A model for in-host viral infection dynamics

Satvik Saha

1 Model descriptions

1.1 Individual based model

This model accommodates n individuals, each of which belongs to one of m environments. The individuals also form the nodes of a contact network. Each individual maintains its own in-host variables T, T^*, V, A which evolve over time. Each environment accumulates viral copies from its members, which decay over time. At the start of every day, each individual is stochastically assigned a subset of its contact neighbours, which together with the individual's environment contribute an 'external pressure' of viral copies. If this crosses a threshold v, then the individual becomes infected and experiences a rapid increase in its in-host viral load V.

Let $n, m \in \mathbb{N}$. For each $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$, let

$$\frac{dT_i}{dt} = b - \delta T_i - \frac{\kappa}{1 + \alpha A_i} T_i V_i, \tag{1.1.1}$$

$$\frac{dT_i^*}{dt} = \frac{\kappa}{1 + \alpha A_i} T_i V_i - q T_i^*, \tag{1.1.2}$$

$$\frac{dV_i}{dt} = p T_i^* - c V_i - c_A A_i V_i - X_i(t) + g\left(W_i + \sum_{j=1}^{m} \eta_{ij} Z_j\right), \tag{1.1.3}$$

$$\frac{dA_i}{dt} = b_A - \delta_A A_i + \kappa_A A_i(t - \tau) V_i(t - \tau), \tag{1.1.4}$$

$$\tag{1.1.5}$$

$$\frac{dZ_j}{dt} = \sum_{i=1}^{n} \xi_{ij} V_i - \delta_Z Z_j, \tag{1.1.6}$$

$$W_i(t) = \zeta \sum_{k=1}^{n} Y_{ik}(\lfloor t \rfloor) V_k(\lfloor t \rfloor), \tag{1.1.7}$$

$$g(x) = \begin{cases} x, & \text{if } x > v, \\ 0, & \text{if } x \leq v. \end{cases} \tag{1.1.8}$$

Here, we define random variables $X_i(t) \sim \text{Exp}(\lambda)$ drawn independently for each $t \geq 0$, and $Y_{ik}(\ell) \sim \text{Bernoulli}(s_{ik}\, p_{\text{inf}})$ drawn independently for each $\ell \in \mathbb{Z}_{\geq 0}$.

Note that the random functions $X_i(t)$ as well as the random variables $T_{ik}(\ell)$ are completely independent of state, and thus may be computed/sampled beforehand. It may help to think of a particular instance/choice for these quantities to be part of the initial conditions for this model.

The parameters ξ_{ij}, η_{ij} are to be thought of as weights linking individuals with their environments; the parameters s_{ik} are to be thought of as strengths of connections between individuals forming a network.

The role of the stochastic term $X_i(t)$ in the equation for dV_i/dt is to allow the possibility of *complete* removal of viral load V_i from an individual. This stochastic effect is prominent when $V_i \sim 1/\lambda$, and becomes negligible when $V_i \gg 1/\lambda$. This is introduced to prevent V_i from 'bouncing back' after an initial infection.

The model state is described by $(T, T^*, V, A, W, Z) \in \mathcal{S} \equiv \mathbb{R}^{3n}_{\geq 0} \times \mathcal{C}_\tau \times \mathcal{C}_\tau \times \mathbb{R}^m_{\geq 0}$, with $\mathcal{C}_\tau \equiv \mathcal{C}(0, \tau)$, $T \equiv (T_1, \ldots, T_n)$, and so on.

The model parameters are $(b, \delta, \kappa, q, p, c, b_A, \delta_A, \kappa_A, c_A, \alpha, \tau, \delta_Z, v, \lambda, p_{\text{inf}}, \zeta, \eta, \xi, s) \in \mathcal{P} \equiv \mathbb{R}^{16}_{\geq 0} \times [0,1] \times \mathbb{R}^{mn}_{\geq 0} \times \mathbb{R}^{mn}_{\geq 0} \times [0,1]^{n \times n}$, with $\eta \equiv [\eta_{ij}]_{ij}$, $\xi \equiv [\xi_{ij}]_{ij}$, and $s \equiv [s_{ik}]_{ik}$.

Simplifications: Given i, let $\eta_{ij} = \eta$, $\xi_{ij} = \xi$ for precisely one j and $\eta_{ij}, \xi_{ij} = 0$ for the rest. In other words, let each individual belong to precisely one environment. Furthermore, let $s_{ik} \in \{0, 1\}$, with all $s_{ii} = 0$.

With the underlying individual-environment and individual-individual connections fixed, the parameter space reduces to $\mathbb{R}^{19}_{\geq 0}$.

Table 1: State variables for model 1.1

Variable	Units	Interpretation
T	cells/ml	Concentration of target cells
T^*	cells/ml	Concentration of infected cells
V	copies/ml	Concentration of viral copies
A	imm/ml	Antibody/immunity level
W	copies/ml	Contact pressure of viral copies
Z	copies/m^2	Environmental viral copies

After choosing thresholds V' and A', we can count

$$S = \sum_{i=1}^{n} \mathbf{1}(A \leq A')\, \mathbf{1}(V \leq V') \tag{1.1.9}$$

$$I = \sum_{i=1}^{n} \mathbf{1}(V > V'), \tag{1.1.10}$$

$$R = n - S - I. \tag{1.1.11}$$

Here, S denotes the number of susceptible individuals, I denotes the number of infectuous individuals, and R denotes the number of recovered individuals. The threshold V' is chosen such that its contribution to the external pressure of viral copies is enough to cross the barrier v. The threshold A' is chosen such that a typical individual with that level of antibodies is immune to infection.

1.2 Simplified in-host model

Consider the in-host model described below.

Table 2: Parameters for model 1.1

Parameter	Domain	Units	Interpretation
b	$\mathbb{R}_{\geq 0}$	cells $\text{ml}^{-1}\text{day}^{-1}$	Generation rate of target cells
δ	$\mathbb{R}_{\geq 0}$	day^{-1}	Death rate of target cells
κ	$\mathbb{R}_{\geq 0}$	$\text{cells}^{-1}\text{ml day}^{-1}$	Infection rate of target cells
q	$\mathbb{R}_{\geq 0}$	day^{-1}	Death rate of infected cells
p	$\mathbb{R}_{\geq 0}$	copies $\text{cells}^{-1}\text{day}^{-1}$	Production rate of viral copies
c	$\mathbb{R}_{\geq 0}$	day^{-1}	Clearance rate of viral copies
b_A	$\mathbb{R}_{\geq 0}$	imm $\text{ml}^{-1}\text{day}^{-1}$	Generation rate of antibodies
δ_A	$\mathbb{R}_{\geq 0}$	day^{-1}	Clearance rate of antibodies
κ_A	$\mathbb{R}_{\geq 0}$	$\text{copies}^{-1}\text{ml day}^{-1}$	Production rate of antibodies
c_A	$\mathbb{R}_{\geq 0}$	$\text{imm}^{-1}\text{ml day}^{-1}$	Clearance rate of viral copies via antibodies
α	$\mathbb{R}_{\geq 0}$	imm^{-1}ml	Inhibition of viral-target contact
τ	$\mathbb{R}_{\geq 0}$	day	Delay in antibody production
δ_Z	$\mathbb{R}_{\geq 0}$	day^{-1}	Removal rate of viral copies
v	$\mathbb{R}_{\geq 0}$	copies ml^{-1}	Entry threshold of viral concentration
λ	$\mathbb{R}_{\geq 0}$	$\text{copies}^{-1}\text{ml day}$	Reciprocal of mean of stochastic viral removal
ζ	$\mathbb{R}_{\geq 0}$	day^{-1}	Rate of viral load transfer
η_{ij}	$\mathbb{R}_{\geq 0}$	ml^{-1}m^2 day^{-1}	Environment-Individual transmission rate of virus
ξ_{ij}	$\mathbb{R}_{\geq 0}$	ml m^{-2} day^{-1}	Viral shedding rate into environment
s_{ik}	$[0,1]$	–	Strength of contact between individuals
p_{inf}	$[0,1]$	–	Probability of viral load transfer

$$\frac{dT}{dt} = b - \delta T - \frac{\kappa}{1+\alpha A} TV, \tag{1.2.1}$$

$$\frac{dT^*}{dt} = \frac{\kappa}{1+\alpha A} TV - qT^*, \tag{1.2.2}$$

$$\frac{dV}{dt} = pT^* - cV - c_A AV, \tag{1.2.3}$$

$$\frac{dA}{dt} = b_A - \delta_A A + \kappa_A A(t-\tau)V(t-\tau). \tag{1.2.4}$$

Equilibria: Note that $(b/\delta, 0, 0, b_A/\delta_A)$ is a trivial infection-free equilibrium. In general, equilibria satisfy the following equations.

$$\frac{\kappa}{1+\alpha A} TV = b - \delta T, \tag{1.2.5}$$

$$b - \delta T = qT^*, \tag{1.2.6}$$

$$pT^* = (c + c_A A)V, \tag{1.2.7}$$

$$\delta_A A - b_A = \kappa_A AV. \tag{1.2.8}$$

Combining 1.2.6, 1.2.7, and 1.2.8 to eliminate T^* and V,

$$\frac{p}{q}(b - \delta T) = \frac{(c + c_A A)(\delta_A A - b_A)}{\kappa_A A}, \tag{1.2.9}$$

whence

$$T = \frac{b}{\delta} - \frac{q(c + c_A A)(\delta_A A - b_A)}{p\delta\kappa_A A}. \tag{1.2.10}$$

Furthermore, 1.2.5 and 1.2.8 give

$$1 + \alpha A = \frac{\kappa TV}{b - \delta T} = T\frac{(\delta_A A - b_A)/\kappa_A}{(b - \delta T)/\kappa},$$

whence

$$\frac{b - \delta T}{\kappa T} = \frac{\delta_A A - b_A}{\kappa_A(1 + \alpha A)}.$$

Using 1.2.9,

$$T = \frac{q(c + c_A A)(1 + \alpha A)}{p\kappa A}. \tag{1.2.11}$$

Equating 1.2.10 and 1.2.11,

$$\frac{b}{\delta} = \frac{q(c + c_A A)}{pA}\left[\frac{1 + \alpha A}{\kappa} + \frac{\delta_A A - b_A}{\delta\kappa_A}\right].$$

Putting $T_0 = b/\delta$, $A_0 = b_A/\delta_A$, we have

$$pAT_0 = q(c + c_A A)\left[\frac{1 + \alpha A}{\kappa} + \frac{\delta_A(A - A_0)}{\delta\kappa_A}\right],$$

whence

$$p\kappa T_0 A = q(c + c_A A)\left[1 - \frac{\kappa/\delta}{\kappa_A/\delta_A}A_0 + \left(\alpha + \frac{\kappa/\delta}{\kappa_A/\delta_A}\right)A\right].$$

Setting $\beta = (\kappa/\delta)/(\kappa_A/\delta_A)$, $\varphi = p\kappa T_0/qc = pb\kappa/qc\delta$, $\gamma = c_A/c$, we have

$$\varphi A = (1 + \gamma A)[1 - \beta A_0 + (\alpha + \beta)A].$$

Thus,

$$\gamma(\alpha + \beta)A^2 + [\gamma(1 - \beta A_0) + (\alpha + \beta) - \varphi]A + (1 - \beta A_0) = 0,$$

or

$$A^2 + \left[\frac{1 - \beta A_0}{\alpha + \beta} + \frac{1}{\gamma} - \frac{\varphi}{\gamma(\alpha + \beta)}\right]A + \frac{1 - \beta A_0}{\gamma(\alpha + \beta)} = 0.$$

1.3 In-host submodel

This model is a variation of model 1.2, with an additional stochastic term $X(t)$ and an external forcing term $W(t)$ present in the equation for dV/dt, in the manner of model 1.1.

Alternatively, this model may be thought of as a reduction of model 1.1 with $n = 1$, $m = 0$, i.e. one individual not tied to any environment.

$$\frac{dT}{dt} = b - \delta T - \frac{\kappa}{1 + \alpha A}TV, \tag{1.3.1}$$

$$\frac{dT^*}{dt} = \frac{\kappa}{1 + \alpha A}TV - qT^*, \tag{1.3.2}$$

$$\frac{dV}{dt} = pT^* - cV - c_A AV - X(t) + W(t), \tag{1.3.3}$$

$$\frac{dA}{dt} = b_A - \delta_A A + \kappa_A A(t - \tau)V(t - \tau). \tag{1.3.4}$$

Here, the random variables $X(t) \sim \mathrm{Exp}(\lambda)$ are drawn independently for each $t \geq 0$.

The forcing term $W(t)$ may be chosen to induce infection at chosen points in time; see Figure 1.

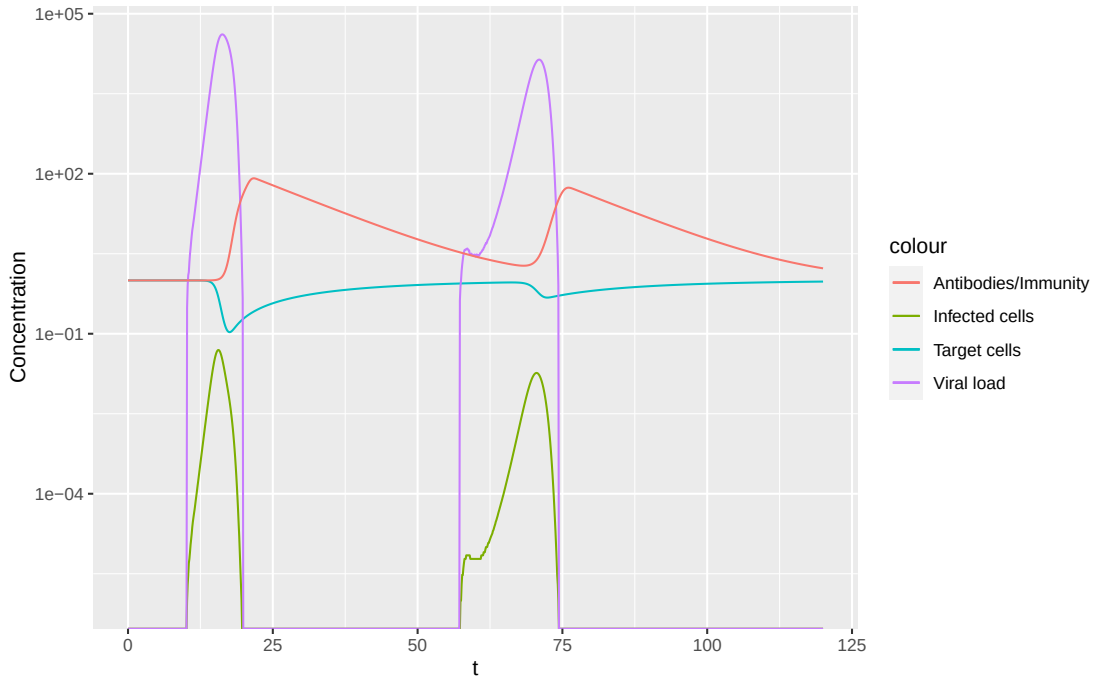

Figure 1: Curves generated by one run of model 1.3, where $W(t) = \mathbf{1}_{[10,11]}(t)+\mathbf{1}_{[57,58]}(t)$. In other words, the external pressure of viral copies $W(t)$ pulses on days 10 and 57 for the duration of one day each. Note that the concentrations of T, T^* have been presented as a fraction of $T_0 = b/\delta$.

1.4 Multiscale model

This model is heavily inspired by [2].

$$\frac{dS}{dt} = -\beta_I(V,I)SI - \beta_Z(Z)SZ + \mu R, \tag{1.4.1}$$

$$\frac{dI}{dt} = \beta_I(V,I)SI + \beta_Z(Z)SZ - \gamma I, \tag{1.4.2}$$

$$\frac{dR}{dt} = \gamma I - \mu R, \tag{1.4.3}$$

$$\tag{1.4.4}$$

$$\frac{dZ}{dt} = \xi I - \delta_Z Z, \tag{1.4.5}$$

$$\tag{1.4.6}$$

$$\epsilon\frac{dT}{dt} = b - \delta T - \frac{\kappa}{1+\alpha A}TV, \tag{1.4.7}$$

$$\epsilon\frac{dT^*}{dt} = \frac{\kappa}{1+\alpha A}TV - qT^*, \tag{1.4.8}$$

$$\epsilon\frac{dV}{dt} = \eta Z + pT^* - cV - c_A AV, \tag{1.4.9}$$

$$\epsilon\frac{dA}{dt} = b_A - \delta_A A + \kappa_A A(t-\tau)V(t-\tau). \tag{1.4.10}$$

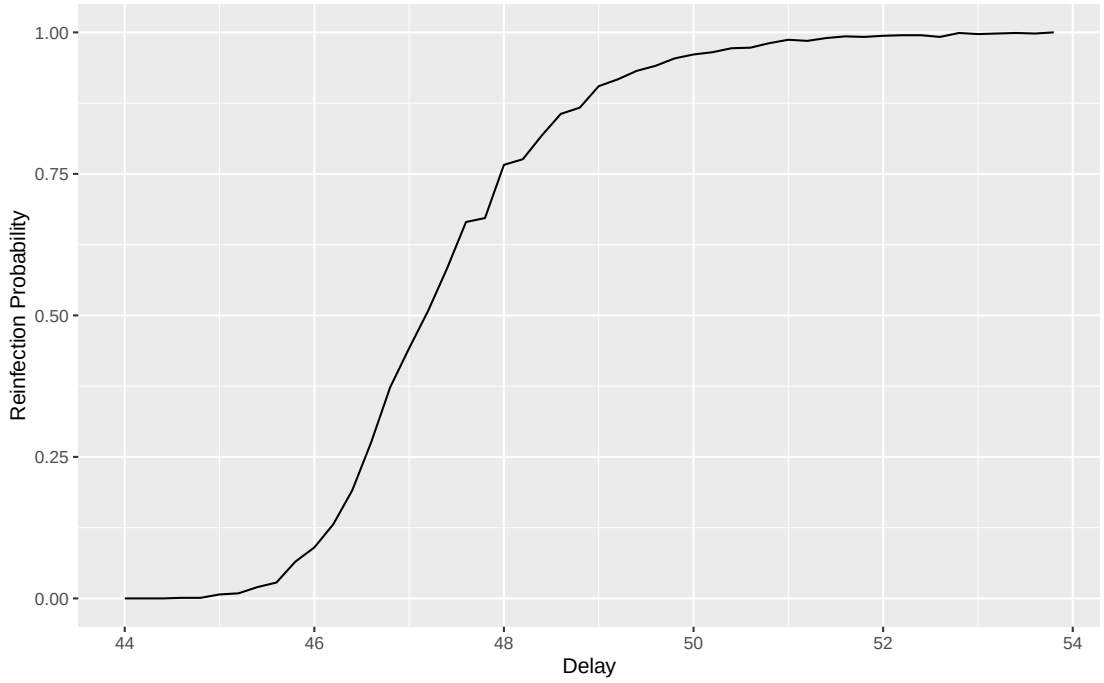

Figure 2: Estimated probabilities of reinfection in model 1.3 as a function of time since first infection t'. For each t', we set $W(t) = \mathbf{1}_{[10,11]}(t) + \mathbf{1}_{[10+t',11+t']}(t)$ and count the proportion of runs (out of 1000) for which a second peak in V is observed. A second peak is characterized by V crossing the threshold $V' = 10^2$ copies/ml.

Here,

$$\beta_I(V, I) = \frac{\beta_{I0} + C_0 V}{1 + C_1 I}, \qquad \beta_Z(Z) = \frac{\beta_{Z0}}{1 + C_2 Z}.$$

1.5 SIRS model

$$\frac{dS}{dt} = -\frac{\beta SI}{N} + \mu R, \tag{1.5.1}$$

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I, \tag{1.5.2}$$

$$\frac{dR}{dt} = \gamma I - \mu R. \tag{1.5.3}$$

Here, $N = S + I + R$.

2 Objectives

1. Compare the S, I, R curves from model 1.1 with those obtained from the multiscale model 1.4 and the SIRS model 1.5.

2. Identify/interpret infection phases (S, I, R) in model 1.1 using the in-host variables (T, T^*, V, A).

3. Investigate the effects of heterogeneity in the individuals and their contact network. For instance,

(a) In-host parameters may be varied across individuals, forming two or more groups.

(b) Groups of individuals may be vaccinated.

4. Investigate the effect of the stochastic term $X_i(t)$ in the in-host model 1.3.

3 Methods

3.1 Variation of infection probability

We run model 1.1 multiple times (200) any given combination of parameters. This process is repeated for $p_{\text{inf}} = x \times 10^{-3}$, for $x \in \{2, 3, 4, 5, 6, 8\}$, keeping all other parameters constant. The resulting infection curves have been illustrated in Figure 3.

3.2 Counting epidemic waves

Given an infection curve generated by model 1.1, we count the number of epidemic waves by identifying peaks. To do this, we first smooth the curve using the `ksmooth` algorithm, using a Gaussian kernel. Peaks in the infection curve are then approximated as peaks in the smoothed curve.

The distribution of the number of epidemic waves has been illustrated in Figure 4.

3.3 Estimating reinfection probabilities

We estimate the probability of reinfection in model 1.3 as a function of time since first infection t'. We say that an individual has been infected if it experiences an increase in viral load V above a threshold V'. We say that an (uninfected) individual has been reinfected by a pulse in the external pressure of viral copies W if it experiences a subsequent infection.

We perform multiple runs of model 1.3. For each t', we set $W(t) = \mathbf{1}_{[10,11]}(t) + \mathbf{1}_{[10+t',11+t']}(t)$ and count the proportion of runs (out of 1000) for which a second peak in V is observed. A second peak is characterized by V crossing the threshold $V' = 10^2$ copies/ml.

The estimates of reinfection probabilities have been illustrated in Figure 2.

3.4 Comparison with an SIRS model

We fit the SIRS model 1.5 to multiple individual infection curves from model 1.1 (with the same set of parameters) simultaneously. The parameters β, γ, μ are common, while the initial conditions $(N - I_{0i}, I_{0i}, 0)$ are allowed to vary across different curves (indexed by i) from 1.1. Infection curves which soon become extinct are discarded, and the remaining ones are truncated up to their first peak.

Let $I(\beta, \gamma, \mu, I_0)$ be the solution of model 1.5, and let I_i for $i \in \{1, 2, \ldots, r\}$ be individual infection curves from model 1.1. If I_i has time points $1, 2, \ldots, t_i$, then we denote the sum of squares distance

$$\|I(\beta, \gamma, \mu, I_0) - I_i\|^2 = \sum_{t=1}^{t_i} |I(\beta, \gamma, \mu, I_0)(t) - I_i(t)|^2.$$

The aforementioned fitting process refers to finding optimal $\beta, \gamma, \mu, I_{01}, I_{02}, \ldots, I_{0r}$ which minimize

$$\sum_{i=1}^{r} \|I(\beta, \gamma, \mu, I_{0i}) - I_i\|^2 = \sum_{i=1}^{r} \sum_{t=1}^{t_i} |I(\beta, \gamma, \mu, I_{0i})(t) - I_i(t)|^2.$$

4 Observations

1. The individual based model 1.1 is capable of producing infection curves with multiple waves/peaks. This has been illustrated in Figure 3.

2. The average viral load across all individuals in model 1.1 is in good agreement with the infection curve for a given run.

3. There is a narrow range of p_{inf}, with all other parameters in model 1.1 fixed, in which a significant proportion of infection curves display multiple prominent peaks without damping. For lower p_{inf}, infection curves become more stochastic in nature. The distribution of the number of peaks has been illustrated in Figure 4.

4. Averaged infection curves from model 1.1 also show multiple peaks; the curve up to the first peak fits well against the SIRS model 1.5.

5. We fit the SIRS model 1.5 to multiple individual infection curves from model 1.1 (with the same set of parameters) simultaneously. The parameters β, γ, μ are common, while the initial conditions $(N - I_0, I_0, 0)$ are allowed to vary across different curves from 1.1. Infection curves which soon become extinct are discarded, and the remaining ones are truncated up to their first peak.

   The SIRS model 1.5 fits reasonably well against such curves from model 1.1. However, the descent from the peak of infection in 1.5 is typically shallower than that in 1.1 (where peaks look somewhat symmetric).

   Keeping all parameters except for p_{inf} in model 1.1 constant and performing this fit for $p_{\text{inf}} = x \times 10^{-3}$ for $x \in \{3, 4, 5, 6, 8\}$, we see that the estimate for μ is typically zero. Furthermore, the estimate for γ is typically $1/7$. This conforms well with the behaviour of individuals in models 1.1 and 1.3, whose infection period is approximately 7 days.

6. Individuals in models 1.1 and 1.3 become 'infected' when a pulse is applied on W. The viral load V rapidly increases, which after a short delay leads to a rapid increase in the antibody/immunity A. This forces V to fall sharply to zero, after which A gradually drops back to its baseline level. This behaviour can be observed in Figure 1.

   A sufficiently elevated A confers 'immunity' to the individual, preventing reinfection. The probability of reinfection, as a function of time since infection, can be calculated. This has been illustrated in Figure 2.

References

[1] S. M. Ciupe and J. M. Heffernan. "In-host modeling". In: *Infect Dis Model* 2.2 (May 2017), pp. 188–202.

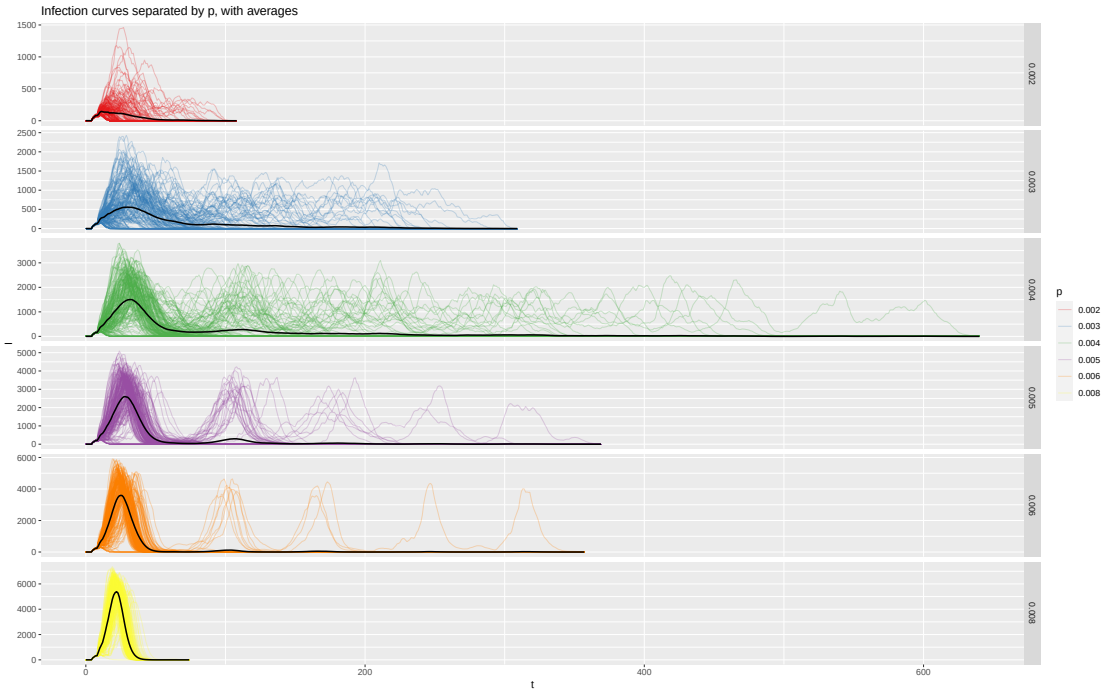

Figure 3: Infection curves generated by model 1.1, by varying infection probabilities p_{inf}. The black curves track the mean number of infected individuals at that time.

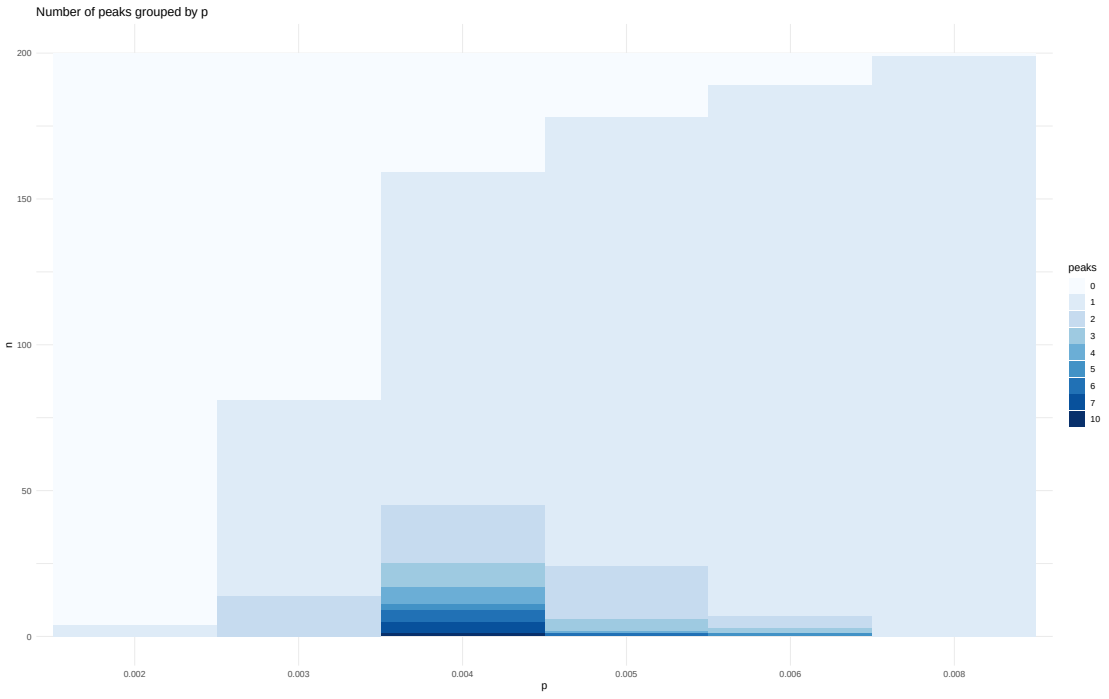

Figure 4: Number of runs (out of 200) of model 1.1 with n peaks in the infection curve, by varying infection probabilities p_{inf}.

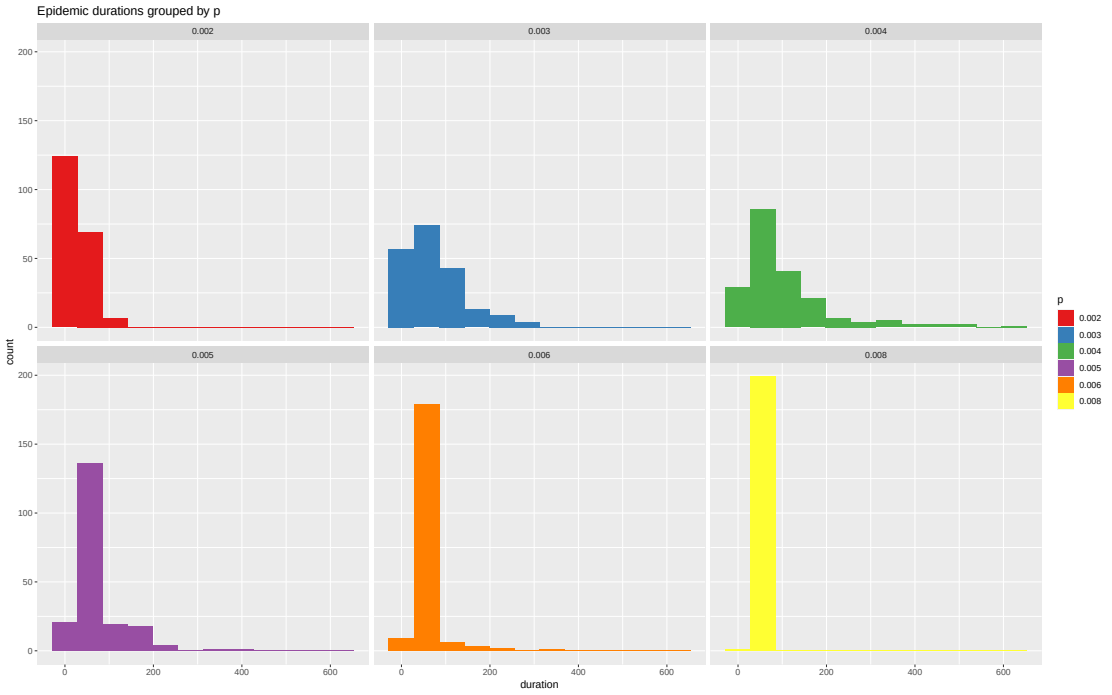

Figure 5: Distribution of epidemic durations in model 1.1, by varying infection probabilities p_{inf}.

[2] X. Wang et al. "A Multiscale Model of COVID-19 Dynamics". In: *Bull Math Biol* 84.9 (Aug. 2022), p. 99.

[3] C. Y. Yang and J. Wang. "A mathematical model for the novel coronavirus epidemic in Wuhan, China". In: *Math Biosci Eng* 17.3 (Mar. 2020), pp. 2708–2724.