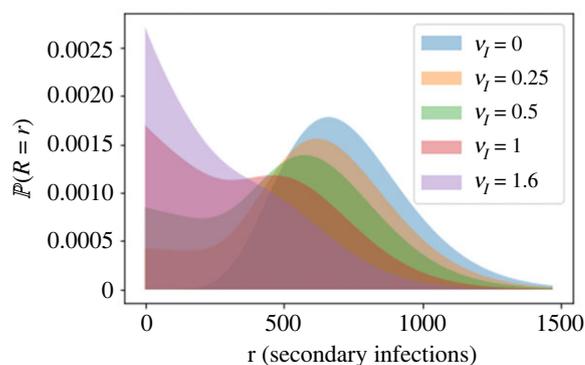


Figure 5. Reproduction number probability distributions calculated using the expression in table 2 for the model of viral release by budding, for different values of the immune killing rate of infectious cells, v_I . Other parameters are set to the values in table 3. The values of v_I used are 0, 0.25, 0.5, 1 and 1.6, per cell per day, corresponding to values of \bar{R} of 735, 623, 535, 407 and 308, respectively.

secondary infections increases. Given that a cell does burst and release virions, the distribution of time until the cell bursts is also affected by the rate of immune clearance. The reason for this is that the immune response introduces a competition between the two mechanisms of death for infected cells. This has the effect of reducing the mean time to cell burst, since cells that would have taken longer to burst are more likely to be cleared by the immune response before they do so. Hence the mean number of virions released from bursting cells and the mean number of secondary infections is also reduced.

These results for the reproduction number distribution can be translated into the probability of viral extinction, using the p.g.f.s presented in §2.6. Figure 7 shows the effect of immune killing on the probability of viral extinction for the bursting model. When $v_I = 0$, corresponding to an *in vitro* situation without the immune clearance of infected cells, the model predicts almost zero chance of viral extinction starting from one infected cell, for the parameters used here. As the immune clearance rate increases, viral extinction becomes much more likely. These results highlight how processes *in vivo*, such as the immune system clearance of infected cells, may alter the expected probability of extinction obtained from models of *in vitro* experiments. There are also many other *in vivo* biological mechanisms to consider, other than the rate at which the immune system clears infected cells, which are not studied here but may also impact the reproduction number distribution. Some of these are discussed in §4.

3.3. Number of infectious stages

We now study the impact of the number of infectious phase stages, given by the parameter n_I , on the reproduction number distribution and probability of viral extinction for the model with bursting. We note that the number of infectious stages, n_I , is varied, but with the mean of the Erlang-distributed time to cell burst, τ_I , kept constant. Thus, as n_I is varied, $\delta_I = n_I / \tau_I$ also changes, since it depends on n_I . Other parameters are set to their values in table 3 and we use an immune clearance rate of $v_I = 1.6$ per cell per day. This implies that the immune system killing of infectious cells occurs on average two times faster than virus-induced cell death, which is assumed to take $\tau_I = 1.25$ days on average.

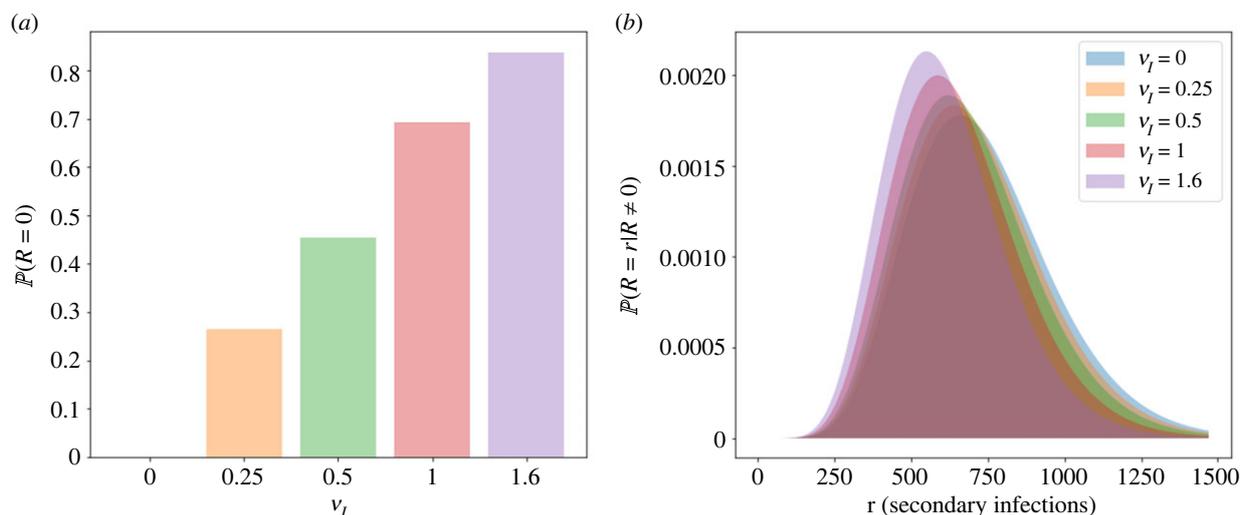


Figure 6. Reproduction number probability distributions calculated using the expression in table 2 for the model of viral release by bursting, for different values of the immune killing rate of infectious cells, v_I . Other parameters are set to the values in table 3. The values of v_I used are 0, 0.25, 0.5, 1 and 1.6, per cell per day, corresponding to values of \bar{R} of 735, 524, 377, 201 and 99, respectively. (a) Probability of zero secondary infections for each value of v_I . (b) Probability distribution of R conditioned on positive values of the reproduction number.

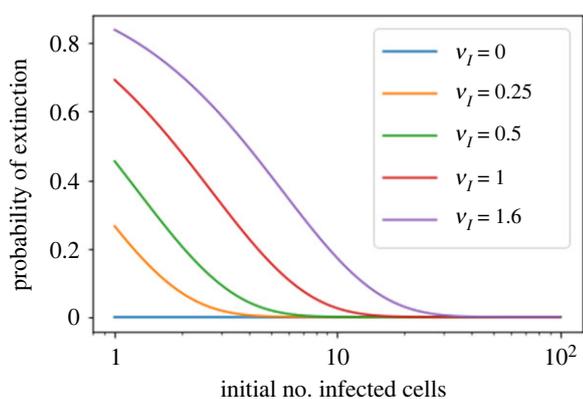


Figure 7. Probability of viral extinction as a function of the initial number of infected cells, for the model of viral release by bursting, for different values of the immune killing rate of infectious cells, v_I . The probability of extinction starting with one cell is calculated by numerically finding the smallest fixed point of the p.g.f. in (2.14). To obtain the probability of extinction starting with i infected cells, one takes the probability for one cell to the power of i . Other parameters are set to the values in table 3.

Figure 8 shows the bursting model reproduction number distribution for different values of n_I . As the number of infectious stages increases, the variance of the Erlang burst time distribution decreases and the distribution becomes more tightly centred around the mean. On the other hand, the exponential killing time distribution is not changed. Since the exponential distribution allows times very close to zero, the killing time is likely to be smaller than the burst time, leading to a higher chance that killing by the immune system will occur first. Therefore, there is a higher probability of zero secondary infections for larger values of n_I (figure 8a). However, if the infected cell does burst, it will release a larger number of virions on average when the value of n_I is higher. Therefore, when conditioned on positive numbers of secondary infections, the reproduction number distribution moves to the right with increasing n_I (figure 8b). For the parameter values in table 3 and the values of n_I considered, the overall

mean reproduction number, given in (2.6), increases with the value of n_I .

Figure 9 presents some interesting results about the probability of viral extinction, which is found by numerically calculating the smallest fixed point of the p.g.f. in (2.14). Figure 9a shows how the probability of extinction starting from one infected cell changes as a function of θ , for different values of n_I , where θ is the probability that a given extracellular virion goes on to infect a new cell rather than be cleared. When θ is very small, extinction is certain because the mean reproduction number is below 1. When θ grows large enough to increase the mean reproduction number above 1 (i.e. $\theta > 1/E[B]$), then the probability of extinction becomes less than 1. As seen in the legend of the plot, the mean burst size increases with n_I . Thus, as n_I increases, the minimum value of θ needed to satisfy the condition $\theta > 1/E[B]$ becomes smaller. However, once this threshold value of θ is exceeded, larger values of n_I cause the probability of extinction to decay more slowly as a function of θ , compared with smaller values of n_I . As a result, we see that for small values of θ , there is a larger probability of extinction for smaller n_I , but as θ grows, the ordering changes. Eventually, for large enough values of θ , the probability of extinction is an increasing function of n_I . This is surprising, considering that the mean reproduction number also increases with n_I . The reason for this counterintuitive result is that the probability of eventual extinction depends strongly on the probability that a cell produces zero secondary infections. This probability, $\mathbb{P}(R=0)$, is more sensitive to θ for smaller values of n_I , which explains why the probability of extinction decays faster as a function of θ for smaller values of n_I . Figure 9b shows how the probability of extinction starting with one infected cell depends on the number of infectious stages (and \bar{R}), for the fixed value of $\theta \approx 0.59$ which comes from the parameter values in table 3. For this value of θ , it can be seen that the probability of extinction is an increasing function of n_I , even though the mean reproduction number \bar{R} grows with n_I . These results provide further evidence that focusing only on the mean reproduction number can be very misleading.

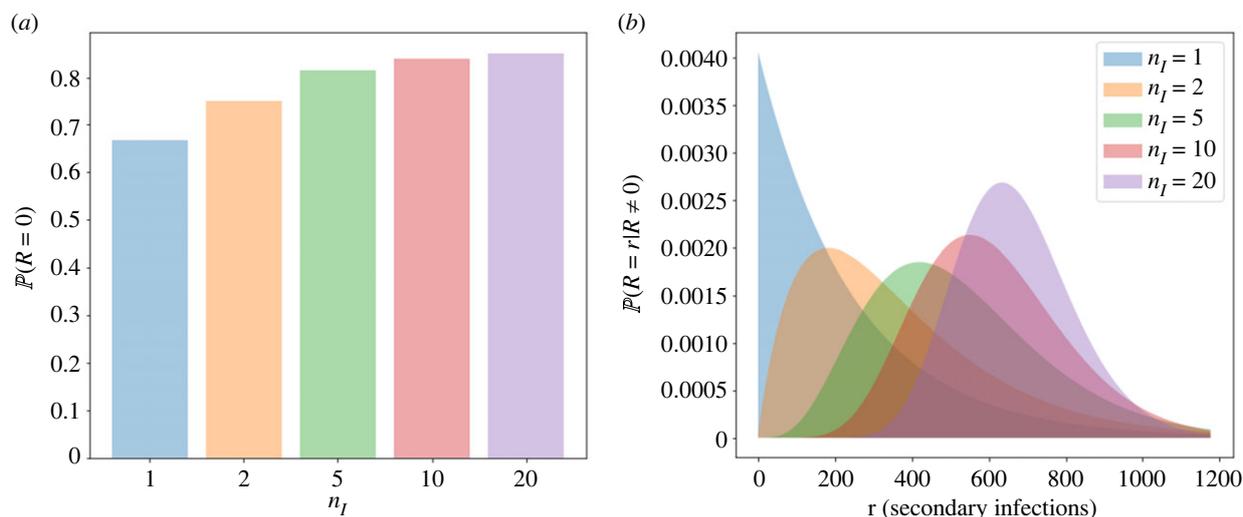


Figure 8. Reproduction number distributions calculated using the expression in table 2 for the bursting model, for different numbers of infectious phase stages, n_I . For each value of n_I , the mean of the Erlang-distributed time until cell burst is kept fixed to the value of τ_I in table 3. Other parameters are also set to the values in table 3, and the rate of immune killing of infectious cells is $\nu_I = 1.6$ per cell per day. (a) Probability of zero secondary infections for each value of n_I . (b) Probability distribution of R conditioned on positive values of the reproduction number.

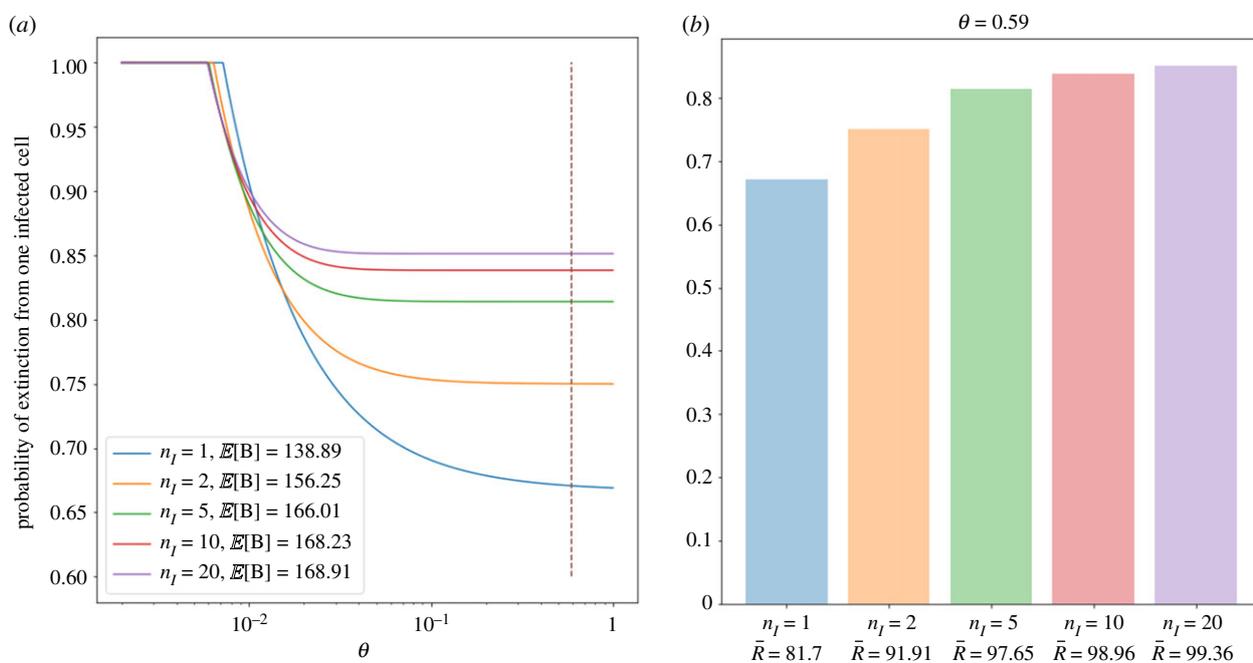


Figure 9. Plots to show how the probability of viral extinction depends on the number of infectious phase stages, n_I , for the model of viral release by bursting. The probability of extinction is calculated by numerically finding the smallest fixed point of the p.g.f. in (2.14). (a) The probability of viral extinction starting from one infected cell as a function of θ , for different values of n_I . The value of θ obtained from the values of β , c and T_0 in table 3 is indicated by the dashed line. For each value of n_I , the mean of the Erlang-distributed time until cell burst is kept fixed to the value of τ_I in table 3. The viral production rate is also fixed to the value of p in table 3, and the rate of immune killing of infectious cells is $\nu_I = 1.6$ per cell per day. (b) The probability of viral extinction starting from one infected cell is shown for different values of n_I , for $\theta = 0.59$ (the value indicated by the dashed line in the left plot). Along with the number of infectious stages, the mean reproduction number increases along the x-axis.

4. Discussion

The mean value of the cellular reproduction number has been computed making use of deterministic mathematical models for a number of viruses [28–31]. In some instances, these values have been compared between different viral strains as a measure of relative viral fitness [32]. Although the value of \bar{R} is often used to compare viruses and predict the outcome of infection, this number alone might not be enough to do so. For example, in the early stages of viral infection, stochastic effects are important. In fact, data on the probability of recovery

after acute HCV infection suggest that about 30% of infected people spontaneously clear infection and do not develop chronic infection [33]. When stochastic effects are considered, there is heterogeneity in the viral burst size, which is the number of viral progeny generated from an infected cell during its lifetime, as well as heterogeneity in the reproduction number, which is the number of secondary infections that a single infected cell produces. One can capture this heterogeneity by investigating the probability distributions of these two variables. Indeed, two viruses may have the same \bar{R} , but different reproduction number probability distributions, for

instance, owing to differences in the lifespans of infected cells or the mechanism of virion release.

Here we have shown how to calculate the probability distributions of the random variables for the burst size and the reproduction number, for two stochastic viral dynamics models. We considered a model in which virions are released from infectious cells by budding at a constant rate. This model, shown in (2.1) and figure 1, includes an eclipse phase to represent the period after a virion infects a target cell, during which the viral genome is replicated but no virions are released. The duration of the eclipse phase follows an Erlang distribution, as does the time spent in the infectious phase until virus-induced cell death. The time until infected cells are killed by the immune system is assumed to follow an exponential distribution. Thus, the time to the death of an infected cell is the minimum of these two competing mechanisms. In addition to the model of viral release by budding, we also considered a model of viral release by bursting. This model assumes that viral particles are generated intracellularly at a constant rate and are eventually released in a burst event, but only if the infected cell is not killed by the immune system.

Given the burst size distribution, we have presented two methods to calculate the reproduction number distribution. The simpler method assumes that the target cell population remains constant (i.e. it is not depleted as these cells become infected). This assumption is valid if the system is well mixed and there is a large number of target cells, which could be the case early in infection. However, for some routes of infection, or for particular viruses, the availability of target cells may be limited. Thus, we introduced a different method that can be used in this instance, in which the number of target cells decreases due to infection events.

For *in vitro* scenarios, the effect of the immune system can be removed by setting $v_E = v_I = 0$ in the model, as in the models by Liao *et al.* [2] and Yan *et al.* [7], which were parametrized using *in vitro* experimental data. In the case where $v_E = v_I = 0$, we showed that the burst size and reproduction numbers follow negative binomial distributions. We can draw a comparison with the epidemic reproduction number, where gamma-distributed individual heterogeneity in the infection process results in a negative binomial distribution of secondary infections from each case [34]. The shape parameter of the negative binomial reproduction number distribution is sometimes called the *dispersion parameter* of the distribution, and it represents the degree of transmission heterogeneity, which arises from a broad range of biological and social factors that influence transmission. A low value of the dispersion parameter corresponds to a high level of dispersion in the distribution and suggests that a small number of infected individuals, known as 'super-spreaders', may contribute to many secondary infections. The value of the dispersion parameter is particularly important for the early dynamics of an epidemic when there are only a few infected individuals. For example, Lloyd-Smith *et al.* [34] investigated the probability of stochastic extinction for an outbreak beginning with one infected individual, for different values of the dispersion parameter. Smaller values of the dispersion parameter were shown to increase the probability of stochastic extinction, but a small dispersion parameter also means that an epidemic can quickly take off due to the possibility of superspreading events, leading to infrequent but explosive epidemics. In the model studied here, the dispersion parameter of the negative binomial

distribution in (2.11), given by n_I , is an integer greater than or equal to 1, since it corresponds to the number of stages in the Erlang-distributed infectious period. However, heterogeneity in the within-host reproduction number can arise from a broad suite of factors other than the infectious period of the cell. Including other sources of variation into the model would impact the distribution of virions released from an infected cell. This could have consequences such as decreasing the predicted chance of successful infection establishment, or introducing the potential for superspreader cells.

For stochastic models of within-host viral dynamics, the distribution of the reproduction number can be used to calculate the probability that infection will become established in an individual, given an initial viral dose. We have shown how to calculate the probability of viral extinction for the budding and bursting models, and that the shape of the reproduction number distribution can have a significant impact on this probability. Previous studies of the probability of within-host viral extinction have focused on reproduction number distributions for which the means are the same in each case, but the shape of the distributions differ. This is similar to the case of epidemic models discussed above, where increased variability in individual infectiousness has been shown to increase the probability of stochastic extinction in an outbreak beginning with one infected individual [34]. For example, Pearson *et al.* [1] studied differences between continuous viral release and bursting, with the reproduction number being geometrically distributed in the continuous release case and assumed to either be a fixed value or Poisson distributed in the bursting case. By comparing the geometric reproduction number distribution and the Poisson distribution (with the same mean), Pearson *et al.* [1] showed that, with the greater variability of the geometric distribution, there was a larger chance of viral extinction and a lower probability of successful infection. Yuan and Allen [11] also considered a similar model to (2.1) but without an eclipse phase and with only one infectious stage. They studied the difference in the probability of viral extinction when comparing a model of the budding strategy with a geometrically distributed burst size, and the bursting strategy with a fixed burst size. They showed that the bursting strategy was more successful for viral invasion when a model without immune response was considered, but that the more successful strategy switched to budding when the immune response was included in the model. The results of both Pearson *et al.* and Yuan and Allen indicate that the distributions of the burst size and reproduction number are important in determining the probability of viral extinction; the mean alone is not enough. We have further shown that an increase in the expected reproduction number does not automatically imply a decrease in the probability of extinction, since the outcome depends on the distribution of the random variable.

In the electronic supplementary material, we studied a stochastic version of a previously published deterministic model of HCV infection in which the rate of virion release by budding is not constant, but depends on the intracellular viral dynamics [17]. The deterministic model by Guedj *et al.*, presented in equation (1) of the electronic supplementary material, is equivalent to a model in which the rate of viral release varies continuously over an infected cell's lifespan. For the parameter values estimated by Guedj *et al.*, the release rate approaches a constant steady state fairly quickly, making the burst size and reproduction number distributions similar

to the negative binomial distributions derived for the model of constant viral release by budding in the absence of immune response. However, the age-dependent viral release rate has a more noticeable impact when the infectious period follows an exponential distribution ($n_I = 1$), since cells that die very soon after infection release rather fewer virions than they would in the case of a constant release rate. The model for HCV considered here only includes simple intracellular dynamics; the model has a single equation for the intracellular viral genome, which increases by viral replication and is lost due to decay or release from the cell by budding. However, the method shown to calculate the burst size distribution for this particular model of HCV dynamics can be directly generalized to other models in which the release of virions from infected cells is a Poisson process and a functional form of the age-dependent viral release rate can be found. Furthermore, the method could be extended and implemented for models with more complex intracellular viral kinetics, such as those to describe co-infection and reassortment of segmented viruses.

A shortcoming of our approach is that we have focused only on stochastic models that have been derived from deterministic equations. *In vitro* experimental techniques that enable visualization of viral replication and virion release from single cells provide the opportunity for future work to focus on developing and parametrizing a much wider variety of stochastic models [35]. However, it will also be necessary to model the impact of *in vivo* processes that may not be observed *in vitro*, such as immune system clearance of infected cells. In §3.2, we showed that this can have important biological implications in terms of the likelihood of infection establishment. While it is challenging to directly measure immune system clearance experimentally, calibrating models with *in vivo* data can allow parameters that are not measurable *in vitro* to be estimated. For example, during treatment of chronic HCV infection, there is strong evidence of an exponential decay of infected cells [36,37], where this rapid death of infected cells is thought to be due to immune responses.

While the theory and examples presented here emphasize the importance of understanding reproduction number variation and modelling processes that contribute to this heterogeneity, there are important mechanisms that are not captured by the simple models presented here. For some viruses, such as HCV and respiratory syncytial virus, infected cells can directly infect target cells via cell–cell contacts [38–41]. Including this infection mechanism into the model may alter the predicted within-host reproduction number distribution. Furthermore, for the models presented here, variation in the reproduction number mainly results from variation in the infectious period of cells. However, there are many other aspects of host-cell biology that can make important contributions to heterogeneity of the within-host reproduction number. For example, cell-to-cell variation in the innate immune response

(e.g. the response to type I interferons (IFN)) can lead to cell differences in the efficiency of viral transcription [42]. While in the models presented here, each cell is assumed to be infected with only one virion, cellular co-infection is common for many viral diseases. Thus, another source of variation across a heterogeneous infected cell population is the multiplicity of infection of individual cells, which can lead to heterogeneity in cell death rates and viral production rates [43]. Finally, cells that are in a certain phase of the cell cycle when they become infected may enable more efficient replication for some viruses [44].

In conclusion, computing the burst size and reproduction number probability distributions for stochastic viral dynamics models is a first step to compare these distributions for a range of viruses (or strains) and model structures. Furthermore, the reproduction number distribution can have surprising consequences for the probability of viral extinction. We have demonstrated this through an example in which the probability of viral extinction actually increases as the mean reproduction number increases.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. Computer codes (in Python) for reproducing the results in figures 3–9 are available from the Zenodo repository: <https://doi.org/10.5281/zenodo.10391662> [45] and from the GitHub repository: <https://github.com/Bevelynn/viral-reproduction-number> [46].

Results for a mathematical model of HCV viral dynamics with an age-dependent viral release rate are provided in electronic supplementary material [47].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. B.W.: conceptualization, formal analysis, investigation, methodology, visualization, writing—original draft, writing—review and editing; J.C.: conceptualization, investigation, writing—review and editing; J.J.G.: funding acquisition, investigation, writing—review and editing; G.L.: formal analysis, funding acquisition, investigation, supervision, writing—review and editing; A.S.P.: investigation, writing—review and editing; R.M.R.: investigation, writing—review and editing; C.M.-P.: conceptualization, funding acquisition, investigation, methodology, supervision, writing—review and editing; M.L.-G.: conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, supervision, visualization, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. B.W. acknowledges support from an EPSRC CASE studentship, project reference 2345914, in partnership with Dstl under contract no. DSTLX-1000142022. This research has also been supported by the UKRI BBSRC through project reference BB/W010755/1, and the USA NIH through project nos. R01-AI167048 (C.M.-P.), R01-AI078881 and R01-OD011095 (A.S.P.), and R01-AI116868 (R.M.R.).

Acknowledgements. This manuscript has been internally reviewed at Los Alamos National Laboratory, and assigned the reference no. LA-UR-21-32242.

References

- Pearson JE, Kravitsky P, Perelson AS. 2011 Stochastic theory of early viral infection: continuous versus burst production of virions. *PLoS Comput. Biol.* **7**, e1001058. (doi:10.1371/journal.pcbi.1001058)
- Liao LE *et al.* 2020 Quantification of Ebola virus replication kinetics *in vitro*. *PLoS Comput. Biol.* **16**, e1008375. (doi:10.1371/journal.pcbi.1008375)
- Rong L, Guedj J, Dahari H, Coffield Jr DJ, Levi M, Smith P, Perelson AS. 2013 Analysis of hepatitis C virus decline during treatment with the protease inhibitor danoprevir using a multiscale model. *PLoS Comput. Biol.* **9**, e1002959. (doi:10.1371/journal.pcbi.1002959)
- Alizon S, Boldin B. 2010 Within-host viral evolution in a heterogeneous environment:

- insights into the HIV co-receptor switch. *J. Evol. Biol.* **23**, 2625–2635. (doi:10.1111/j.1420-9101.2010.02139.x)
5. Hataye JM *et al.* 2019 Principles governing establishment versus collapse of HIV-1 cellular spread. *Cell Host Microbe* **26**, 748–763. (doi:10.1016/j.chom.2019.10.006)
 6. Bacsik DJ, Dadonaite B, Butler A, Greaney AJ, Heaton NS, Bloom JD. 2023 Influenza virus transcription and progeny production are poorly correlated in single cells. *Elife* **12**, RP86852. (doi:10.7554/eLife.86852)
 7. Yan AW, Zhou J, Beauchemin CA, Russell CA, Barday WS, Riley S. 2020 Quantifying mechanistic traits of influenza viral dynamics using *in vitro* data. *Epidemics* **33**, 100406. (doi:10.1016/j.epidem.2020.100406)
 8. Burg D, Rong L, Neumann AU, Dahari H. 2009 Mathematical modeling of viral kinetics under immune control during primary HIV-1 infection. *J. Theor. Biol.* **259**, 751–759. (doi:10.1016/j.jtbi.2009.04.010)
 9. Nowak M, May RM. 2000 *Virus dynamics*. Oxford, UK: Oxford University Press.
 10. Wodarz D, Nowak MA. 2002 Mathematical models of HIV pathogenesis and treatment. *BioEssays* **24**, 1178–1187. (doi:10.1002/bies.10196)
 11. Yuan Y, Allen LJ. 2011 Stochastic models for virus and immune system dynamics. *Math. Biosci.* **234**, 84–94. (doi:10.1016/j.mbs.2011.08.007)
 12. Allen LJ, Jang SR, Roeger LI. 2017 Predicting population extinction or disease outbreaks with stochastic models. *Lett. Biomath.* **4**, 1–22. (doi:10.30707/LiB4.1Allen)
 13. Perelson AS. 2002 Modelling viral and immune system dynamics. *Nat. Rev. Immunol.* **2**, 28–36. (doi:10.1038/nri700)
 14. Mittler JE, Sulzer B, Neumann AU, Perelson AS. 1998 Influence of delayed viral production on viral dynamics in HIV-1 infected patients. *Math. Biosci.* **152**, 143–163. (doi:10.1016/S0025-5564(98)10027-5)
 15. Beauchemin CA, Miura T, Iwami S. 2017 Duration of SHIV production by infected cells is not exponentially distributed: implications for estimates of infection parameters and antiviral efficacy. *Sci. Rep.* **7**, 1–13. (doi:10.1038/srep42765)
 16. Holder BP, Beauchemin CA. 2011 Exploring the effect of biological delays in kinetic models of influenza within a host or cell culture. *BMC Public Health* **11**, 1–15. (doi:10.1186/1471-2458-11-S1-S10)
 17. Guedj J, Dahari H, Rong L, Sansone ND, Nettles RE, Cotler SJ, Layden TJ, Uprichard SL, Perelson AS. 2013 Modeling shows that the NS5A inhibitor daclatasvir has two modes of action and yields a shorter estimate of the hepatitis C virus half-life. *Proc. Natl Acad. Sci. USA* **110**, 3991–3996. (doi:10.1073/pnas.1203110110)
 18. Bai F, Huff KE, Allen LJ. 2019 The effect of delay in viral production in within-host models during early infection. *J. Biol. Dyn.* **13**, 47–73. (doi:10.1080/17513758.2018.1498984)
 19. Baral S, Raja R, Sen P, Dixit NM. 2019 Towards multiscale modeling of the CD8⁺ T cell response to viral infections. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **11**, e1446. (doi:10.1002/wsbm.1446)
 20. Sacha JB, Buechler MB, Newman LP, Reed J, Wallace LT, Loffredo JT, Wilson NA, Watkins DI. 2010 Simian immunodeficiency virus-specific CD8⁺ T cells recognize Vpr- and Rev-derived epitopes early after infection. *J. Virol.* **84**, 10907–10912. (doi:10.1128/JVI.01357-10)
 21. Perelson AS, Nelson PW. 1999 Mathematical analysis of HIV-1 dynamics *in vivo*. *SIAM Rev.* **41**, 3–44. (doi:10.1137/S0036144598335107)
 22. Canini L, Perelson AS. 2014 Viral kinetic modeling: state of the art. *J. Pharmacokinet. Pharmacodyn.* **41**, 431–443. (doi:10.1007/s10928-014-9363-3)
 23. de Jong LC, Crnko S, Ten Broeke T, Bovenschen N. 2021 Noncytotoxic functions of killer cell granzymes in viral infections. *PLoS Pathog.* **17**, e1009818. (doi:10.1371/journal.ppat.1009818)
 24. Andrade F, Fellows E, Jenne DE, Rosen A, Young C. 2007 Granzyme H destroys the function of critical adenoviral proteins required for viral DNA replication and granzyme B inhibition. *EMBO J.* **26**, 2148–2157. (doi:10.1038/sj.emboj.7601650)
 25. Czuppon P, Débarre F, Gonçalves A, Tenaillon O, Perelson AS, Guedj J, Blanquart F. 2021 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**, e1008752. (doi:10.1371/journal.pcbi.1008752)
 26. Allen LJ. 2010 *An introduction to stochastic processes with applications to biology*. Boca Raton, FL: CRC press.
 27. Williams BF. 2022 Mechanistic intracellular and within-host models of bacterial and viral infections. PhD thesis, University of Leeds, UK.
 28. González-Parra G, Dobrovolsky HM, Aranda DF, Chen-Charpentier B, Rojas RAG. 2018 Quantifying rotavirus kinetics in the REH tumor cell line using *in vitro* data. *Virus Res.* **244**, 53–63. (doi:10.1016/j.virusres.2017.09.023)
 29. Iwami S, Holder BP, Beauchemin CA, Morita S, Tada T, Sato K, Igarashi T, Miura T. 2012 Quantification system for the viral dynamics of a highly pathogenic simian/human immunodeficiency virus based on an *in vitro* experiment and a mathematical model. *Retrovirology* **9**, 1–12. (doi:10.1186/1742-4690-9-18)
 30. Baccam P, Beauchemin C, Macken CA, Hayden FG, Perelson AS. 2006 Kinetics of influenza A virus infection in humans. *J. Virol.* **80**, 7590–7599. (doi:10.1128/JVI.01623-05)
 31. Stafford MA, Corey L, Cao Y, Daar ES, Ho DD, Perelson AS. 2000 Modeling plasma virus concentration during primary HIV infection. *J. Theor. Biol.* **203**, 285–301. (doi:10.1006/jtbi.2000.1076)
 32. Iwanami S, Kakizoe Y, Morita S, Miura T, Nakaoka S, Iwami S. 2017 A highly pathogenic simian/human immunodeficiency virus effectively produces infectious virions compared with a less pathogenic virus in cell culture. *Theor. Biol. Med. Modell.* **14**, 1–13. (doi:10.1186/s12976-017-0055-8)
 33. World Health Organization. 2022 Hepatitis C. See <https://www.who.int/news-room/fact-sheets/detail/hepatitis-c>.
 34. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. 2005 Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355–359. (doi:10.1038/nature04153)
 35. Teufel AI, Liu W, Draghi JA, Cameron CE, Wilke CO. 2021 Modeling poliovirus replication dynamics from live time-lapse single-cell imaging data. *Sci. Rep.* **11**, 9622. (doi:10.1038/s41598-021-87694-x)
 36. Neumann AU, Lam NP, Dahari H, Gretch DR, Wiley TE, Layden TJ, Perelson AS. 1998 Hepatitis C viral dynamics *in vivo* and the antiviral efficacy of interferon- α therapy. *Science* **282**, 103–107. (doi:10.1126/science.282.5386.103)
 37. Dixit NM, Layden-Almer JE, Layden TJ, Perelson AS. 2004 Modelling how ribavirin improves interferon response rates in hepatitis C virus infection. *Nature* **432**, 922–924. (doi:10.1038/nature03153)
 38. Brimacombe CL *et al.* 2011 Neutralizing antibody-resistant hepatitis C virus cell-to-cell transmission. *J. Virol.* **85**, 596–605. (doi:10.1128/JVI.01592-10)
 39. Graw F, Martin DN, Perelson AS, Uprichard SL, Dahari H. 2015 Quantification of hepatitis C virus cell-to-cell spread using a stochastic modeling approach. *J. Virol.* **89**, 6551–6561. (doi:10.1128/JVI.00016-15)
 40. Huong TN, Iyer Ravi L, Tan BH, Sugrue RJ. 2016 Evidence for a biphasic mode of respiratory syncytial virus transmission in permissive Hep2 cell monolayers. *J. Virol.* **13**, 1–17. (doi:10.1186/s12985-016-0467-9)
 41. Zeng C, Evans JP, King T, Zheng YM, Oltz EM, Whelan SP, Saif LJ, Peebles ME, Liu SL. 2022 SARS-CoV-2 spreads through cell-to-cell transmission. *Proc. Natl Acad. Sci. USA* **119**, e2111400119. (doi:10.1073/pnas.2111400119)
 42. Wang C *et al.* 2020 Cell-to-cell variation in defective virus expression and effects on host responses during influenza virus infection. *MBio* **11**, 10–1128. (doi:10.1128/mBio.02880-19)
 43. Koelle K, Farrell AP, Brooke CB, Ke R. 2019 Within-host infectious disease models accommodating cellular coinfection, with an application to influenza. *Virus Evol.* **5**, vez018. (doi:10.1093/ve/vez018)
 44. Bagga S, Bouchard MJ. 2014 Cell cycle regulation during viral infection. In *Cell cycle control* (eds E Noguchi, M Gadaleta), pp. 165–227. Methods in molecular biology, vol 1170. New York, NY: Humana Press. (doi:10.1007/978-1-4939-0888-2_10)
 45. Williams B, Carruthers J, Gillard JJ, Lythe G, Perelson AS, Ribeiro RM, Molina-París C, López-García M. 2024 The reproduction number and its probability distribution for stochastic viral dynamics. Zenodo. (doi:10.5281/zenodo.10391662)
 46. Williams B, Carruthers J, Gillard JJ, Lythe G, Perelson AS, Ribeiro RM, Molina-París C, López-García M. 2024 The reproduction number and its probability distribution for stochastic viral dynamics. GitHub repository. (<https://github.com/Bevelynn/viral-reproduction-number>)
 47. Williams B, Carruthers J, Gillard JJ, Lythe G, Perelson AS, Ribeiro RM, Molina-París C, López-García M. 2024 The reproduction number and its probability distribution for stochastic viral dynamics. Figshare. (doi:10.6084/m9.figshare.c.7007799)