

**Heterogeneous SARS-CoV-2 kinetics due to variable timing and intensity of immune responses
Supplementary Material**

Katherine Owens¹, Shadisadat Esmaeili¹, and Joshua T. Schiffer^{1,2*}

1. Fred Hutchinson Cancer Center, Vaccine and Infectious Diseases Division
2. University of Washington, Department of Medicine

*corresponding author:

Fred Hutchinson Cancer Center | 1100 Fairview Ave. N. | Seattle, WA 98109 | Mail Stop E4-100
206.667.7359

jschiffe@fredhutch.org

Table S1: Selecting the best model for fit to longitudinal viral load data by comparing AIC score. Initial conditions for all models were $S_0 = 1 \times 10^7$, $I_0 = 0$, $R_0 = 0$, and $E_0 = 0$. The initial viral inoculum was estimated from a logit distribution with lower bound of 0 and upper bound of 250 viral RNA copies/ml. The delay between infection and detection was estimated from a logit distribution with lower bound of 0 and upper bound of 20 days. The standard deviation of measurement error was fixed at $\sigma = 0.5 \log_{10}$ copies viral RNA/mL for all runs to ensure that variation in AIC scores came from differences in the ODE models. A subset of the full NBA data consisting of 266 well-documented infections was used for model comparison. An approximation of the AIC score was computed for each model by running the SAEM algorithm from six different initial parameter guesses and then taking the lowest score, indicated with bold and underline for each model. The best model overall was M7.

Model	AIC
<p>M1: Target-cell limited model</p> $\frac{dS}{dt} = -\beta SV \quad (\text{S1a})$ $\frac{dI_P}{dt} = \beta SV - \delta I_P \quad (\text{S1b})$ $\frac{dV}{dt} = \pi I_P - \gamma V \quad (\text{S1c})$	<p>16102, <u>16032</u>, 16063, 16069, 16042, 16049</p>
<p>M2: Target cell limited model with eclipse phase (TEIV)</p> $\frac{dS}{dt} = -\beta SV \quad (\text{S2a})$ $\frac{dI_E}{dt} = \beta SV - kI_E \quad (\text{S2b})$ $\frac{dI_P}{dt} = kI_E - \delta I_P \quad (\text{S2c})$ $\frac{dV}{dt} = \pi I_P - \gamma V \quad (\text{S2d})$	<p>15716, 15706, 15697, <u>15647</u>, 15698, 15685</p>
<p>M3: TEIV model with time-dependent clearance of infected cells</p> $\frac{dS}{dt} = -\beta SV \quad (\text{S3a})$ $\frac{dI_E}{dt} = \beta SV - kI_E \quad (\text{S3b})$ $\frac{dI_P}{dt} = kI_E - m(t)I_P \quad (\text{S3c})$ $\frac{dV}{dt} = \pi I_P - \gamma V \quad (\text{S3d})$ <p style="text-align: center;">where $\begin{cases} m(t) = 0 & t < \tau \\ m(t) = m & t \geq \tau. \end{cases}$</p>	<p>14823, 14792, 14846, 14799, <u>14782</u>, 14799</p>

<p>M4: TEIV model with refractory cell compartment (TREIV)</p> $\frac{dS}{dt} = -\beta SV - \phi I_P S + \rho R \quad (S4a)$ $\frac{dR}{dt} = \phi I_P S - \rho R \quad (S4b)$ $\frac{dI_E}{dt} = \beta SV - kI_E \quad (S4c)$ $\frac{dI_P}{dt} = kI_E - \delta I_P \quad (S4d)$ $\frac{dV}{dt} = \pi I_P - \gamma V \quad (S4e)$	<p>14865, 14879, 14861, 14888, 14879, 14862</p>
<p>M5: TIEV model with density dependent clearance of infected cells</p> $\frac{dS}{dt} = -\beta SV \quad (S5a)$ $\frac{dI_E}{dt} = \beta SV - kI_E \quad (S5b)$ $\frac{dI_P}{dt} = kI_E - \delta I_P^h I_P \quad (S5c)$ $\frac{dV}{dt} = \pi I_P - \gamma V \quad (S5d)$	<p>15680, 15652, 15736, 15629, 15730, 15669</p>
<p>M6: TEIV model with time and density-dependent clearance of infected cells</p> $\frac{dS}{dt} = -\beta SV \quad (S6a)$ $\frac{dI_E}{dt} = \beta SV - kI_E \quad (S6b)$ $\frac{dI_P}{dt} = kI_E - \delta I_P^h I_P - m(t)I_P \quad (S6c)$ $\frac{dV}{dt} = \pi I_P - \gamma V \quad (S6d)$ <p style="text-align: center;">where $\begin{cases} m(t) = 0 & t < \tau \\ m(t) = m & t \geq \tau. \end{cases}$</p>	<p>14376, 14360, 14433, 14589, 14534, 14570</p>
<p>M7: TREIV model with time-dependent clearance of infected cells</p> $\frac{dS}{dt} = -\beta SV - \phi I_P S + \rho R \quad (S7a)$ $\frac{dR}{dt} = \phi I_P S - \rho R \quad (S7b)$	<p>13748, 13785, 13775, 13776, 13778, 13738*</p>

$\frac{dI_E}{dt} = \beta SV - kI_E \quad (S7c)$ $\frac{dI_P}{dt} = kI_E - \delta I_P - m(t)I_P \quad (S7d)$ $\frac{dV}{dt} = \pi I_P - \gamma V \quad (S7e)$ <p style="text-align: center;">where $\begin{cases} m(t) = 0 & t < \tau \\ m(t) = m & t \geq \tau. \end{cases}$</p>	
<p>M8: TREIV model with density-dependent clearance of infected cells</p> $\frac{dS}{dt} = -\beta SV - \phi I_P S + \rho R \quad (S8a)$ $\frac{dR}{dt} = \phi I_P S - \rho R \quad (S8b)$ $\frac{dI_E}{dt} = \beta SV - kI_E \quad (S8c)$ $\frac{dI_P}{dt} = kI_E - \delta I_P^h I_P \quad (S8d)$ $\frac{dV}{dt} = \pi I_P - \gamma V \quad (S8e)$	<p>14916, 15166, 15225, 15361, 15373, 15077</p>
<p>M9: TREIV model with time and density-dependent clearance of infected cells</p> $\frac{dS}{dt} = -\beta SV - \phi I_P S + \rho R \quad (S9a)$ $\frac{dR}{dt} = \phi I_P S - \rho R \quad (S9b)$ $\frac{dI_E}{dt} = \beta SV - kI_E \quad (S9c)$ $\frac{dI_P}{dt} = kI_E - \delta I_P^h I_P - m(t)I_P \quad (S9d)$ $\frac{dV}{dt} = \pi I_P - \gamma V \quad (S9e)$ <p style="text-align: center;">where $\begin{cases} m(t) = 0 & t < \tau \\ m(t) = m & t \geq \tau. \end{cases} \quad (S9f)$</p>	<p>13797, 13827, 13800, 13837, 13841, 13821</p>

Table S2: Incorporating covariates. To further improve model fits, we considered including covariates accounting for the strain of the infection, age group, the vaccination status or cumulative prior exposures status at the time of infection. The fitting conditions were the same as described in Table S1, and the model used was M7. The best set of co-variates included no dose / no record versus at least one vaccine dose, as well all non-omicron variants versus omicron variants.

Covariates	AIC scores across 5 runs	Notes
vaccination status, τ	13710, 13627 , 13787, 13849, 13794	Wald test τ – difference from boosted <ul style="list-style-type: none"> - 0 dose: 0.73, $p < 2.2e-16$ - 1 dose: -0.36, $p = 3.6e-1$ - 2 dose: -0.076, $p = 3.18e-1$ - No record: 0.44, $p = 6.77e-7$
0 dose v. >1 dose v. no record, τ	13700, 13677, 13665 , 13713, 13690	Wald test τ – difference from 0 doses <ul style="list-style-type: none"> - >1 doses: -0.8, $p < 2.2e-16$ - No record: -0.2, $p = 2.03e-2$ Next most likely covariate suggested is ϕ for lineage (ANOVA: 13.92, $p = 1.89e-8$)
0 dose/no record v. >1 dose, τ ; lineage, ϕ	13652, 13608 , 13658, 13665, 13634	Wald test τ – difference from 0 doses/no record <ul style="list-style-type: none"> - 1+ doses: -0.7, $p = 3.61e-5$ Wald test ϕ – difference from delta <ul style="list-style-type: none"> - Other: 0.46, $p = 2.04e-1$ - Omicron: 1.5, $p = 2.72e-5$ - Unknown: 2.23, $p = 2.92e-6$
0 v. >1 dose, τ ; other/delta v. omicron, ϕ	13666, 13637 , 13649, 13673, 13647	Wald test ϕ – difference from delta/other <ul style="list-style-type: none"> - Omicron: 1.08, $p = 1.02e-9$ - Unknown: 1.88, $p = 2.31e-7$ Next most likely covariate suggested is τ for other/delta v. omicron v. unknown (ANOVA: 6.23, $p = 2.26e-3$)
0 dose/no record v. >1 dose, τ ; other/delta v. omicron, τ, ϕ	13616, 13589* , 13601, 13619, 13593	Wald test—difference from delta/other <ul style="list-style-type: none"> - Omicron: -.54, $p = 1.92e-12$ - Unknown: -.6, $p = 8.91e-7$ Next most likely covariate suggested is m for 0 v. >1 dose (ANOVA: 3.21, $p = 7.42e-2$)
0 dose/no record v. >1 dose, τ, m ; other/delta v. omicron, τ, ϕ	13596 , 13603, 13617, 13610, 13633	Wald test—difference from 0 doses/no record <ul style="list-style-type: none"> - 1+ doses: 0.484, $p = 8.8e-2$

Table S3: Incorporating correlations. To further improve model fits, we used the best model from Table S2 and incorporated linear dependencies between correlated model parameters into the statistical model for the random effects. The Pearson correlation between the random effects for each pair of parameters was calculated, and the benefit of accounting for the most prominent correlations at the population level was tested by selecting the correlation groups indicated in parenthesis in column 2 in the monolix interface, re-running the SAEM algorithm five times, and comparing the resulting AIC scores.

Covariates	Correlated parameters	AIC				
		Run 1	Run 2	Run 3	Run 4	Run 5
0 v. >1 dose, τ ; other/delta v. omicron, τ, ϕ	none	13616	<u>13589</u>	13601	13619	13593
0 v. >1 dose, τ ; other/delta v. omicron, τ, ϕ	(π, ϕ, δ), (τ, ρ)	13504	13499	<u>13478*</u>	13531	13541
0 v. > 1 dose, τ ; other/delta v. omicron, τ, ϕ	($\pi, \phi, \delta, \tau, \rho$)	13566	13540	13570	<u>13536</u>	13540
0 v. >1 dose, τ ; other/delta v. omicron, τ, ϕ	(π, ϕ, δ)	13553	13559	<u>13531</u>	13565	13564
0 v. >1 dose, τ ; other/delta v. omicron, τ, ϕ	(τ, ρ)	13556	13585	<u>13538</u>	13564	13568

Table S4: Population parameters from final model fits. Using M7 with fixed initial conditions $S_0 = 1 \times 10^7$, $I_0 = 0$, $R_0 = 0$, and $E_0 = 0$, and $V_0 = 97$ viral RNA copies/mL, and estimating the delay between infection and detection from a logit distribution with lower bound of 0 and upper bound of 20 days, we estimated model parameters to describe 1510 infections in the NBA cohort. We further included correlations between model parameters π , ϕ , δ , and τ , ρ . For the mixed-effect model, the standard deviation of measurement error was fixed at $a = 0.4 \log_{10}$ copies viral RNA/ml. The population parameters are recorded here and estimated individual parameters are available at <https://github.com/lacyk3/SARS-CoV-2Kinetics>. The corresponding model fits are included in Figure S7.

Parameter (unit)	Symbol	Population Mean	Standard Error	Standard deviation of random effects	Distribution Of random effects	Source
viral infectivity (\log_{10} (RNAcopies/mL) $^{-1}$ day $^{-1}$)	$\log_{10}\beta$	-7.31	3.17e-3	4.69e-2	normal	estimated
viral production rate (\log_{10} day $^{-1}$)	$\log_{10}\pi$	2.74	9.7e-3	0.316	normal	estimated
rate at which refractory cells revert to susceptible state (\log_{10} day $^{-1}$)	$\log_{10}\rho$	-1.62	3.33e-2	0.972	normal	estimated
rate constant for conversion of target cells to a refractory state (\log_{10} cell $^{-1}$ day $^{-1}$)	$\log_{10}\phi$	-5.29	6.52e-2	1.2	normal	estimated
Deviation from $\log_{10}\phi$ for delta/other	$\beta_{\phi_omicron}$	0.324	7.38e-2	--	--	--
Deviation from $\log_{10}\phi$ for delta/other	β_{ϕ_none}	1.53	0.134	--	--	--
infected cell clearance rate when $I = 1$ (day $^{-1}$ cells $^{-h}$)	δ	1.38	2.64e-2	0.565	log normal	estimated
onset of acquired immunity relative to detection (days)	τ	15.9	0.453	0.461	log normal	estimated
Deviation from τ for delta/other	$\beta_{\tau_omicron}$	-0.364	3.53e-2	--	--	--
Deviation from τ for delta/other	β_{τ_none}	-0.422	6.14e-2	--	--	--
Deviation from τ for unvaccinated/no record	$\beta_{\tau_>1\ vax}$	-0.141	3.16e-2	--	--	--
Increase in clearance rate of infected cells due to acquired immunity (day $^{-1}$)	m	16.4	0.542	0.502	log normal	estimated
delay between infection in nasal tissue and detection (days)	t_0	2.08	2.8e-2	0.900	logit [0,20]	estimated
viral clearance rate (day $^{-1}$)	γ	15	--	--	--	Goyal et al.
mean eclipse phase duration (days $^{-1}$)	$1/k$	1/4	--	--	--	Ke et al.

Agreement between data and model-based quantitative features.

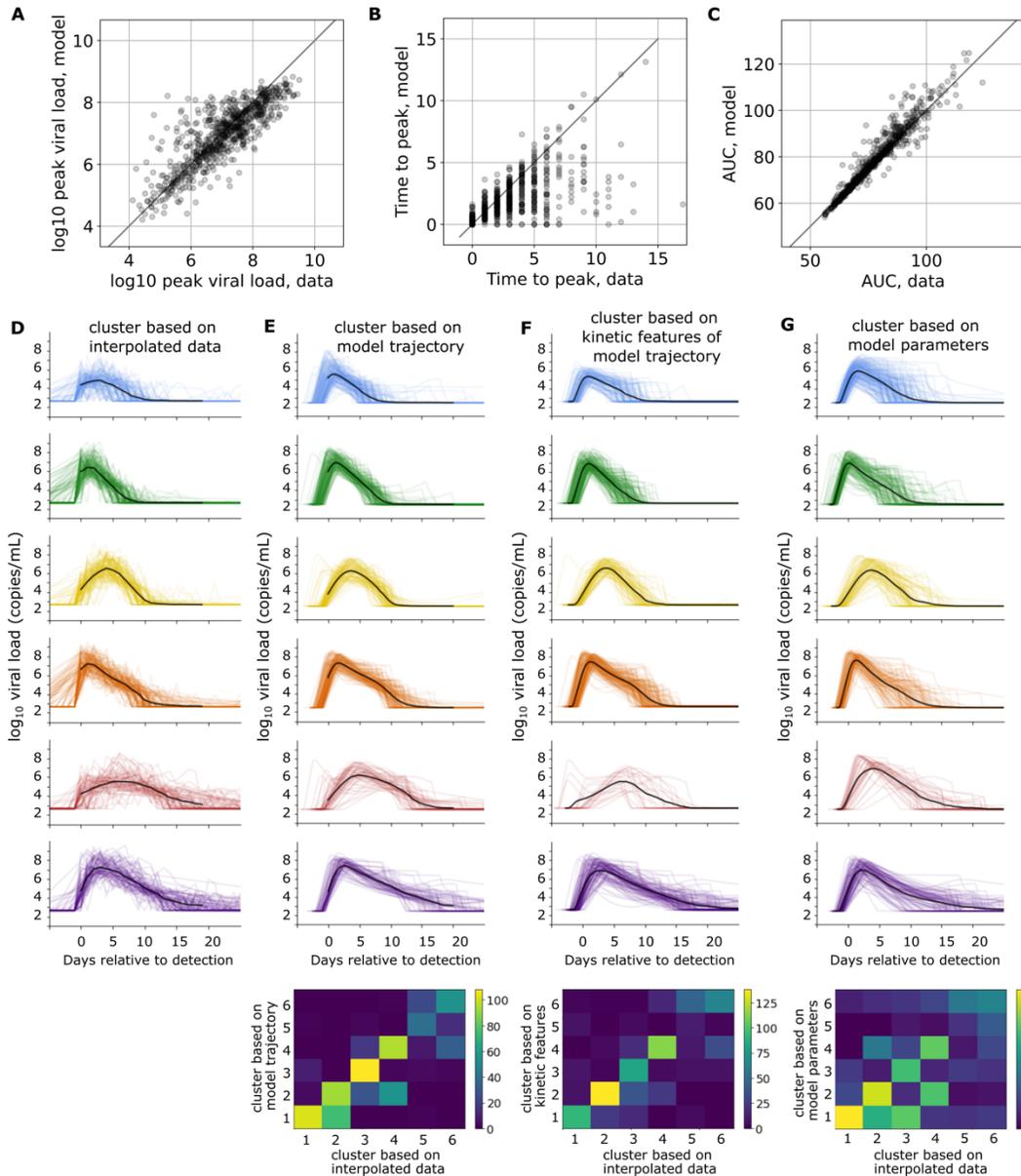


Figure S1: Model validation. We compared 3 quantities computed from the data with estimates from model simulation. (A) There is reasonable agreement between peak viral loads. (B) The time to peak estimated by the model is often shorter than what is estimated from the data alone, as many true infections lack early measurements. (C) AUC predicted by the model and AUC predicted by the data have strong agreement (when we use the linearly interpolated data and sample modeled viral load daily). We compared results using k -means clustering with $k=6$ based on (D) time-series data from the NBA cohort against three other possible feature sets from model output: (E) viral load trajectories predicted by the model simulations: most cluster centers closely match their counterparts in (D), though group 1 had slightly earlier peak viral loads; the results agreed with the time-series data based clustering for 72% of cases. (F) viral kinetic features (peak, time to peak, duration, AUC) extracted from each simulation: again, a similar set of 6 characteristic clusters emerged; however, group 5 was smaller; clustering results agreed with the data-based clusters (D) in only 54% of cases. (G) normalized estimated model parameters as features; in this case, clusters did not agree as well with the clustering of the cohort data in (D). The bottom row shows degree of agreement with each column and the clustered data in (D).

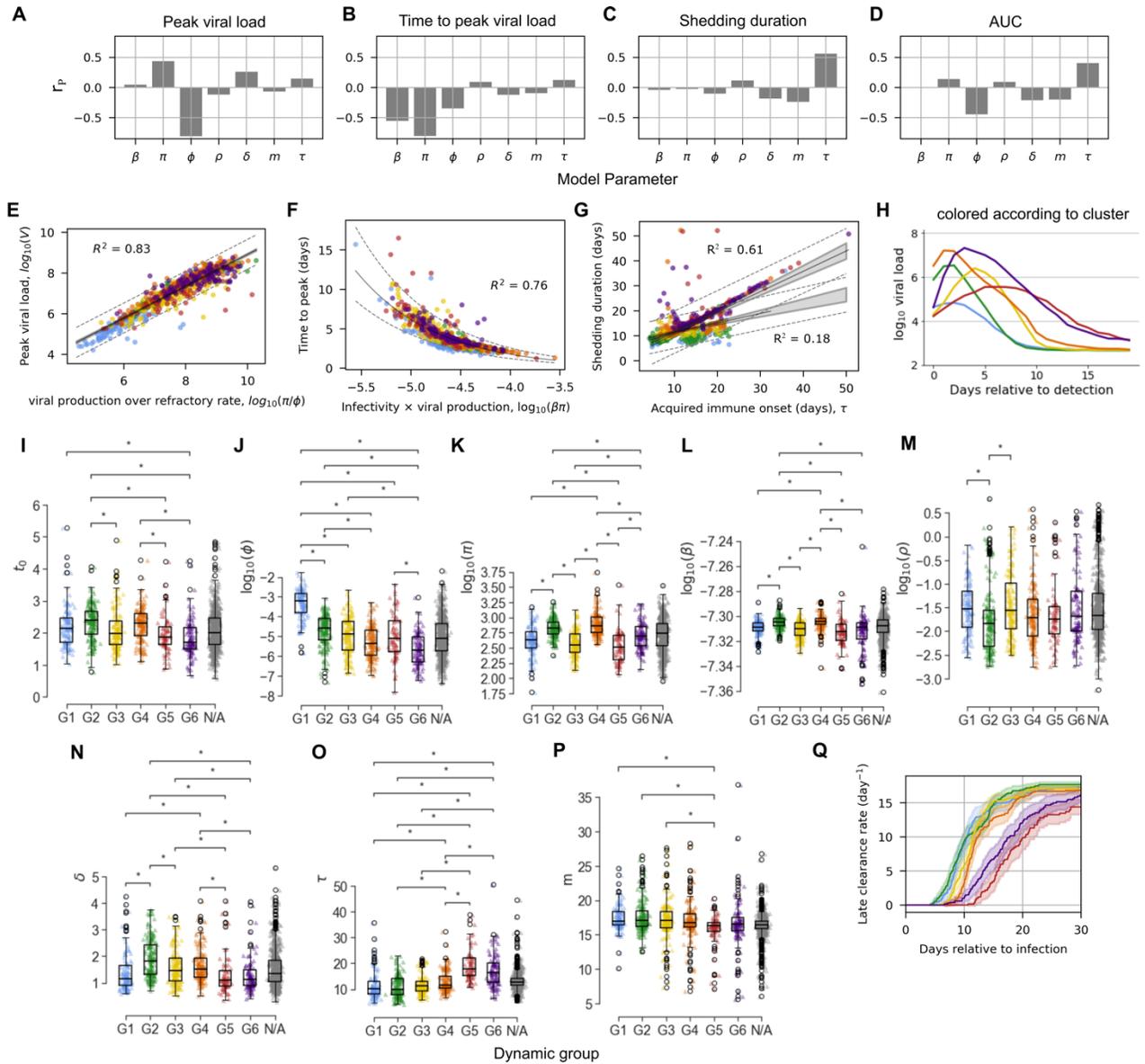


Figure S2: Differences between dynamic clusters. The first row shows the Pearson correlation coefficient, r_p , between model parameters and kinetic quantities (A) peak viral load, (B) time to peak viral load, (C) shedding duration and (D) area under the viral load curve. The second row shows how well combinations of parameters predict (E) peak viral load, (F) time to peak viral load, or (G) shedding duration. The coefficient of determination, R^2 , is indicated, the best fit line is shown in black with a bootstrapped 95% confidence interval shaded, and the prediction interval is shown with dotted lines. For shedding duration, the plot shows predictors calculated separately with groups 1, 5 and 6 yielding $R^2 = 0.61$ and groups 2-4 yielding $R^2 = 0.18$. Points are colored according to their dynamic group, the centers of which are reproduced in panel (H). Panels (I-P) show the distributions of estimated individual parameter values stratified by dynamic group. Here pairs of groups indicated by a star have significantly different mean parameter values as identified by the Mann-Whitney U-test with $p_{adjusted} < .05$, where $p_{adjusted}$ is scaled using the Bonferroni correction for multiple comparisons. Panel (Q) shows the average magnitude of the infected cell clearance rate due to delayed acquired immunity over time, stratified by dynamic group.

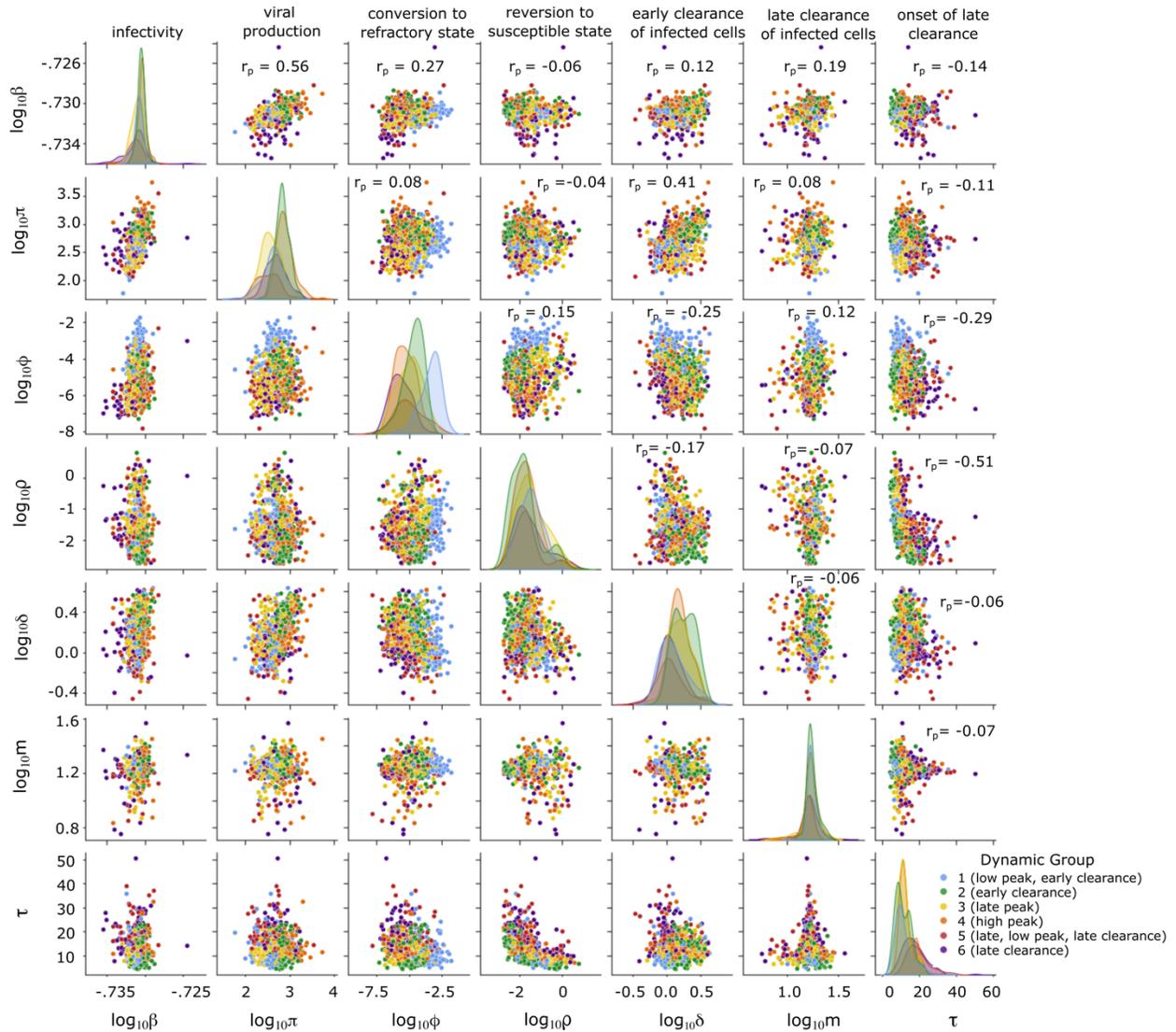


Figure S3: Distributions of model parameters and pair-wise correlations. Along the diagonal, the distribution of each parameter is shown for the six dynamic groups. Off-diagonal, for each pair of parameters, estimated individual parameter values are plotted with dots colored according to dynamic group. Panels are annotated with the Pearson correlation coefficient for that pair of parameters. The delay between infection and detection were excluded as it did not show strong correlations with other parameters.

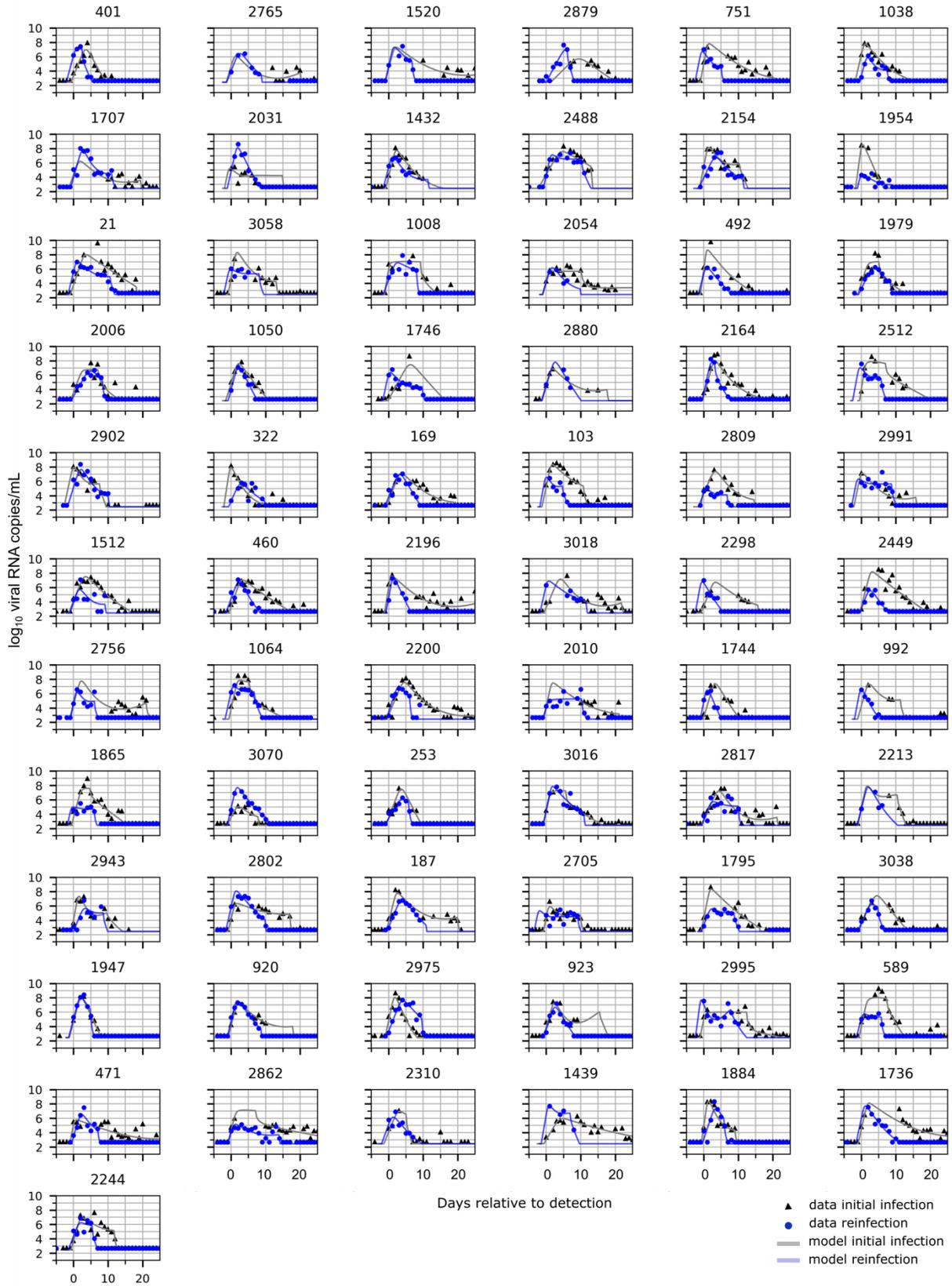


Figure S4: 67 cases of primary infection and reinfection in the same individual documented in the NBA cohort.

Tuning k-means clustering hyperparameters.

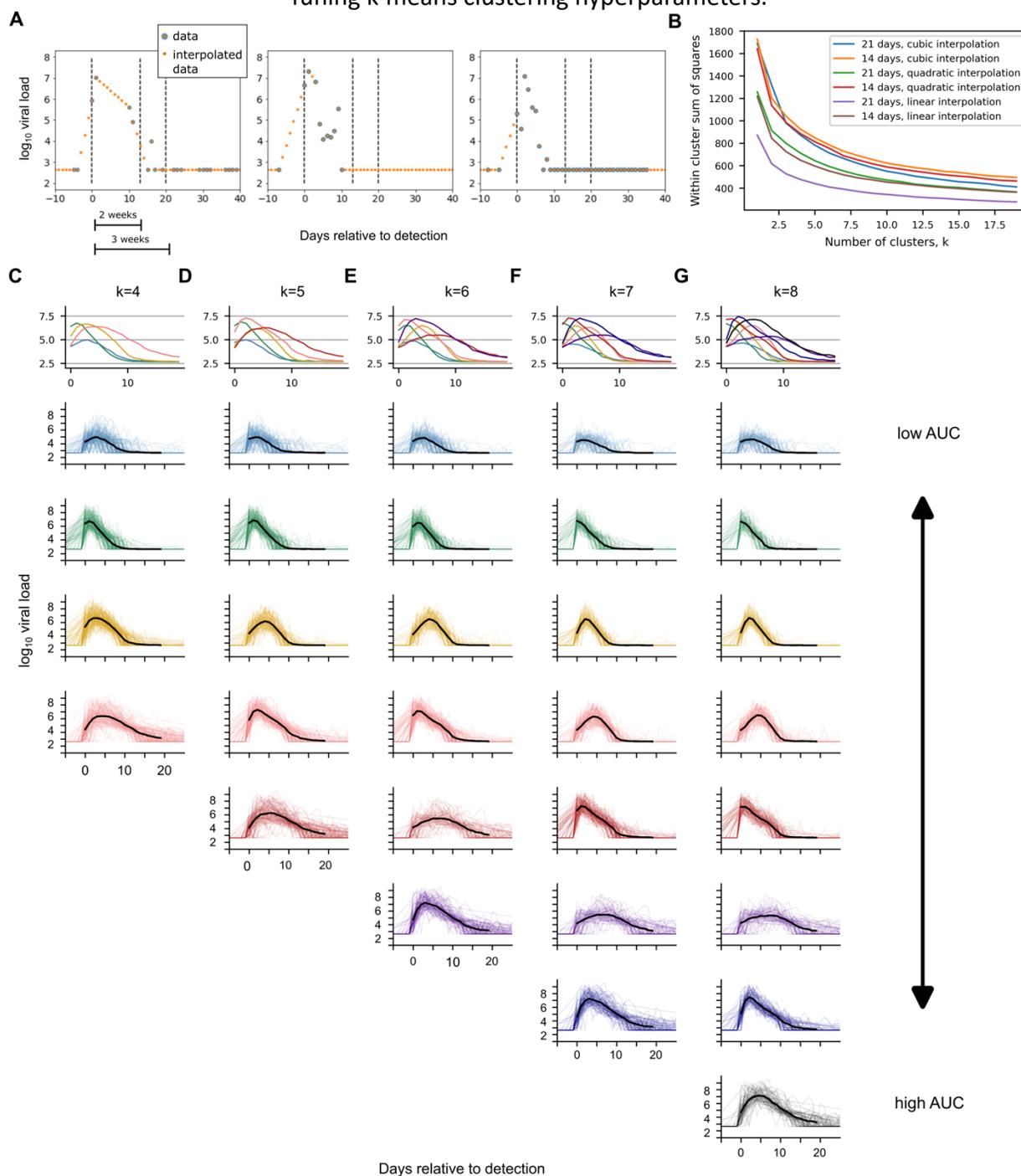


Figure S5: Examining various cluster sizes. We partitioned 768 infections from the NBA cohort that were documented over at least 3 weeks into subgroups with similar dynamics using k-means clustering. Before settling on using a 3-week documentation window and linear interpolation to impute any missing data points (A), we also considered a 2-week documentation window and interpolation using a quadratic or cubic spline (B). We examined the clustering results for $k = 4$ through $k = 8$, as this spanned the apparent “elbow” in the within-cluster sum of squares. The resulting clusters were ordered from low to high AUC for ease of comparison across scenarios (C-G).

Figure S6: Viral load data and predicted viral load for 1510 SARS-CoV-2 infections documented in the NBA cohort. Cases of rebound are indicated with an R in the panel title and are included first. The number prior to the underscore corresponds to the person ID assigned by Hay et al. (eLife 2022) while the number after the underscore indicates whether this is a first, second, 3rd infection etc. The color indicates which dynamic group this data was assigned to by k-means clustering, as color-coded in Figures 2. Infections plotted in black/gray were excluded from clustering.

