

SUPPLEMENTARY INFORMATION:  
MATHEMATICAL DETAILS OF THE MODELS DISCUSSED IN THE PAPER

INTRODUCTION

This document presents the model and derivation of the various  $R_0$  presented in the paper. First, we introduce the overall model with both synaptic and free-virus transmission. We calculate  $R_0$  for the sub-model which only includes free-virus transmission. Then, we calculate  $R_0$  for the model with just synaptic transmission. We provide numerical results to support the combined  $R_0$  for the complete model. Finally, we explore different formulations for the rate of synaptic transmission.

COMPLETE MODEL WITH SYNAPTIC AND FREE VIRUS TRANSMISSION

$$\frac{dy_{0,0}}{dt} = \lambda - dy_{0,0} - \beta y_{0,0}v - y_{0,0} \sum_{k=0}^N \sum_{l=0}^M y_{kl} \sum_{r=0}^N \sum_{s=0}^M \gamma_{rs}^{kl}$$

$$\frac{dy_{i,j}}{dt} = (1-p)\beta v y_{i-1,j} + p\beta v y_{i,j-1} - a_{i,j} y_{i,j} - \beta v y_{i,j} + \sum_{k=0}^N \sum_{l=0}^M y_{kl} \left( \sum_{r=0}^i \sum_{s=0}^j \gamma_{rs}^{kl} y_{i-r,j-s} - y_{ij} \sum_{r=0}^{N-i} \sum_{s=0}^{M-j} \gamma_{rs}^{kl} \right)$$

$$\frac{dv}{dt} = \sum_{i=0}^N \sum_{j=0}^M k_{ij} y_{ij} - uv$$

Note that  $\gamma_{rs}^{00}$ ,  $\gamma_{00}^{kl}$ , and  $k_{00}$  are all zero. These are synaptic transmission rates and viral production, respectively.

Notation:

$y_{ij}$  A cell with Unintegrated (i), Integrated (j) viruses  
 N, maximum number of unintegrated virus per cell  
 M, maximum number of integrated virus per cell  
 k,r: dummy variables used as counters for i (integrated virus)  
 l,s: dummy variables used as counters for j (unintegrated virus)

Parameters:

$\lambda$  : Birth rate of uninfected CD4+ T-cells  
 $d$  : Death rate of uninfected CD4+ T-cells.  
 $\beta$  : Infectivity - effectiveness of the interaction between virus and a CD4+ T cell

$p$  : Probability of integration.

$k_{i,j}$  : Virus replication parameter - rate of free virus production by an infected cell

$u$  : Clearance rate of free virus

$a_{i,j}$  : Clearance rate of infected cells

$a_{i,j}$  can be defined as:

$$a_{ij} = \begin{cases} d & \text{if } j = 0, i < c \\ a_u & \text{if } j = 0, i \geq c \\ a_i & \text{if } j > 0 \end{cases}$$

$k_{i,j}$  can be defined as:

$$k_{ij} = \begin{cases} 0 & \text{if } j = 0, i < c \\ k_u & \text{if } j = 0, i \geq c \\ k_i & \text{if } j > 0 \end{cases}$$

$c$  : the threshold number of uDNA viruses necessary for uDNA-only cells to begin production.

$\gamma_{rs}^{kl}$  : the probability that a cell with  $k$  integrated,  $l$  unintegrated viruses transfers  $r$  integrated and  $s$  unintegrated viruses. We set  $\gamma_{00}^{kl} = 0$  and  $\gamma_{rs}^{00} = 0$

$\sum_{k=1}^N \sum_{l=1}^M y_{kl} (\sum_{r=1}^i \sum_{s=1}^j \gamma_{rs}^{kl} y_{i-r,j-s})$  Represents cells transmitting to other cells to create  $y_{ij}$ . The second set of summations only reach  $i$  and  $j$  because a cell can only transmit up to  $i$  and  $j$  viruses to an empty cell to get  $y_{ij}$ . The  $y_{kl}$  term represents the transmitting (source) cell. Any productively infected cell can be a source cell.  $y_{i-r,j-s}$  represents the target cell that accepts viruses.

$\sum_{k=1}^N \sum_{l=1}^M y_{kl} (y_{ij} \sum_{r=1}^{N-i} \sum_{s=1}^{M-j} \gamma_{rs}^{kl})$  : Represents cells transmitting to  $y_{ij}$

## $R_0$ CALCULATION: FREE-VIRUS TRANSFER ALONE

$$\frac{dy_{0,0}}{dt} = \lambda - dy_{0,0} - \beta y_{0,0}v$$

$$\frac{dy_{i,j}}{dt} = (1-p)\beta v y_{i-1,j} + p\beta v y_{i,j-1} - a_{i,j} y_{i,j} - \beta v y_{i,j}$$

$$\frac{dv}{dt} = \sum_{i=1}^N \sum_{j=1}^M k_{ij} y_{ij} - uv$$

The virus-free equilibrium of the above system is  $y_{00} = \lambda/d, y_{ij} = 0, v = 0$ . With standard methods, we can calculate the Jacobian of this system at this steady state to arrive at  $R_0$ . Below is the Jacobian of the system with the assumption that there are two distinct death rates and two distinct virus production rates. Cells infected with only unintegrated viruses produce virus and die at a lower rate than the cells infected with provirus.

$$J_{NM+1 \times NM+1} = \begin{pmatrix} -d & 0 & 0 & \dots & 0 & 0 & \dots & 0 & -\frac{\beta\lambda}{d} \\ 0 & -a_{\text{un}} & 0 & \dots & 0 & 0 & \dots & 0 & -\frac{(-1+p)\beta\lambda}{d} \\ 0 & 0 & -a_{\text{un}} & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & -a_{\text{in}} & 0 & \dots & 0 & \frac{p\beta\lambda}{d} \\ 0 & 0 & 0 & \dots & 0 & -a_{\text{in}} & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & -a_{\text{in}} & 0 \\ 0 & k_{\text{un}} & k_{\text{un}} & \dots & k_{\text{in}} & k_{\text{in}} & \dots & k_{\text{in}} & -u \end{pmatrix}$$

All but 3 of the eigenvalues for this system are  $-d, -a_i$  and  $-a_u$ . The remaining eigenvalues are the roots of the 3rd degree polynomial:

$$p(x) = a_3x^3 + a_2x^2 + a_1x + a_0$$

Where

$$\begin{aligned} a_3 &= d \\ a_2 &= du + da_{\text{in}} + da_{\text{un}} \\ a_1 &= dua_{\text{un}} + da_{\text{in}}(u + a_{\text{un}}) - \beta\lambda(pk_{\text{in}} - (p-1)k_{\text{un}}) \\ a_0 &= p\beta\lambda a_{\text{un}}k_{\text{in}} - a_{\text{in}}(dua_{\text{un}} + (-1+p)\beta\lambda k_{\text{un}}) \end{aligned}$$

These eigenvalues are nonnegative when they satisfy the Routh-Hurwitz criterion, which occurs whenever the following inequality is satisfied:

$$R_0 \text{ Free Virus Only, No Threshold} : \frac{p\beta\lambda k_i}{ua_i d} + \frac{(1-p)\beta\lambda k_u}{ua_u d} > 1$$

$R_0$  CALCULATION: FREE-VIRUS TRANSFER WITH THRESHOLD FOR ACTIVE UDNA

If a cell without an integrated provirus requires more than one unintegrated virus to begin production then the system of ODE remains the same, except the virus production rate ( $k_u$ ) is zero for cells without integrated viruses and below a threshold number of unintegrated viruses. In addition, these cells have a clearance rate equal to that of uninfected cells.

The Jacobian for this system becomes:

$$J_{NM+1 \times NM+1} = \begin{pmatrix} -d & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0 & -\frac{\beta\lambda}{d} \\ 0 & -d & \dots & 0 & \dots & 0 & 0 & \dots & 0 & -\frac{(-1+p)\beta\lambda}{d} \\ 0 & 0 & \ddots & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & -a_{un} & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \ddots & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & -a_{in} & 0 & \dots & 0 & \frac{p\beta\lambda}{d} \\ 0 & 0 & \dots & 0 & \dots & 0 & -a_{in} & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 & \dots & -a_{in} & 0 \\ 0 & 0 & \dots & k_{un} & \dots & k_{in} & k_{in} & \dots & k_{in} & -u \end{pmatrix}$$

With the requirement that more than one uDNA virus is necessary to begin virus production, all but 2 of the eigenvalues are  $-d$ ,  $-a_i$  or  $-a_u$ . The last remaining eigenvalues are:

$$\frac{1}{2} \left( -u - a_{in} \pm \frac{\sqrt{du^2 - 2dua_{in} + da_{in}^2 + 4p\beta\lambda k_{in}}}{\sqrt{d}} \right)$$

which is non-negative whenever:

$$R_0 \text{ Free Virus Only, With Threshold} = \frac{p\beta\lambda k_i}{ua_i d} > 1$$

## $R_0$ CALCULATION: SYNAPTIC TRANSFER ALONE

$$\begin{aligned} \frac{dy_{0,0}}{dt} &= \lambda - dy_{0,0} - y_{0,0} \sum_{k=0}^N \sum_{l=0}^M y_{kl} \sum_{r=0}^N \sum_{s=0}^M \gamma_{rs}^{kl} \\ \frac{dy_{i,j}}{dt} &= -a_{i,j} y_{i,j} + \sum_{k=0}^N \sum_{l=0}^M y_{kl} \left( \sum_{r=0}^i \sum_{s=0}^j \gamma_{rs}^{kl} y_{i-r,j-s} - y_{ij} \sum_{r=0}^{N-i} \sum_{s=0}^{M-j} \gamma_{rs}^{kl} \right) \end{aligned}$$

The virus-free equilibrium of the above system is  $y_{00} = \lambda/d, y_{ij} = 0$ . With standard methods, we can calculate the Jacobian of this system at this steady state to arrive at  $R_0$ . Below is the Jacobian of the system with the assumption that there are two death rates: death rates for cells infected with uDNA only and a separate death rate for iDNA infected cells. Note that we make no assumption about the values of  $\gamma$ .

The Jacobian for the system with synaptic transfer alone is an NM by NM matrix:

$$J = \begin{pmatrix} -d & -\frac{\lambda}{d} \sum \frac{\lambda}{d} \gamma_{ij}^{01} & \sum \frac{\lambda}{d} \gamma_{ij}^{02} & \dots & \sum \frac{\lambda}{d} \gamma_{ij}^{10} & \sum \frac{\lambda}{d} \gamma_{ij}^{11} & \dots & \sum \frac{\lambda}{d} \gamma_{ij}^{NM} \\ 0 & -a_u + \frac{\lambda}{d} \gamma_{01}^{01} & \frac{\lambda}{d} \gamma_{01}^{02} & \dots & \frac{\lambda}{d} \gamma_{01}^{10} & \frac{\lambda}{d} \gamma_{01}^{11} & \dots & \frac{\lambda}{d} \gamma_{01}^{NM} \\ 0 & \frac{\lambda}{d} \gamma_{02}^{01} & -a_u + \frac{\lambda}{d} \gamma_{02}^{02} & \dots & \frac{\lambda}{d} \gamma_{02}^{10} & \frac{\lambda}{d} \gamma_{02}^{11} & \dots & \frac{\lambda}{d} \gamma_{02}^{NM} \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \frac{\lambda}{d} \gamma_{10}^{01} & \frac{\lambda}{d} \gamma_{10}^{02} & \dots & -a_i + \frac{\lambda}{d} \gamma_{10}^{10} & \frac{\lambda}{d} \gamma_{10}^{11} & \dots & \frac{\lambda}{d} \gamma_{10}^{NM} \\ 0 & \frac{\lambda}{d} \gamma_{11}^{01} & \frac{\lambda}{d} \gamma_{11}^{02} & \dots & \frac{\lambda}{d} \gamma_{11}^{10} & -a_i + \frac{\lambda}{d} \gamma_{11}^{11} & \dots & \frac{\lambda}{d} \gamma_{11}^{NM} \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \frac{\lambda}{d} \gamma_{NM}^{01} & \frac{\lambda}{d} \gamma_{NM}^{02} & \dots & \frac{\lambda}{d} \gamma_{NM}^{10} & \frac{\lambda}{d} \gamma_{NM}^{11} & \dots & -a_i + \frac{\lambda}{d} \gamma_{NM}^{NM} \end{pmatrix}$$

where ,  $\sum = \sum_{i=0}^N \sum_{j=0}^M$ :

The first eigenvalue of the system is -d. To determine the remaining eigenvalues we look at the remaining NM-1 by NM-1 sub matrix:

$$\hat{J} = A_u + \begin{pmatrix} \frac{\lambda}{d} \gamma_{01}^{01} & \frac{\lambda}{d} \gamma_{01}^{02} & \dots & \frac{\lambda}{d} \gamma_{01}^{10} & \frac{\lambda}{d} \gamma_{01}^{11} & \dots & \frac{\lambda}{d} \gamma_{01}^{NM} \\ \frac{\lambda}{d} \gamma_{02}^{01} & \frac{\lambda}{d} \gamma_{02}^{02} & \dots & \frac{\lambda}{d} \gamma_{02}^{10} & \frac{\lambda}{d} \gamma_{02}^{11} & \dots & \frac{\lambda}{d} \gamma_{02}^{NM} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\lambda}{d} \gamma_{10}^{01} & \frac{\lambda}{d} \gamma_{10}^{02} & \dots & q + \frac{\lambda}{d} \gamma_{10}^{10} & \frac{\lambda}{d} \gamma_{10}^{11} & \dots & \frac{\lambda}{d} \gamma_{10}^{NM} \\ \frac{\lambda}{d} \gamma_{11}^{01} & \frac{\lambda}{d} \gamma_{11}^{02} & \dots & \frac{\lambda}{d} \gamma_{11}^{10} & q + \frac{\lambda}{d} \gamma_{11}^{11} & \dots & \frac{\lambda}{d} \gamma_{11}^{NM} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\lambda}{d} \gamma_{NM}^{01} & \frac{\lambda}{d} \gamma_{NM}^{02} & \dots & \frac{\lambda}{d} \gamma_{NM}^{10} & \frac{\lambda}{d} \gamma_{NM}^{11} & \dots & q + \frac{\lambda}{d} \gamma_{NM}^{NM} \end{pmatrix}$$

where  $q = a_u - a_i$  and  $A_u = a_u I$ .

This matrix has (NM-1)(NM-1)+3 different variables and to simplify it we assume that the transfer rate  $\gamma_{rs}$  is constant for all iDNA-infected cells, but for uDNA-infected cells the rate is  $\eta \gamma_{rs}$ , where  $\eta < 1$

Under these assumptions, we can replace the above Jacobian by:

$$\hat{J} = A_u + \begin{pmatrix} \frac{\lambda}{d}\gamma_{01} & \frac{\lambda}{d}\gamma_{01} & \dots & \frac{\lambda}{d}\gamma_{01} & \frac{\lambda}{d}\gamma_{01} & \dots & \frac{\lambda}{d}\gamma_{01} \\ \frac{\lambda}{d}\gamma_{02} & \frac{\lambda}{d}\gamma_{02} & \dots & \frac{\lambda}{d}\gamma_{02} & \frac{\lambda}{d}\gamma_{02} & \dots & \frac{\lambda}{d}\gamma_{02} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\lambda}{d}\gamma_{10} & \frac{\lambda}{d}\gamma_{10} & \dots & q + \frac{\lambda}{d}\gamma_{10} & \frac{\lambda}{d}\gamma_{10} & \dots & \frac{\lambda}{d}\gamma_{10} \\ \frac{\lambda}{d}\gamma_{11} & \frac{\lambda}{d}\gamma_{11} & \dots & \frac{\lambda}{d}\gamma_{11} & q + \frac{\lambda}{d}\gamma_{11} & \dots & \frac{\lambda}{d}\gamma_{11} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\lambda}{d}\gamma_{NM} & \frac{\lambda}{d}\gamma_{NM} & \dots & \frac{\lambda}{d}\gamma_{NM} & \frac{\lambda}{d}\gamma_{NM} & \dots & q + \frac{\lambda}{d}\gamma_{NM} \end{pmatrix}$$

Note that if we require a threshold number (c) of uDNA viruses to start production, the leftmost c columns become zero.

All but 2 eigenvalues of this matrix is q or 0. The remaining two eigenvalues are:

$$\frac{1}{2} * \left( \frac{\lambda}{d}\beta_{su} + \frac{\lambda}{d}\beta_{si} + q \pm \sqrt{-4 * \frac{\lambda}{d}\beta_{su}q + \left( \frac{\lambda}{d}\beta_{su} + \frac{\lambda}{d}\beta_{si} + q \right)^2} \right) - a_u$$

Where  $\beta_{su} = \sum_{r=1}^N \eta\gamma_{r,0}$  (zero integrated, r unintegrated,  $r \geq c$ ) and  $\beta_{si} = \sum_{r=0}^N \sum_{s=1}^N \gamma_{r,s}$ . This accounts for the reduction in transmission from uDNA-only cells.

These eigenvalues are exactly 0 when:

$$\frac{\lambda}{d}a_u\beta_i + \frac{\lambda}{d}a_i\beta_u = a_i a_u$$

And the new  $R_0$  is now:

$$R_0 \text{ Synaptic Only} : \frac{\beta_{si}\lambda}{a_i d} + \frac{\beta_{su}\lambda}{a_u d}$$

#### $R_0$ CALCULATION: COMBINED CELL FREE AND SYNAPTIC TRANSMISSION

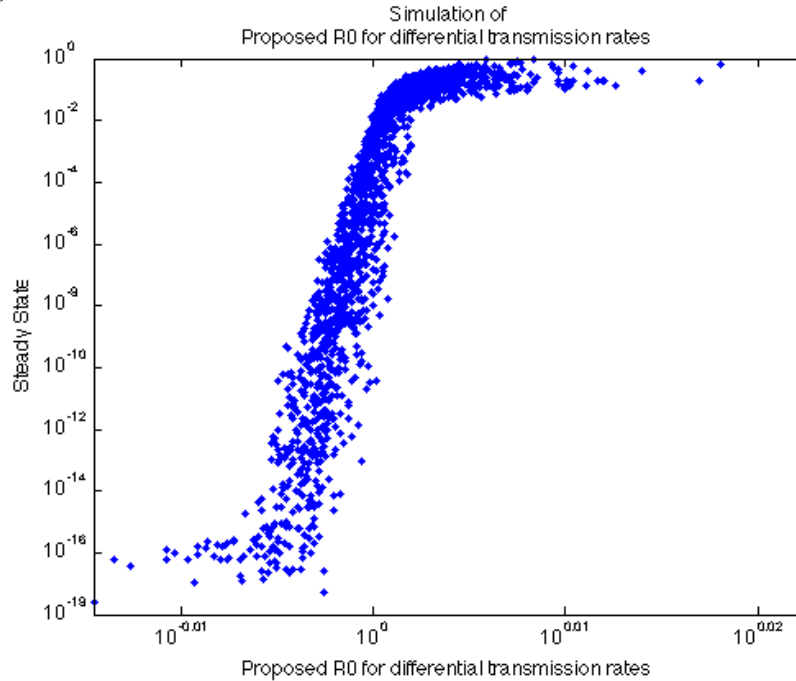
Although the analytical solution of  $R_0$  of the combined system of ODEs is unknown, numerical simulations (See figure) show that the critical value is:

$$R_0 \text{ Complete Model, No Threshold} : \frac{p\beta\lambda k_i}{ua_i d} + \frac{(1-p)\beta\lambda k_u}{ua_u d} + \frac{\beta_{si}\lambda}{a_i d} + \frac{\beta_{su}\lambda}{a_u d}$$

If more than one integrated viruses are required to begin infection, this becomes:

$$R_0 \text{ Complete Model, Threshold} : \frac{p\beta\lambda k_i}{ua_i d} + \frac{\beta_{si}\lambda}{a_i d} + \frac{\beta_{su}\lambda}{a_u d}$$

Figure 1: Simulation of the combined system show  $R_0$  approximately as predicted analytically - all parameters randomized. As  $R_0$  approaches 1 from below, the system requires more time to approach equilibrium. There is a sharp transition between zero steady states and non-zero steady states when the proposed  $R_0$  is above 1.



### EXPANDING $\gamma_{rs}$

If we assume that viruses integrate independently, then for any number of viruses transferred,  $h = r + s$ , the chance of any one of these viruses becoming integrated is  $p$  and the chance of any single one of these viruses becoming unintegrated is  $(1-p)$ . If we assume that these are independent events, then the resulting distribution of these  $h$  viruses follows a binomial distribution. In addition, each  $\gamma$  is can be decomposed to the probability to transfer  $h$  viruses and the probability that some portion of the  $h$  viruses become integrated.

Then we can define a set of transfer rates for any given number of cells:

$$\hat{\gamma}^h \text{ for } h = 1 \dots \max(N, M)$$

Note that for transfer from cells with only unintegrated viruses, the transfer rate is reduced by  $\eta$ . This defines the transfer rate of  $h$  viruses that become  $r$  unintegrated and  $s$  integrated viruses, where the transfer rates are:

$$\gamma_{r,0} = \hat{\gamma}^r (1-p)^r$$

$$\gamma_{r,s} = \hat{\gamma}^{r+s} \binom{r+s}{r} (1-p)^r p^s$$

The first represents r viruses being transferred and all r becoming unintegrated. The second represents a total of r+s viruses transferred through the synapse and s viruses becoming integrated while r viruses are unintegrated.

Then:

$$\beta_{su} = \eta \sum_{r=c}^N \hat{\gamma}^r (1-p)^r$$

$$\beta_{si} = \sum_{r=0}^N \sum_{s=1}^M \hat{\gamma}^{r+s} \binom{r+s}{r} (1-p)^r p^s$$

Which can be simplified to:

$$\beta_{su} = \eta \sum_{r=c}^N \hat{\gamma}^r (1-p)^r$$

$$\beta_{si} = \sum_{r=1}^N \hat{\gamma}^r (1 - (1-p)^r)$$

Therefore  $R_0$  of the model with synaptic transfer only is:

$$R_0 = \frac{\lambda}{a_i d} \sum_{r=1}^N \hat{\gamma}^r (1 - (1-p)^r) + \frac{\eta \lambda}{a_u d} \sum_{r=c}^N \hat{\gamma}^r (1-p)^r$$

And  $R_0$  of the complete model is:

$$R_0 = \frac{p\beta\lambda k_i}{ua_i d} + \frac{\lambda}{a_i d} \hat{\gamma}^r \sum_{r=1}^N (1 - (1-p)^r) + \frac{\eta \lambda}{a_u d} \hat{\gamma}^r \sum_{r=c}^N (1-p)^r$$

Finally, we explore different distributions for  $\hat{\gamma}^h$ , for  $h=1\dots N$ . It is unclear how many viruses are transferred per synapse, so we explore different distributions for the number of viruses transferred.

Uniform:  $\gamma^h$  is the same for all h

$$\text{Normal: } \gamma^h \sim \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{(h-\hat{h})^2}{2\sigma^2}}$$

$$\text{Poisson: } \gamma^h \sim \frac{\hat{h}^h \exp^{-\hat{h}}}{h!}$$



For example, for the poisson distribution, the final value for  $R_0$  becomes:

$$R_0 = \frac{p\beta\lambda k_i}{ua_id} + \frac{\lambda}{a_id} \frac{\hat{h}^r \exp^{-\hat{h}}}{r!} \sum_{r=1}^N (1 - (1-p)^r) + \frac{\eta\lambda}{a_ud} \frac{\hat{h}^r \exp^{-\hat{h}}}{r!} \sum_{r=c}^N (1-p)^r$$