The effect of defective viral genomes during respiratory syncytial virus infection

Zakarya Noffel and Hana M. Dobrovolny

Department of Physics and Astronomy, Texas Christian University, Fort Worth, USA

Background

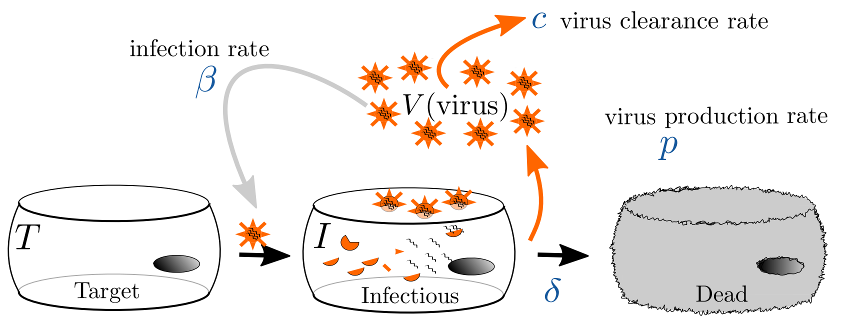
- Respiratory syncytial virus (RSV) can cause a severe respiratory illnesses particularly in young children and the elderly.
- Defective viral genomes (DVGs) have recently been found during RSV infections and are thought to be linked to the severity of the illness.
- While we derive our foundational understanding of the spread of RSV from experimental data, we may see helpful trends and correlations beyond the surface level through mathematical modeling.
- We used mathematical models to estimate parameter values so that we could pinpoint where DVGs were affecting the viral replication process.
- In this study, we present a number of examples of how ODE mathematical modeling has been used to help understand the spread of RSV under different scenarios of DVG appearance.

Mathematical model

We used a basic mathematical model of viral infection,

$$\begin{aligned}\frac{\mathrm{d}T}{\mathrm{d}t} &= -\beta TV\\ \frac{\mathrm{d}I}{\mathrm{d}t} &= \beta TV - \delta I\\ \frac{\mathrm{d}V}{\mathrm{d}t} &= pI - cV,\end{aligned}$$

- The model does not include an eclipse phase and assumes that cells produce and release virus immediately upon infection.
- T, I, and V are the variables of the model.
- β, δ,c, and p are the parameters.
- T represents the number of target cells infected with virus V at a rate β.
- Once infected, the cells enter the infectious phase I, in which they produce virus at a rate p, die at rate δ, and are cleared at rate c.



Methods

- To fit the data and calculate the most accurate parameter estimates, we used Python's built-in optimizing function scipy.optimize.minimize.
- We minimized the SSR

$$\mathrm{SSR} = \sum_{i=1}^{n} (y_i - y_m)^2$$

in order to find the parameters that would produce a curve closest to the actual data.

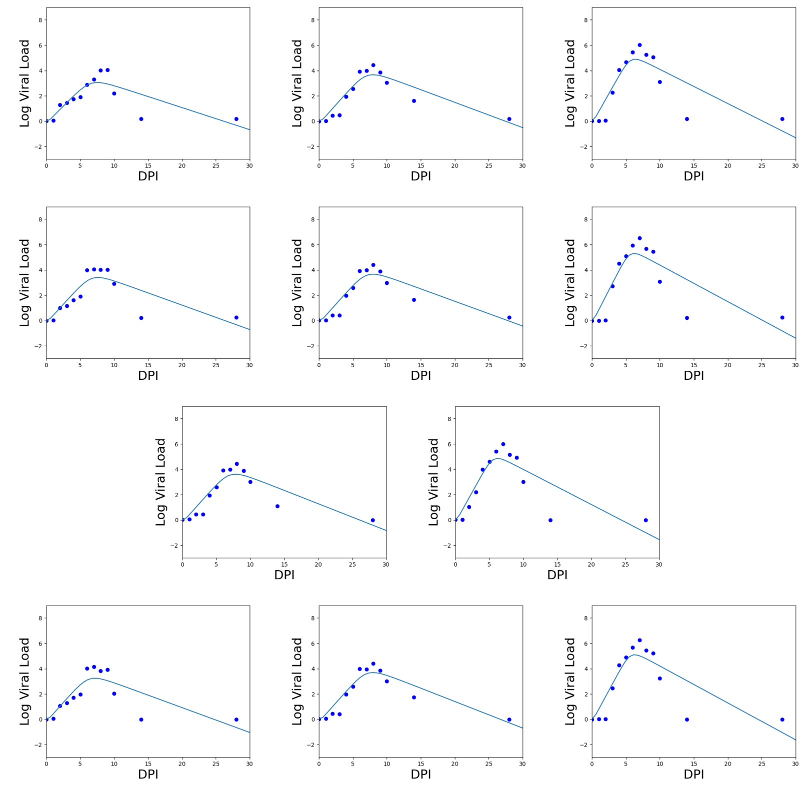
- We used bootstrapping to estimate the posterior distributions of our parameters.
- We shuffled the residuals at each given time point, then repeated the minimization process to produce another set of parameter estimates.
- After repeating this process 1000 times, we calculated the 95% confidence interval for each set of parameters.

Experimental data sets

We have two main data set from Felt et al. (2021) Nature Microbiol. consisting of a challenge study of RSV infections in healthy adults. The appearance of DVGs was then classified in different ways: early vs. late, prolonged vs. transient.

Fitting results

The experimental data and the best fit model predictions are shown below:



Parameter values

The best fit parameter estimates are given below:

1	β	p	δ	c	R_0	t_{inf}
Early	4.38e-05	5.06e+06	7.62	2.31e+01	1.26	0.095
95% CI	(-3.37, -5.97)	(14.8, 4.39)	(2.07e+07, 1.28)	(472, 1.64)	(0.00472, 2.95e-282)	(0.443, 3.88e-05)
DVG	2.93e-06	1.77e+09	2.69e+01	1.81e+02	1.07	0.020
95% CI	(-3.65, -6.25)	(15.0, 4.58)	(4.01e+08, 0.864)	(96.1, 0.614)	(0.074, 7.18)	(0.506, 1.96)
Late	2.21e-08	4.95e+12	2.26e+02	4.76e+02	1.01	0.00428
95% CI	(-5.73, -7.68)	(13.1, 8.49)	(408, 2.34)	(1.35e+03, 50.6)	(4.13e-46 0.0)	(0.0703, 0.003)

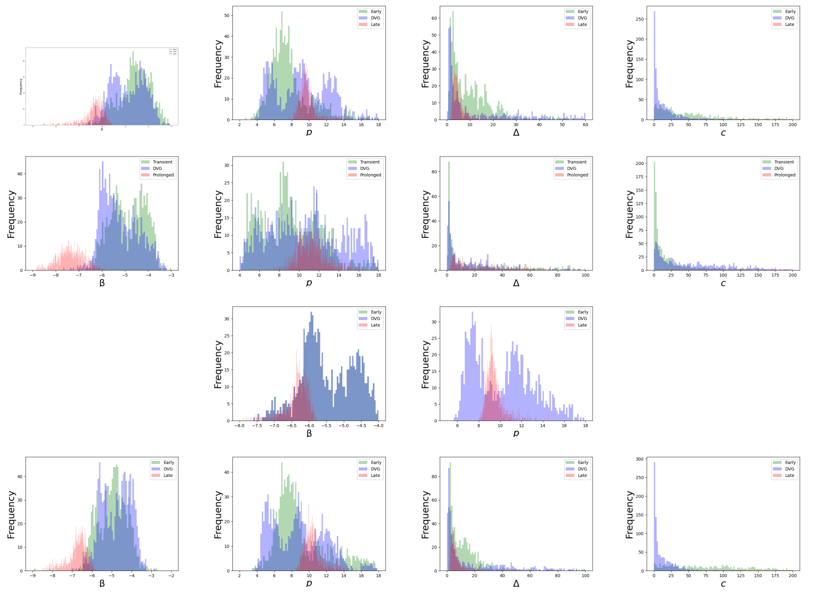
2	β	p	δ	c	R_0	t_{inf}
Transient	1.01e-05	1.78e+08	1.18e+01	1.31e+02	1.16	0.0334
95% CI	(-3.69, -6.18)	(16.0, 4.74)	(1.82e+09, 1.17)	(139, 0.961)	(0.0322, 2.24e-292)	(0.461, 7.22e-06)
DVG	2.55e-07	1.89e+12	3.87e+02	1.24e+03	1.01	0.002
95% CI	(-3.62, -6.42)	(17.2, 4.63)	(3.91e+09, 0.853)	(486, 1.61)	(0.0131, 0.0)	(0.455, 2.76e-06)
Prolonged	1.02e-08	2.18e+13	6.11e+01	3.46e+03	1.05	0.00300
95% CI	(-6.30, -8.60)	(14.9, 8.73)	(1.33e+03, 3.56)	(2.600e+03, 18.1)	(7.25e-20 0.0)	(0.0908, 0.00106)

3	β	p	δ	c	R_0	t_{inf}
95% CI	(-4.16, -7.02)	(-11.1, -11.1)	(1.00, 1.00)	(1.0, 1.0)	(6.11e-18, 8.44e-21)	(1.54e+09 5.72e+07)
DVG	3.10e-05	3.66e+07	2.41e+00	2.51e+02	1.88	0.0420
95% CI	(-4.16, -7.02)	(16.2, 6.41)	(5.58e+08 1.38)	(3.88e+03, 16.5)	(1.41e-24 0.0)	(0.117, 9.41e-06)
Late	3.27e-07	5.48e+09	9.01e+00	1.51e+02	1.32	0.0334
95% CI	(-5.91, -7.34)	(12.1, 8.42)	(159, 3.47)	(365, 31.2)	(4.28e-36, 7.36e-287)	(0.084, 0.00653)

4	β	p	δ	c	R_0	t_{inf}
Early	5.00e-06	2.19e+09	3.17e+01	3.26e+02	1.06	0.0135
95% CI	(-3.74, -6.29)	(16.8, 5.11)	(1.76e+09, 1.58)	(869, 2.70)	(0.000177, 1.90e-283)	(0.296, 4.10e-06)
DVG	2.20e-05	1.82e+07	3.25e+00	7.80e+01	1.58	0.0707
95% CI	(-3.60, -6.17)	(16.0, 4.54)	(1.84e+09, 0.771)	(92.3, 0.542)	(0.0670, 1.45e-262)	(0.521, 8.26e-06)
Late	1.68e-08	1.31e+13	6.90e+01	3.06e+03	1.04	0.00301
95% CI	(-6.04, -8.06)	(13.6, 8.96)	(476, 2.67)	(2.08e+03, 57.7)	(3.34e-51, 0.0)	(0.0562, 0.00257)

Parameter distributions

Using bootstrapping, we found distributions for the parameter estimates. We compared the distributions in each set of experiments to see if there were any dynamical differences in the time course of viral load that depended on the appearance of DVGs.



- In many cases, the posterior parameter distributions overlap indicating that there is no statistically significant difference between the parameter estimates for different cases.
- Some of the distributions, however appear to be separated and might be distinct for the different experimental groups.

Statistical comparison

We used a Mann-Whitney U test to check if parameter distributions were distinct. p values are given in the tables below. Statistically significant values are in bold.

1	β	p	δ	c	R_0	t_{inf}
Early/DVG	0.389	0.356	0.247	0.076	0.405	0.431
Early/Late	**0.0005**	0.098	0.365	**0.008**	0.440	0.269
DVG/Late	**0.001**	0.478	0.141	$\mathbf{5.82}\times10^{-5}$	0.436	0.475

2	β	p	δ	c	R_0	t_{inf}
Transient/DVG	0.371	0.411	0.510	0.129	0.469	0.443
Transient/Prolonged	$\mathbf{1.30}\times10^{-5}$	0.206	0.249	**0.001**	0.493	0.519
DVG/Prolonged	$\mathbf{2.37}\times10^{-5}$	0.428	0.227	0.034	0.343	0.382

3	β	p	δ	c	R_0	t_{inf}
DVG/Late	0.015	0.483	0.228	0.521	0.167	0.281

4	β	p	δ	c	R_0	t_{inf}
Early/DVG	0.441	0.488	0.514	**0.005**	0.426	0.476
Early/Late	**0.005**	0.238	0.365	0.024	0.468	0.446
DVG/Late	$\mathbf{7.44}\times10^{-5}$	0.216	0.197	$\mathbf{2.17}\times10^{-5}$	0.503	0.478

Conclusions

- Parameter differences show up primarily in the infection rate, β and the clearance rate c.
- When DVGs appeared late or were prolonged, the infection rate tended to be lower.
- When DVGs were early, clearance rate tended to be higher.

Future directions

- Create a mathematical model that includes DVGs to study how they affect the viral infection.

RSV can be an important cause of severe respiratory illness and is often found in infants, the elderly, or immunosuppressed patients. Here, we use mathematical models of infections within a host in conjunction with experiments to further our understanding of viruses. In studying the spread of RSV and its reproduction within target cells, this study aims to compare parameters governing viral spread among RSV infections with different amounts of defective viral genomes (DVGs). The final parameter distributions are plotted on the given histograms, which show that DVGs tend to affect infection rate and clearance rate.