Densely sampled viral trajectories for SARS-CoV-2 variants alpha (B.1.1.7) and epsilon (B.1.429)

Stephen M. Kissler*[1], Joseph R. Fauver*[2], Christina Mack*[3,4], Caroline G. Tai[3], Mallery I. Breban[2], Anne E. Watkins[2], Radhika M. Samant[3], Deverick J. Anderson[5], David D. Ho[6], Jessica Metti[7], Gaurav Khullar[7], Rachel Baits[7], Matthew MacKay[7], Daisy Salgado[7], Tim Baker[7], Joel T. Dudley[7], Christopher E. Mason[7], Nathan D. Grubaugh†[2], Yonatan H. Grad†[1]

[1] Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA
[2] Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT
[3] IQVIA, Real World Solutions, Durham, NC
[4] Department of Epidemiology, University of North Carolina-Chapel Hill, Chapel Hill, NC
[5] Duke Center for Antimicrobial Stewardship and Infection Prevention, Durham, NC
[6] Aaron Diamond AIDS Research Center, Columbia University Vagelos College of Physicians and Surgeons, New York, NY
[7] TEMPUS, Chicago, IL

* denotes equal contribution
† denotes co-senior authorship

Correspondence and requests for materials should be addressed to:
Email: ygrad@hsph.harvard.edu
Telephone: 617.432.2275

Abstract.

We estimated (1) the time from first detectable virus to peak viral concentration (proliferation time), (2) the time from peak viral concentration to initial return to the limit of detection (clearance time), and (3) the peak viral concentration separately for 69 individuals infected with SARS-CoV-2 variants alpha (n=14), epsilon (n=10), and non-variants of interest/variants of concern (VOI/VOCs) (n=45). For individuals infected with alpha, the mean duration of the proliferation phase was 4.3 days (95% credible interval [3.0, 6.0]), the mean duration of the clearance phase was 7.4 days [6.0, 9.0], and the mean overall duration of infection (proliferation plus clearance) was 11.7 days [9.8, 13.8]. For individuals infected with epsilon, the mean duration of the proliferation phase was 5.6 days [4.0, 7.4], the mean duration of the clearance phase was 8.6 days [6.7, 10.6], and the mean overall duration of infection was 14.2 days [11.8, 16.7]. These compare to a mean proliferation phase of 4.3 days [3.4, 5.4], a mean clearance phase of 6.9 days [6.0, 7.9], and a mean duration of infection of 11.3 days [10.1, 12.5] for non-VOI/VOC infections. The peak viral concentration was 20.9 Ct [18.3, 23.5] for alpha, 20.6 Ct [17.8, 23.5] for epsilon, and 20.8 Ct [19.1, 22.4] for VOI/VOCs. This converts to 7.9 log10 RNA copies/ml [7.2, 8.7] for alpha, 8.0 log10 RNA copies/ml [7.2, 8.8] for epsilon, and 8.0 log10 RNA copies/ml [7.5, 8.4] for non-VOI/VOCs. The distributions of individual-level means are suggestive of longer clearance times for the VOCs. The overall durations of acute infection varied widely across individuals, with individual posterior mean acute infection durations ranging from 5.5 - 16.1 days for alpha, 9.4 - 20.3 days for epsilon, and 4.8 - 17.6 days for non-VOI/VOCs. These data offer evidence that infections with SARS-CoV-2 variants alpha and epsilon feature broadly similar viral trajectories as infections with non-VOI/VOCs.

Main text.

The reasons for the enhanced transmissibility of SARS-CoV-2 variants of concern (VOCs)[1] are unclear. Of special interest are the VOCs B.1.1.7 (alpha) and B.1.429 (epsilon), which until April 2021 were the most prevalent VOCs in the United States.[1] Variant alpha features multiple mutations in the spike protein receptor binding domain[2] that may enhance ACE-2 binding,[3] thus increasing the efficiency of virus transmission. Variant epsilon partially evades neutralization by sera from convalescent patients and vaccine recipients.[4] In addition to, and perhaps due to, these attributes, the viral trajectories for infections with alpha or epsilon could feature a higher peak viral load or longer duration of carriage, both of which could increase transmissibility. To test whether acute infections with SARS-CoV-2 VOCs alpha or epsilon are associated with higher or more sustained nasopharyngeal viral concentrations relative to non-variants of interest and non-variants of concern (non-VOI/VOCs), we assessed longitudinal densely sampled PCR tests performed in a cohort of 69 individuals (**Supplementary Table 1**) infected with SARS-CoV-2 undergoing daily surveillance testing, including 14 infected with alpha and 10 infected with epsilon, with lineage assignments confirmed by whole genome sequencing.

We estimated (1) the time from first detectable virus to peak viral concentration (proliferation time), (2) the time from peak viral concentration to initial return to the limit of detection (clearance time), and (3) the peak viral concentration for each individual (**Supplementary Appendix**).[5] We estimated the means of these quantities separately for individuals infected with alpha, epsilon, and non-VOI/VOCs (**Figure 1; Supplementary Table 2**). For individuals infected with alpha, the mean duration of the proliferation phase was 4.3 days (95% credible interval [3.0, 6.0]), the mean duration of the clearance phase was 7.4 days [6.0, 9.0], and the mean overall duration of infection (proliferation plus clearance) was 11.7 days [9.8, 13.8]. For individuals infected with epsilon, the mean duration of the proliferation phase was 5.6 days [4.0, 7.4], the mean duration of the clearance phase was 8.6 days [6.7, 10.6], and the mean overall duration of infection was 14.2 days [11.8, 16.7]. These compare to a mean proliferation phase of 4.3 days [3.4, 5.4], a mean clearance phase of 6.9 days [6.0, 7.9], and a mean duration of infection of 11.3 days [10.1, 12.5] for non-VOI/VOC infections. The peak viral concentration was 20.9 Ct [18.3, 23.5] for alpha, 20.6 Ct [17.8, 23.5] for epsilon, and 20.8 Ct [19.1, 22.4] for VOI/VOCs. This converts to 7.9 $\log_{10}$ RNA copies/ml [7.2, 8.7] for alpha, 8.0 $\log_{10}$ RNA copies/ml [7.2, 8.8] for epsilon, and 8.0 $\log_{10}$ RNA copies/ml [7.5, 8.4] for non-VOI/VOCs. We found no evidence of variant-specific differences in the population means for these parameters, as assessed by overlapping 95% credible intervals, and no

differences in empirical distribution, as assessed by a Kolmogorov-Smirnov test[6] with significance threshold $\alpha = 0.05$. However, the distributions of individual-level means are suggestive of longer clearance times for the VOCs (**Figure 1C**). The overall durations of acute infection varied widely across individuals, with individual posterior mean acute infection durations ranging from 5.5 – 16.1 days for alpha, 9.4 – 20.3 days for epsilon, and 4.8 – 17.6 days for non-VOI/VOCs (**Supplementary Table 3**). Data and code are available online.[7]

These data offer evidence that infections with SARS-CoV-2 variants alpha and epsilon feature broadly similar viral trajectories as infections with non-VOI/VOCs. Our ability to detect differences in the trajectories was limited by small sample sizes and a high degree of interpersonal variation. The findings should be seen as preliminary, as they are based on fourteen alpha cases and ten epsilon cases. The possibility of an extended clearance time for SARS-CoV-2 VOCs merits further investigation; if borne out by additional data, a longer isolation period than the currently recommended 10 days after symptom onset[8] may be needed to effectively interrupt secondary infections by some VOCs. Collection of longitudinal PCR and test positivity data in larger and more diverse cohorts is urgently needed.

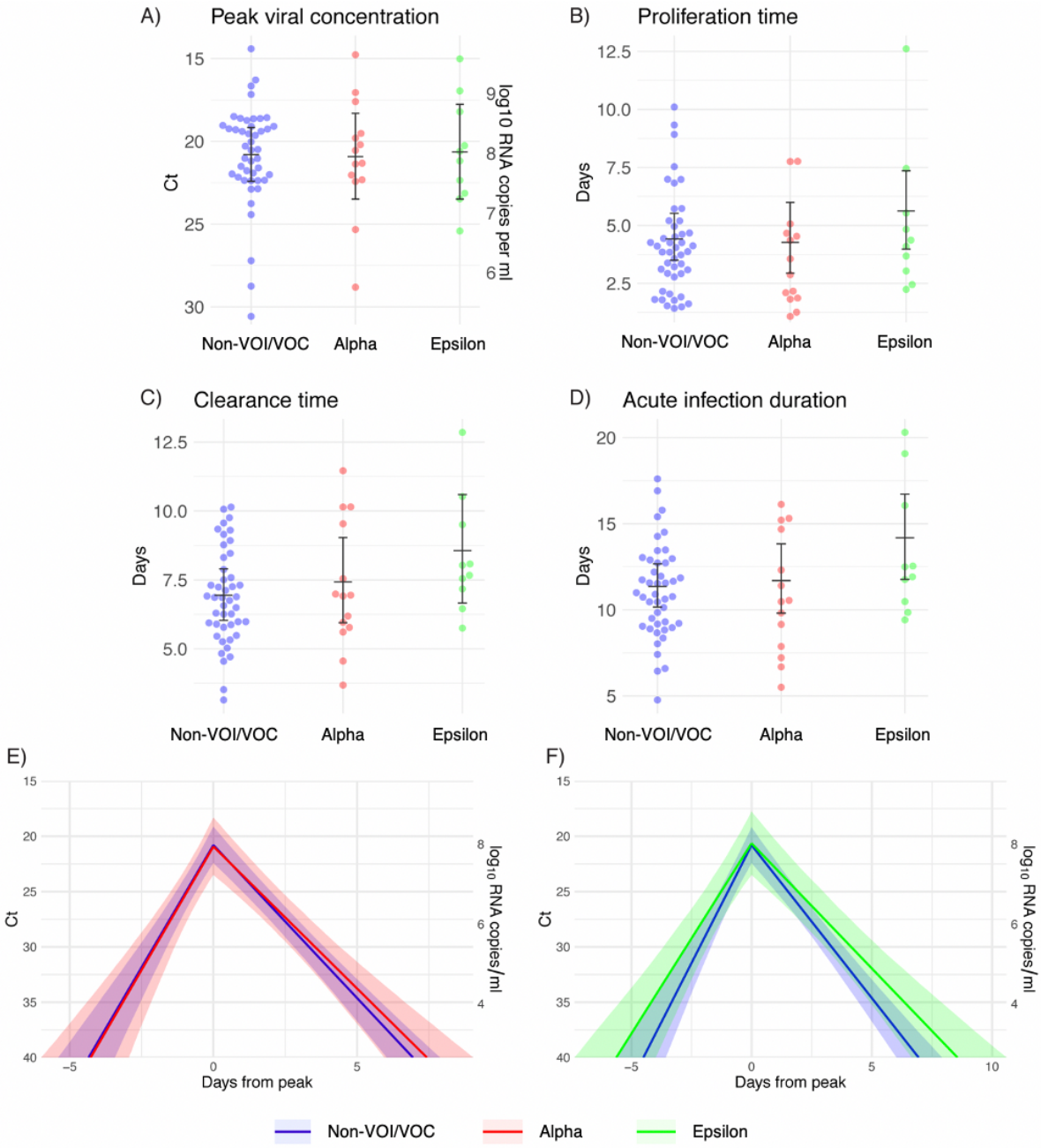


Figure 1. Estimated viral trajectory parameters for SARS-CoV-2 infections with alpha, epsilon, and non-variants of interest/non-variants of concern. Individual posterior means (points) with population means and 95% credible intervals (hatched lines) for (A) the mean peak viral concentration, (B) the mean proliferation duration, (C) the mean clearance duration, and (D) the mean total duration of acute infection. The points are jittered horizontally to avoid overlap. Solid lines in panels (E)-(F) depict the mean posterior viral trajectories for alpha infections (red) and epsilon infections (green) respectively, relative to non-VOI/VOC infections (blue), as specified by the population means and credible intervals in (A)-(D). The shaded regions represent 95% credible areas for the mean population trajectories.

References

1. Centers for Disease Control and Prevention. COVID Data Tracker. Published 2021. Accessed May 20, 2021. https://covid.cdc.gov/covid-data-tracker/
2. Galloway SE, Paul P, MacCannell DR, Johansson MA, Brooks JT, MacNeil A, et al. Emergence of SARS-CoV-2 B.1.1.7 Lineage — United States, December 29, 2020–January 12, 2021. *MMWR Morb Mortal Wkly Rep*. 2021;70(3):95-99. doi:10.15585/mmwr.mm7003e2
3. Yi C, Sun X, Ye J, Ding L, Liu M, Yang Z, et al. Key residues of the receptor binding motif in the spike protein of SARS-CoV-2 that interact with ACE2 and neutralizing antibodies. *Cell Mol Immunol*. 2020;17(6):621-630. doi:10.1038/s41423-020-0458-z
4. Deng X, Garcia-Knight MA, Khalid MM, Servellita V, Wang C, Morris MK, et al. Transmission, infectivity, and antibody neutralization of an emerging SARS-CoV-2 variant in California carrying a L452R spike protein mutation. *medRxiv*. Published online 2021.
5. Kissler SM, Fauver JR, Mack C, Olesen SW, Tai C, Shiue KY, et al. SARS-CoV-2 viral dynamics in acute infections. *medRxiv*. Published online 2020:1-13. doi:10.1101/2020.10.21.20217042
6. Conover WJ. *Practical Nonparametric Statistics*. John Wiley & Sons; 1971.
7. Kissler SM. Github Repository: CtTrajectories_B117. Published 2021. Accessed June 14, 2021. https://github.com/gradlab/CtTrajectories_AllVariants
8. Centers for Disease Control and Prevention. Duration of Isolation and Precautions for Adults with COVID-19. COVID-19. Published 2020. Accessed February 8, 2020. https://www.cdc.gov/coronavirus/2019-ncov/hcp/duration-isolation.html
9. United States Food and Drug Administration. *Emergency Use Authorization for TaqPath COVID-19 Combo Kit*.; 2020. https://www.fda.gov/media/136113/download
10. Loman N, Rowe W, Rambaut A. nCoV-2019 novel coronavirus bioinformatics protocol.
11. Illumina. *Illumina COVIDSeq Test Instructions for Use*.; 2021. https://www.fda.gov/media/138776/download
12. Illumina. NextSeq 550 System Documentation. Published 2021. Accessed June 10, 2021. https://support.illumina.com/sequencing/sequencing_instruments/nextseq-550/documentation.html
13. BaseSpace Labs. DRAGEN COVID Lineage. Published online 2021.
14. Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol*. 2020;5(11):1403-1407. doi:10.1038/s41564-020-0770-5
15. Aksamentov I, Neher R. NextClade. Published 2021. Accessed June 10, 2021. https://clades.nextstrain.org/
16. Kudo E, Israelow B, Vogels CBF, Lu P, Wyllie AL, Tokuyama M, et al. Detection of SARS-CoV-2 RNA by multiplex RT-qPCR. Sugden B, ed. *PLOS Biol*. 2020;18(10):e3000867. doi:10.1371/journal.pbio.3000867
17. Vogels C, Fauver J, Ott IM, Grubaugh N. *Generation of SARS-COV-2 RNA Transcript Standards for QRT-PCR Detection Assays*.; 2020. doi:10.17504/protocols.io.bdv6i69e
18. Cleary B, Hay JA, Blumenstiel B, Gabriel S, Regev A, Mina MJ. Efficient prevalence estimation and infected sample identification with group testing for SARS-CoV-2. *medRxiv*. Published online 2020.
19. Tom MR, Mina MJ. To Interpret the SARS-CoV-2 Test, Consider the Cycle Threshold Value. *Clin Infect Dis*. 2020;02115(Xx):1-3. doi:10.1093/cid/ciaa619
20. Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, et al. Stan : A Probabilistic Programming Language. *J Stat Softw*. 2017;76(1). doi:10.18637/jss.v076.i01

21. R Development Core Team R. R: A Language and Environment for Statistical Computing. Team RDC, ed. *R Found Stat Comput*. 2011;1(2.11.1):409. doi:10.1007/978-3-540-74686-7

Supplementary Appendix.

Ethics.

Residual de-identified viral transport media from anterior nares and oropharyngeal swabs collected from players, staff, vendors, and associated household members from a professional sports league were obtained from BioReference Laboratories. In accordance with the guidelines of the Yale Human Investigations Committee, this work with de-identified samples was approved for research not involving human subjects by the Yale Internal Review Board (HIC protocol # 2000028599). This project was designated exempt by the Harvard IRB (IRB20-1407).

Study population. The data reported here represent a convenience sample including team staff, players, arena staff, and other vendors (e.g., transportation, facilities maintenance, and food preparation) affiliated with a professional sports league. Clinical samples were obtained by combined swabs of the anterior nares and oropharynx administered by a trained provider. Viral concentration was measured using the cycle threshold (Ct) according to the Roche cobas target 1 assay. For an initial pool of 589 participants who first tested positive for SARS-CoV-2 infection during the study period (between November 28th, 2020 and May 4th, 2021), a diagnosis of "novel" or "persistent" infection was recorded. "Novel" denoted a likely new infection while "persistent" indicated the presence of virus in a clinically recovered individual. A total of 69 individuals (90% male) had novel infections that met our inclusion criteria: at least five positive PCR tests ($Ct < 40$), at least one negative PCR test ($Ct = 40$), at least one test with $Ct < 32$, and a genetic lineage of either B.1.1.7, B.1.429, or non-variant of interest/non-variant of concern (that is, we excluded lineages B.1.427, P.1, B.1.351, B.1.526, B.1.526.1, B.1.526.2, and P.2, of which there were 8 infections in total) as confirmed by whole genome sequencing. Fourteen of the individuals who met the inclusion criteria were infected with B.1.1.7 (alpha) and ten were infected with B.1.429 (epsilon).

Genome sequencing and lineage assignments

RNA was extracted and confirmed as SARS-CoV-2 positive by RT-qPCR with the Thermo Fisher TaqPath SARS-CoV-2 assay.[9] Next Generation Sequencing with the Illumina COVIDSeq ARTIC primer set[10] was used for viral amplification. Library preparation was performed using the amplicon-based Illumina COVIDseq Test v03[11] and sequenced 2x74 on Illumina NextSeq 550 following the protocol as described in Illumina's documentation.[12] The resulting FASTQs were

processed and analyzed on Illumina BaseSpace Labs using the Illumina DRAGEN COVID Lineage Application;[13] versions included are 3.5.0, 3.5.1, 3.5.2, and 3.5.3. The DRAGEN COVID Lineage pipeline was run with default parameters recommended by Illumina. Samples were considered SARS-COV-2 positive if at least 5 viral amplicon targets were detected at 20x coverage. Each SARS-COV-2 positive sample underwent lineage assignment and phylogenetics analysis using the most updated version of Pangolin[14] and NextClade,[15] respectively.

Converting Ct values to viral genome equivalents. To convert Ct values to viral genome equivalents, we first converted the Roche cobas target 1 Ct values to equivalent Ct values on a multiplexed version of the RT-qPCR assay from the US Centers for Disease Control and Prevention.[16] We did this following our previously described methods.[5] Briefly, we adjusted the Ct values using the best-fit linear regression between previously collected Roche cobas target 1 Ct values and CDC multiplex Ct values using the following regression equation:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad \text{(S1)}$$

Here, y_i denotes the i^{th} Ct value from the CDC multiplex assay, x_i denotes the i^{th} Ct value from the Roche cobas target 1 test, and ε_i is an error term with mean 0 and constant variance across all samples. The coefficient values are $\beta_0 = -6.25$ and $\beta_1 = 1.34$.

Ct values were fitted to a standard curve in order to convert Ct value data to RNA copies. Synthetic T7 RNA transcripts corresponding to a 1,363 b.p. segment of the SARS-CoV-2 nucleocapsid gene were serially diluted from 10^6-10^0 RNA copies/µl in duplicate to generate a standard curve[17] **(Supplementary Table 4)**. The average Ct value for each dilution was used to calculate the slope (-3.60971) and intercept (40.93733) of the linear regression of Ct on log-10 transformed standard RNA concentration, and Ct values from subsequent RT-qPCR runs were converted to RNA copies using the following equation:

$$\log_{10}([\text{RNA}]) = (Ct - 40.93733)/(-3.60971) + \log_{10}(250) \quad \text{(S2)}$$

Here, [RNA] represents the RNA copies /ml. The $\log_{10}(250)$ term accounts for the extraction (300 µl) and elution (75 µl) volumes associated with processing the clinical samples as well as the 1,000 µl/ml unit conversion.

Model fitting.

For the statistical analysis, we removed any sequences of 3 or more consecutive negative tests to avoid overfitting to these trivial values. Following our previously described methods,[5] we assumed that the viral concentration trajectories consisted of a proliferation phase, with exponential growth in viral RNA concentration, followed by a clearance phase characterized by exponential decay in viral RNA concentration.[18] Since Ct values are roughly proportional to the negative logarithm of viral concentration[19], this corresponds to a linear decrease in Ct followed by a linear increase. We therefore constructed a piecewise-linear regression model to estimate the peak Ct value, the time from infection onset to peak (*i.e.* the duration of the proliferation stage), and the time from peak to infection resolution (*i.e.* the duration of the clearance stage). The trajectory may be represented by the equation

$$E[Ct(t)] = \begin{cases} \text{l.o.d.} & t \leq t_o \\ \text{l.o.d.} - \frac{\delta}{t_p - t_o}(t - t_o) & t_o < t \leq t_p \\ \text{l.o.d.} - \delta + \frac{\delta}{t_r - t_p}(t - t_p) & t_p < t \leq t_r \\ \text{l.o.d.} & t > t_r \end{cases} \quad \text{(S3)}$$

Here, E[*Ct(t)*] represents the expected value of the Ct at time *t,* "l.o.d" represents the RT-qPCR limit of detection, δ is the absolute difference in Ct between the limit of detection and the peak (lowest) Ct, and t_o, t_p, and t_r are the onset, peak, and recovery times, respectively.

Before fitting, we re-parametrized the model using the following definitions:

- $\Delta Ct(t)$ = l.o.d. – $Ct(t)$ is the difference between the limit of detection and the observed Ct value at time *t.*
- $\omega_p = t_p - t_o$ is the duration of the proliferation stage.
- $\omega_r = t_r - t_p$ is the duration of the clearance stage.

We constrained $0.25 \leq \omega_p \leq 14$ days and $2 \leq \omega_r \leq 30$ days to prevent inferring unrealistically small or large values for these parameters for trajectories that were missing data prior to the peak and

after the peak, respectively. We also constrained $0 \leq \delta \leq 40$ as Ct values can only take values between 0 and the limit of detection (40).

We next assumed that the observed $\Delta Ct(t)$ could be described the following mixture model:

$$\Delta Ct(t) \sim \lambda \, \text{Normal}(E[\Delta Ct(t)], \sigma(t)) + (1 - \lambda) \, \text{Exponential}(\log(10))\Big]_{0}^{\text{l.o.d}} \quad \text{(S4)}$$

where E[$\Delta Ct(t)$] = l.o.d. - E[$Ct(t)$] and λ is the sensitivity of the q-PCR test, which we fixed at 0.99. The bracket term on the right-hand side of the equation denotes that the distribution was truncated to ensure Ct values between 0 and the limit of detection. This model captures the scenario where most observed Ct values are normally distributed around the expected trajectory with standard deviation $\sigma(t)$, yet there is a small (1%) probability of an exponentially distributed false negative near the limit of detection. The log(10) rate of the exponential distribution was chosen so that 90% of the mass of the distribution sat below 1 Ct unit and 99% of the distribution sat below 2 Ct units, ensuring that the distribution captures values distributed at or near the limit of detection. We did not estimate values for λ or the exponential rate because they were not of interest in this study; we simply needed to include them to account for some small probability mass that persisted near the limit of detection to allow for the possibility of false negatives.

We used a hierarchical structure to describe the distributions of ω_p, ω_r, and δ for each individual based on their respective population means $\mu_{\omega p}$, $\mu_{\omega r}$, and μ_δ and population standard deviations $\sigma_{\omega p}$, $\sigma_{\omega r}$, and σ_δ such that

$$\begin{aligned} \omega_p &\sim \text{Normal}(\mu_{\omega p}, \sigma_{\omega p}) \\ \omega_r &\sim \text{Normal}(\mu_{\omega r}, \sigma_{\omega r}) \\ \delta &\sim \text{Normal}(\mu_\delta, \sigma_\delta) \end{aligned} \quad \text{(S5)}$$

We inferred population means (μ.) separately for individuals infected with alpha, epsilon, and non-VOI/VOCs. We used a Hamiltonian Monte Carlo fitting procedure implemented in Stan (version 2.24)[20] and R (version 3.6.2)[21] to estimate the individual-level parameters ω_p, ω_r, δ, and t_p as well as the population-level parameters σ^*, $\mu_{\omega p}$, $\mu_{\omega r}$, μ_δ, $\sigma_{\omega p}$, $\sigma_{\omega r}$, and σ_δ. We used the following priors:

Hyperparameters:

$$\sigma^* \sim \text{Cauchy}(0, 5)\ [0, \infty] \tag{S6}$$

$$\begin{aligned} \mu_{\omega p} &\sim \text{Normal}(14/2, 14/6)\ [0.25, 14] \\ \mu_{\omega r} &\sim \text{Normal}(30/2, 30/6)\ [2, 30] \\ \mu_{\delta} &\sim \text{Normal}(40/2, 40/6)\ [0, 40] \end{aligned} \tag{S7}$$

$$\begin{aligned} \sigma_{\omega p} &\sim \text{Cauchy}(0, 14/\tan(\pi(0.95\text{-}0.5)))\ [0, \infty] \\ \sigma_{\omega r} &\sim \text{Cauchy}(0, 30/\tan(\pi(0.95\text{-}0.5)))\ [0, \infty] \\ \sigma_{\delta} &\sim \text{Cauchy}(0, 40/\tan(\pi(0.95\text{-}0.5)))\ [0, \infty] \end{aligned} \tag{S8}$$

Individual-level parameters:

$$\begin{aligned} \omega_p &\sim \text{Normal}(\mu_{\omega p}, \sigma_{\omega p})\ [0.25,14] \\ \omega_r &\sim \text{Normal}(\mu_{\omega r}, \sigma_{\omega r})\ [2,30] \\ \delta &\sim \text{Normal}(\mu_{\delta}, \sigma_{\delta})\ [0,40] \\ t_p &\sim \text{Normal}(0, 2) \end{aligned} \tag{S9}$$

The values in square brackets denote truncation bounds for the distributions. We chose a vague half-Cauchy prior with scale 5 for the observation variance σ^*. The priors for the population mean values ($\mu_\cdot$) are normally distributed priors spanning the range of allowable values for that parameter; this prior is vague but expresses a mild preference for values near the center of the allowable range. The priors for the population standard deviations ($\sigma_\cdot$) are half Cauchy-distributed with scale chosen so that 90% of the distribution sits below the maximum value for that parameter; this prior is vague but expresses a mild preference for standard deviations close to 0.

We ran four MCMC chains for 1,000 iterations each with a target average proposal acceptance probability of 0.8. The first half of each chain was discarded as the warm-up. The Gelman R-hat statistic was less than 1.1 for all parameters. This indicates good overall mixing of the chains. There were no divergent iterations, indicating good exploration of the parameter space. The posterior distributions for μ_δ, $\mu_{\omega p}$, and $\mu_{\omega r}$, were estimated separately for individuals infected with alpha, epsilon, and non-VOI/VOCs. These are depicted in **Figure 1** (main text). Draws from the

individual posterior viral trajectory distributions are depicted in **Supplementary Figures 1-2.** The mean posterior viral trajectories for each individual are depicted in **Supplementary Figure 3.**

Assessing sensitivity to different priors.

To ensure that our findings were not overly influenced by the prior distributions, we re-fit the model using two different sets of priors. The first set used the posterior population means from a previous study in a similar population as the prior values for $\mu_{\omega p}$, $\mu_{\omega r}$, and μ_{δ}., These priors were defined by

$$\begin{aligned} \mu_{\omega p} &\sim \text{Normal}(2.7, 14/6)\ [0.25, 14] \\ \mu_{\omega r} &\sim \text{Normal}(7.4, 30/6)\ [2, 30] \\ \mu_{\delta} &\sim \text{Normal}(20, 40/6)\ [0, 40]. \end{aligned} \tag{S10}$$

The second set used unrealistically low prior means for $\mu_{\omega p}$, $\mu_{\omega r}$, and μ_{δ} to verify that the suggestion of longer clearance times for alpha and epsilon infections was informed by the data and not solely by a biased prior distribution. These priors were defined by

$$\begin{aligned} \mu_{\omega p} &\sim \text{Normal}(0, 14/6)\ [0.25, 14] \\ \mu_{\omega r} &\sim \text{Normal}(0, 30/6)\ [2, 30] \\ \mu_{\delta} &\sim \text{Normal}(20, 40/6)\ [0, 40]. \end{aligned} \tag{S11}$$

Note that we updated the prior means but kept the prior variances at their original wide values to avoid encoding over-confidence in the priors into the model. The posterior population means for these new sets of priors are depicted in **Supplementary Figures 4-5** (compare to **Figure 1**). Overall, the findings were consistent across choices of prior.

	Alpha (%)	Epsilon (%)	Non-VOI/VOC (%)	Total (%)
Total	14 (20)	10 (14)	45 (65)	69 (100)
Age				
<18	0 (0)	0 (0)	0 (0)	0 (0)
18-29	12 (17)	4 (6)	26 (38)	42 (61)
30-39	2 (3)	3 (4)	9 (13)	14 (20)
40-49	0 (0)	2 (3)	3 (4)	5 (7)
50-59	0 (0)	1 (1)	4 (6)	5 (7)
≥60				3 (4)
Symptoms reported				
Yes	6 (9)	5 (7)	17 (25)	28 (41)
No	8 (12)	5 (7)	28 (41)	41 (59)

Supplementary Table 1. Characteristics of the study population. Number and percent (in parentheses) of individuals in the study population by age group and reported symptoms, stratified by variant.

	Minimum Ct	Maximum viral concentration (log_{10} RNA copies/ml)	Proliferation duration (days)	Clearance duration (days)	Acute infection duration (days)
Alpha	20.92 [18.31, 23.49]	7.94 [7.23, 8.67]	4.27 [2.95, 5.99]	7.43 [5.95, 9.03]	11.7 [9.80, 13.84]
Epsilon	20.64 [17.76, 23.49]	8.02 [7.23, 8.82]	5.62 [3.98, 7.36]	8.56 [6.66, 10.60]	14.19 [11.76, 16.71]
Non-VOI/VOC	20.80 [19.14, 22.42]	7.98 [7.53, 8.44]	4.34 [3.43, 5.39]	6.94 [6.03, 7.89]	11.28 [10.11, 12.54]

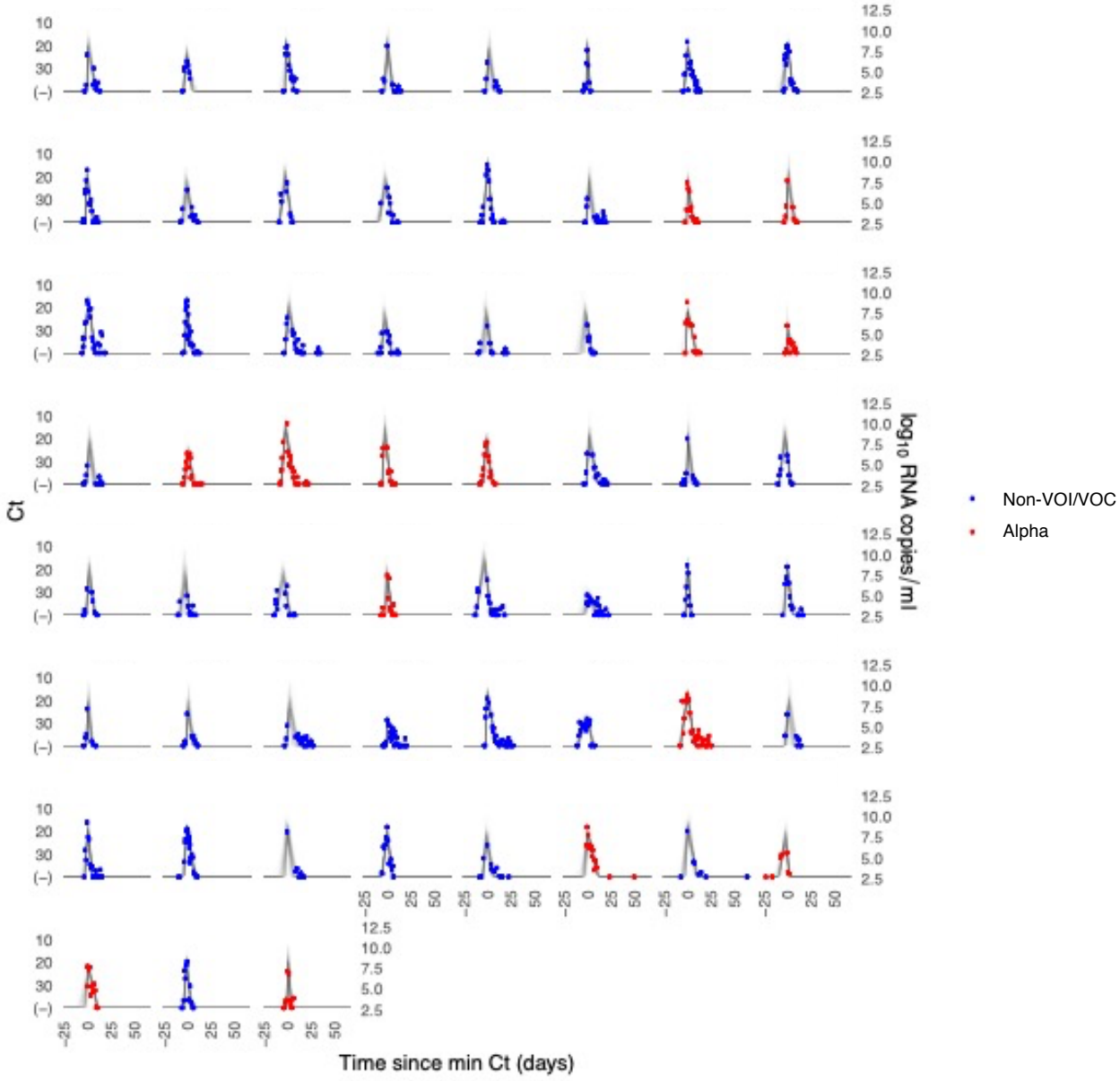
Supplementary Table 2. Posterior population viral trajectory parameters for SARS-CoV-2 infections with alpha, epsilon, and non-variants of interest/variants of concern. Reported values represent the posterior mean and 95% credible intervals (brackets) for each parameter.

	Minimum Ct	Maximum viral concentration (log_{10} RNA copies/ml)	Proliferation duration (days)	Clearance duration (days)	Acute infection duration (days)
Alpha					
Median	20.9	7.94	3.22	6.93	10.5
IQR	[19.6, 22.2]	[7.58, 8.31]	[1.93, 4.63]	[5.82, 9.04]	[8.19, 14.1]
Range	[14.8, 28.8]	[5.76, 9.65]	[1.08, 7.76]	[3.68, 11.5]	[5.50, 16.1]
Epsilon					
Median	20.9	7.95	4.23	7.85	12.2
IQR	[18.7, 22.9]	[7.38, 8.55]	[3.20, 5.36]	[7.27, 9.15]	[10.80, 15.2]
Range	[15.0, 25.4]	[6.70, 9.58]	[2.24, 12.60]	[5.75, 12.9]	[9.42, 20.3]
Non-VOI/VOC					
Median	20.5	8.06	3.87	6.86	10.9
IQR	[19.1, 22.2]	[7.60, 8.45]	[2.92, 4.96]	[5.87, 7.88]	[9.17, 12.7]
Range	[14.4, 30.6]	[5.27, 9.75]	[1.42, 10.10]	[3.14, 10.1]	[4.77, 17.6]

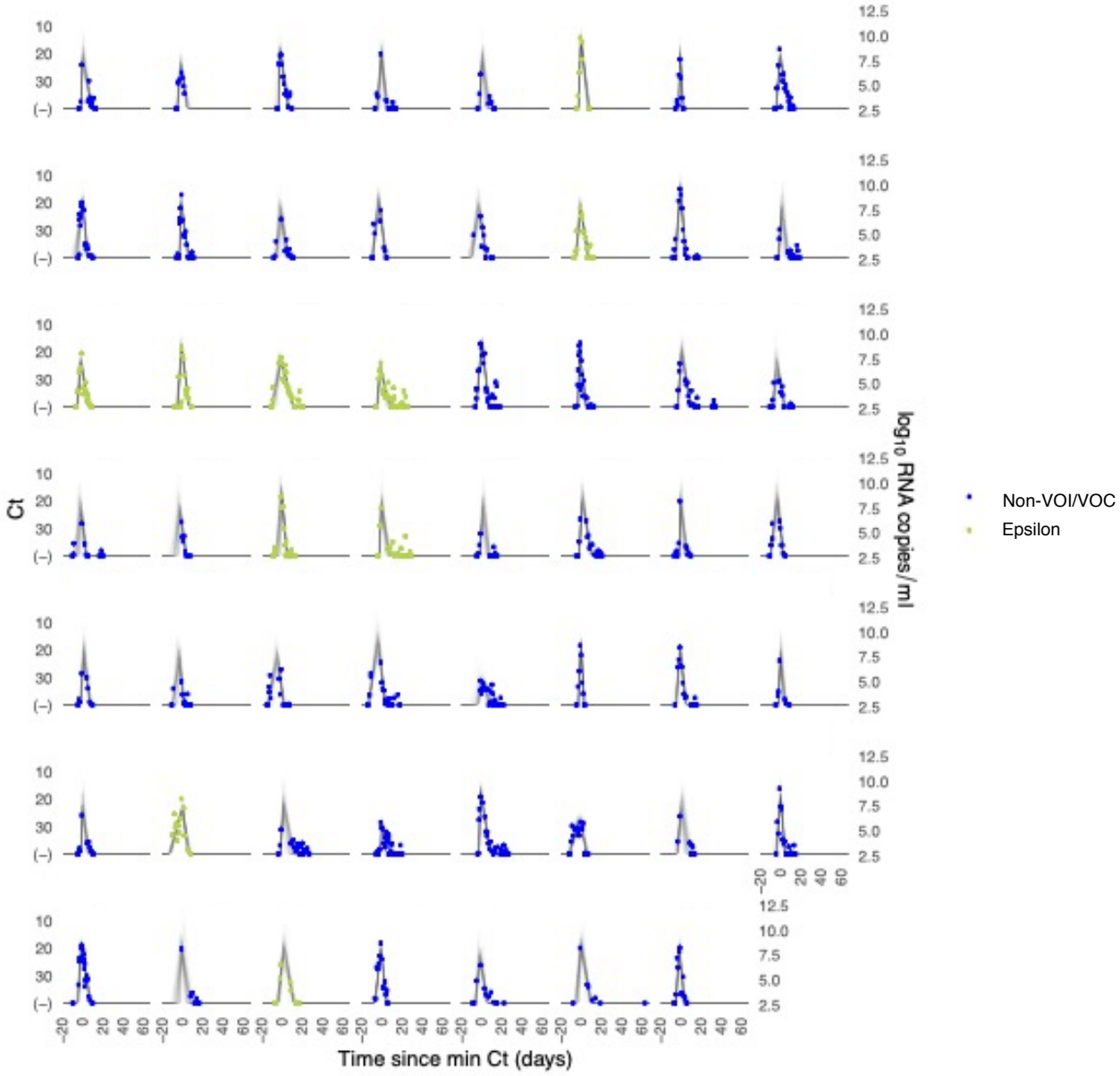
Supplementary Table 3. Summary of individual-level viral trajectory parameter means for SARS-CoV-2 infections with alpha, epsilon, and non-variants of interest/variants of concern. Reported values represent the median, inner quartile range, and full range of individual-level posterior means for each parameter.

Standard (copies/ul)	Replicate 1 (Ct)	Replicate 2 (Ct)	Average Ct
10^6	19.3	19.7	19.5
10^5	23.0	21.2	22.1
10^4	26.9	26.7	26.8
10^3	30.6	30.4	30.5
10^2	34.0	34.0	34.0
10^1	37.2	36.6	36.9
10^0	N/A	39.9	39.9

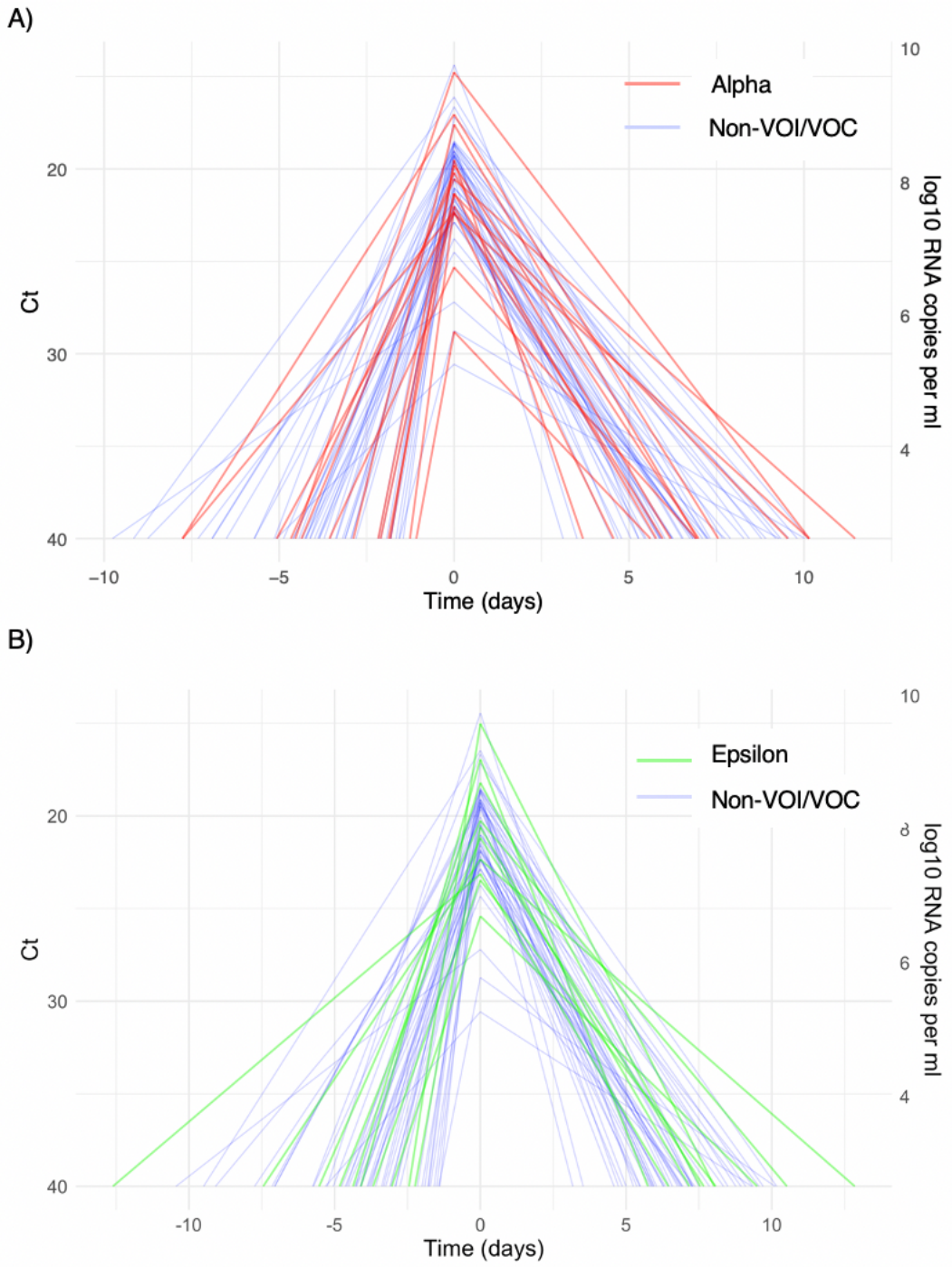
Supplementary Table 4. Standard curve relationship between virus RNA copies and Ct values. Synthetic T7 RNA transcripts corresponding to a 1,363 base pair segment of the SARS-CoV-2 nucleocapsid gene were serially diluted from 10^6-10^0 and evaluated in duplicate with RT-qPCR. The best-fit linear regression of the average Ct on the log10-transformed standard values had slope -3.60971 and intercept 40.93733 (R^2 = 0.99).



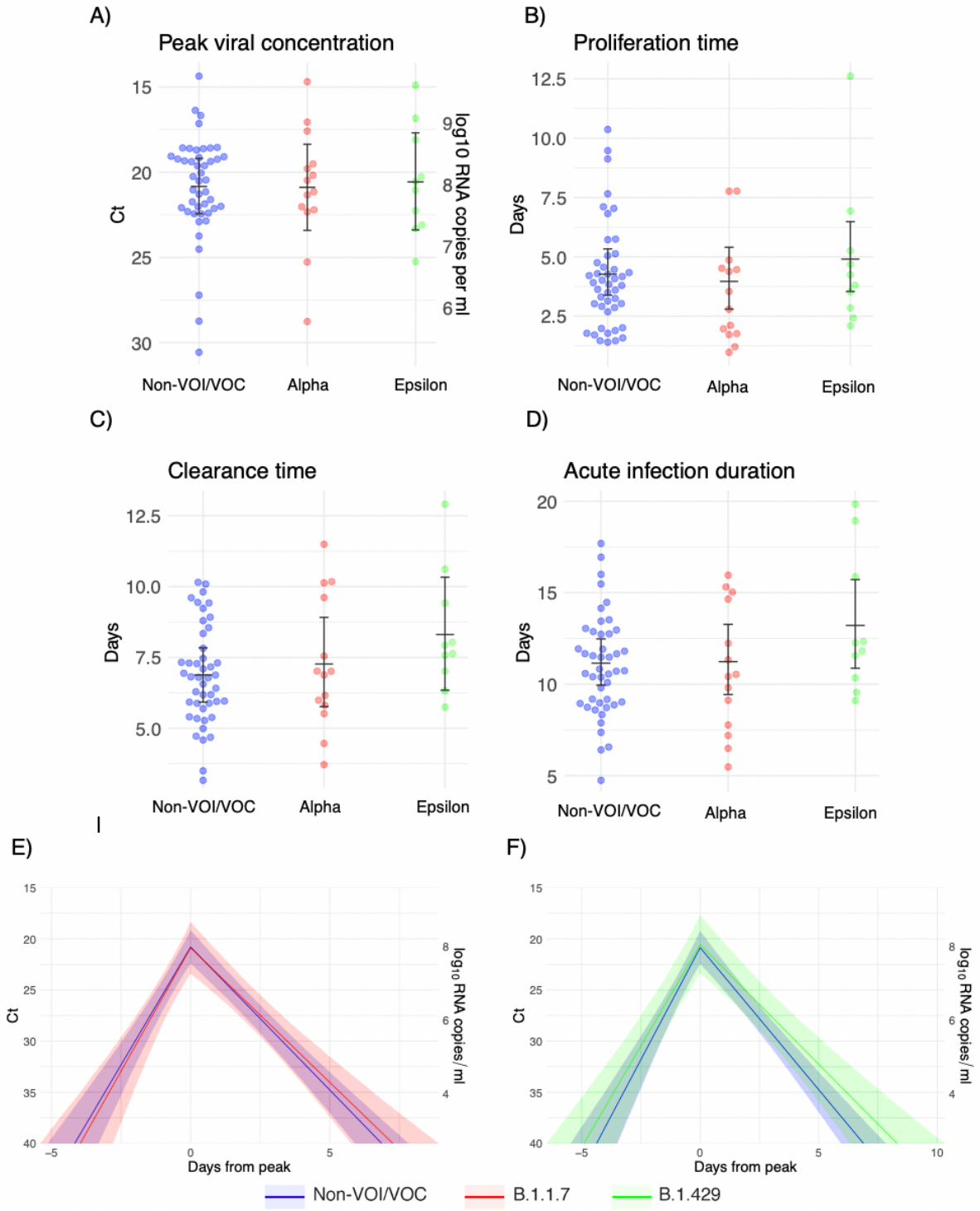
Supplementary Figure 1. Ct values and estimated trajectories for alpha variant and non-VOI/VOC SARS-CoV-2 infections. Each pane depicts the recorded Ct values (points) and derived log-10 genome equivalents per ml (log(ge/ml)) for a single person during the study period. Points along the horizontal axis represent negative tests. Time is indexed in days since the minimum recorded Ct value (maximum viral concentration). Individuals with confirmed alpha infections are depicted in red. Non-VOI/VOC infections are depicted in blue. Lines depict 100 draws from the posterior distribution for each person's viral trajectory.



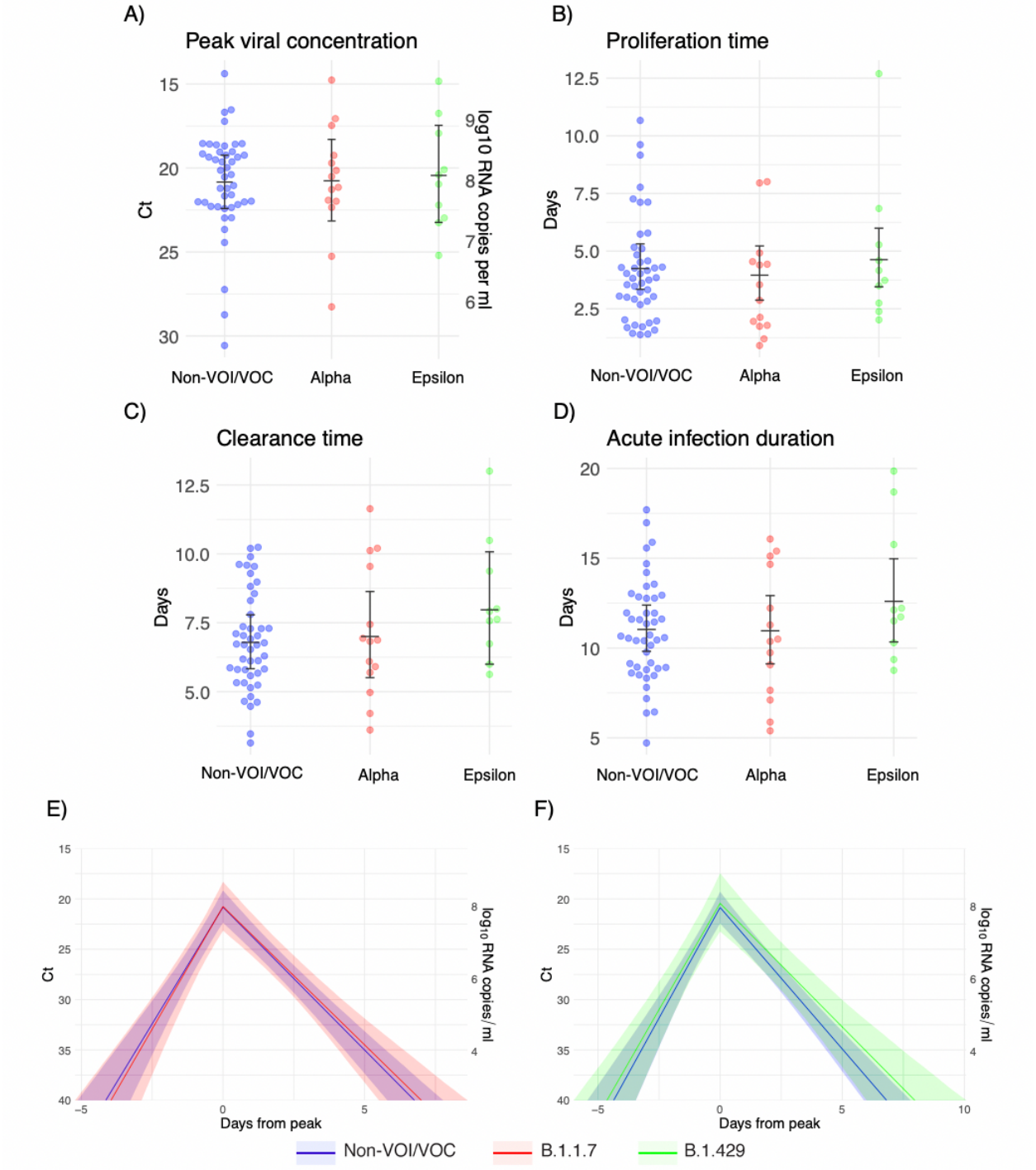
Supplementary Figure 2. Ct values and estimated trajectories for epsilon variant and non-VOI/VOC SARS-CoV-2 infections. Each pane depicts the recorded Ct values (points) and derived log-10 genome equivalents per ml (log(ge/ml)) for a single person during the study period. Points along the horizontal axis represent negative tests. Time is indexed in days since the minimum recorded Ct value (maximum viral concentration). Individuals with confirmed epsilon infections are depicted in red. Non-VOI/VOC infections are depicted in blue. Lines depict 100 draws from the posterior distribution for each person's viral trajectory.



Supplementary Figure 3. Mean posterior viral trajectories for each individual. Panel (A) depicts alpha infections (red) against non-VOI/VOC infections (blue). Panel (B) depicts epsilon infections (green) against non-VOI/VOC infections (blue). Trajectories are aligned temporally to have the same peak time.



Supplementary Figure 4. Estimated viral trajectory parameters for SARS-CoV-2 infections with alpha, epsilon, and non-variants of interest/non-variants of concern using informative priors. Individual posterior means (points) with population means and 95% credible intervals (hatched lines) for (A) the mean peak viral concentration, (B) the mean proliferation duration, (C) the mean clearance duration, and (D) the mean total duration of acute infection. Points are jittered horizontally to avoid overlap. Solid lines in panels (E)-(F) depict the mean posterior viral trajectories for alpha infections (red) and epsilon infections (green) respectively, relative to non-VOI/VOC infections (blue), as specified by the population means and credible intervals in (A)-(D). The shaded regions represent 95% credible areas for the mean population trajectories. Priors were informed by a previous analysis and are defined in Eq. (S10).



Supplementary Figure 5. Estimated viral trajectory parameters for SARS-CoV-2 infections with alpha, epsilon, and non-variants of interest/non-variants of concern using low priors. Individual posterior means (points) with population means and 95% credible intervals (hatched lines) for (A) the mean peak viral concentration, (B) the mean proliferation duration, (C) the mean clearance duration, and (D) the mean total duration of acute infection. Points are jittered horizontally to avoid overlap. Solid lines in panels (E)-(F) depict the mean posterior viral trajectories for alpha infections (red) and epsilon infections (green) respectively, relative to non-VOI/VOC infections (blue), as specified by the population means and credible intervals in (A)-(D). The shaded regions represent 95% credible areas for the mean population trajectories. Priors were chosen to be unrealistically low and are defined in Eq. (S11).