

The impact of immune history and variant on SARS-CoV-2 viral kinetics and infection rebound

Supplementary methods

Logistic regression models

Model fitting.

We fitted Bayesian logistic regression models for the probability of an individual having Ct value <30 on each day post detection using the *brms* package version 2.14.4. Models were run on the Harvard FAS Research Computing cluster using R version 4.0.2. For each model, we ran 4 chains for 2000 iterations each. Weakly informative priors (normal distributions with means of 0 and standard deviations of 10) were used for all model parameters. We assessed convergence based on all estimated parameters having a Gelman R-hat statistic less than 1.1.

Viral kinetics model

Statistical analysis.

Following previously described methods,^{1,2} we used a Bayesian hierarchical model to estimate the proliferation duration, clearance duration, and peak viral concentration for acute SARS-CoV-2 infections, stratified by variant (Omicron, Delta, Other) and immune status (vaccination history, including unexposed, 1-2 doses, or boosted; and antibody titer, including unexposed, titer ≤ 250 AU, and titer > 250 AU). The model describes the log₁₀ viral concentration during an acute infection using a continuous piecewise-linear curve with control points that specify the time of acute infection onset, the time and magnitude of peak viral concentration, and the time of acute infection clearance. The assumption of piecewise linearity is equivalent to assuming exponential viral growth during the proliferation period followed by exponential viral decay during the clearance period. The control points were inferred using the Hamiltonian Monte Carlo algorithm as implemented in Stan (version 2.24).³ We used priors informed by a previous analyses.^{1,2} Data and code are available online.

Model fitting.

To restrict to a set of well-observed acute infections for model fitting, we first removed any sequences of 3 or more consecutive negative tests (Ct = 40) from each acute infection to avoid overfitting to these trivial values. We kept only acute infections with at least one Ct value < 32 and at least 3 Ct values < 40 (the limit of detection).

We 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). This is represented by the equation

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

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:

- $\Delta Ct(t) = \text{l.o.d.} - Ct(t)$ 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 next characterized the likelihood of observing a given $\Delta Ct(t)$ using the following mixture model:

$$L(\Delta Ct(t) = x | \delta, t_p, \omega_p, \omega_r) = (1 - \lambda) \left[f_N(x | E[\Delta Ct(t)], \sigma(t)) + I_{lod} F_N(0 | E[\Delta Ct(t)], \sigma(t)) \right] + \lambda f_{Exp}(x | \kappa)$$

The left-hand side of the equation denotes the likelihood (L) that the observed viral load, as measured by Ct deviation from the limit of detection ($\Delta Ct(t)$), is equal to some quantity x given the model parameters δ (peak viral load), t_p (time of peak viral load), ω_p (proliferation time), and ω_r (clearance time). This likelihood is equal to the sum of two main components: the likelihood that the observed value was generated by the modeled viral kinetic process, denoted by the bracketed term preceded by a $(1-\lambda)$; and the likelihood that the observed value was a false negative, denoted by the term preceded by a λ . In the bracketed term representing the modeled viral kinetic process, $f_N(x | E[\Delta Ct(t)], \sigma(t))$ represents the Normal PDF evaluated at x with mean $E[\Delta Ct(t)]$ (generated by the model equations above) and observation noise $\sigma(t)$. $F_N(0 | E[\Delta Ct(t)], \sigma(t))$ is the Normal CDF evaluated at 0 with the same mean and standard deviation. This represents the scenario where the true viral load goes below the limit of detection, so that the observation sits at the limit of detection. I_{lod} is an indicator function that is 1 if $\Delta Ct(t) = 0$ and 0 otherwise; this way, the F_N term acts as a point mass concentrated at $\Delta Ct(t) = 0$. Last, $f_{Exp}(x | \kappa)$ is the Exponential PDF evaluated at x with rate κ . We set $\kappa = \log(10)$ 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. A schematic of the likelihood function is depicted in **Supplementary Figure 15**.

We used a hierarchical structure with a non-centered parameterization to describe the distributions of ω_p , ω_r , and δ for each person:

$$\begin{aligned} \omega_p[i] &= \text{Exp}[\mu_{\omega_p} + \zeta_{\omega_p}^i + \sigma_{\omega_p} N^i_{\omega_p}] \omega_p^* \\ \omega_r[i] &= \text{Exp}[\mu_{\omega_r} + \zeta_{\omega_r}^i + \sigma_{\omega_r} N^i_{\omega_r}] \omega_r^* \\ \delta[i] &= \text{Exp}[\mu_{\delta} + \zeta_{\delta}^i + \sigma_{\delta} N^i_{\delta}] \delta^* \end{aligned}$$

Here, ω_p^* , ω_r^* , and δ^* are user-defined estimated values for the means of ω_p , ω_r , and δ , so that the exponential terms represent an adjustment factor relative to that midpoint defined by μ (a shared adjustment factor for the entire population), ζ^i (an adjustment factor shared among individuals of a given variant/immune category), and σ (a shared standard deviation for the entire population). The N^i terms represent individual-level random effects. The prior distributions for the μ , ζ , and σ terms were all Normal(0, 0.25) (with σ truncated to have support on the positive reals). These prior distributions define LogNormal adjustment factors that have ~99% of their probability mass between 0.5 and 2, so that the prior distributions for ω_p , ω_r , and δ cover roughly half to twice their prior estimated midpoint values.

We used a Hamiltonian Monte Carlo fitting procedure implemented in Stan (version 2.24) and R (version 3.6.2) to estimate the parameters. We ran four MCMC chains for 2,000 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

indicated good overall mixing of the chains. There were no divergent iterations, indicating good exploration of the parameter space.

References

1. Kissler, S. M. *et al.* Viral dynamics of acute SARS-CoV-2 infection and applications to diagnostic and public health strategies. *PLoS Biol.* **19**, e3001333 (2021).
2. Kissler, S. M. *et al.* Viral Dynamics of SARS-CoV-2 Variants in Vaccinated and Unvaccinated Persons. *N. Engl. J. Med.* **385**, 2489–2491 (2021).
3. Carpenter, B. *et al.* Stan: A Probabilistic Programming Language. *J. Stat. Softw.* **76**, 1–32 (2017).

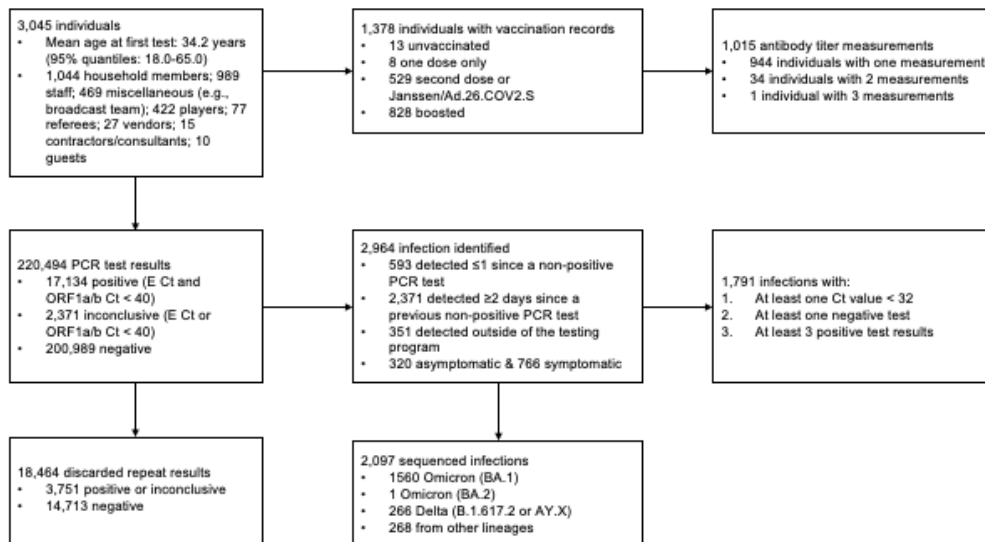


Figure S1. Summary of cohort. Top row describes cohort demographics and data on immune histories. Middle row describes infection data. Bottom row provides additional information on the infection data.

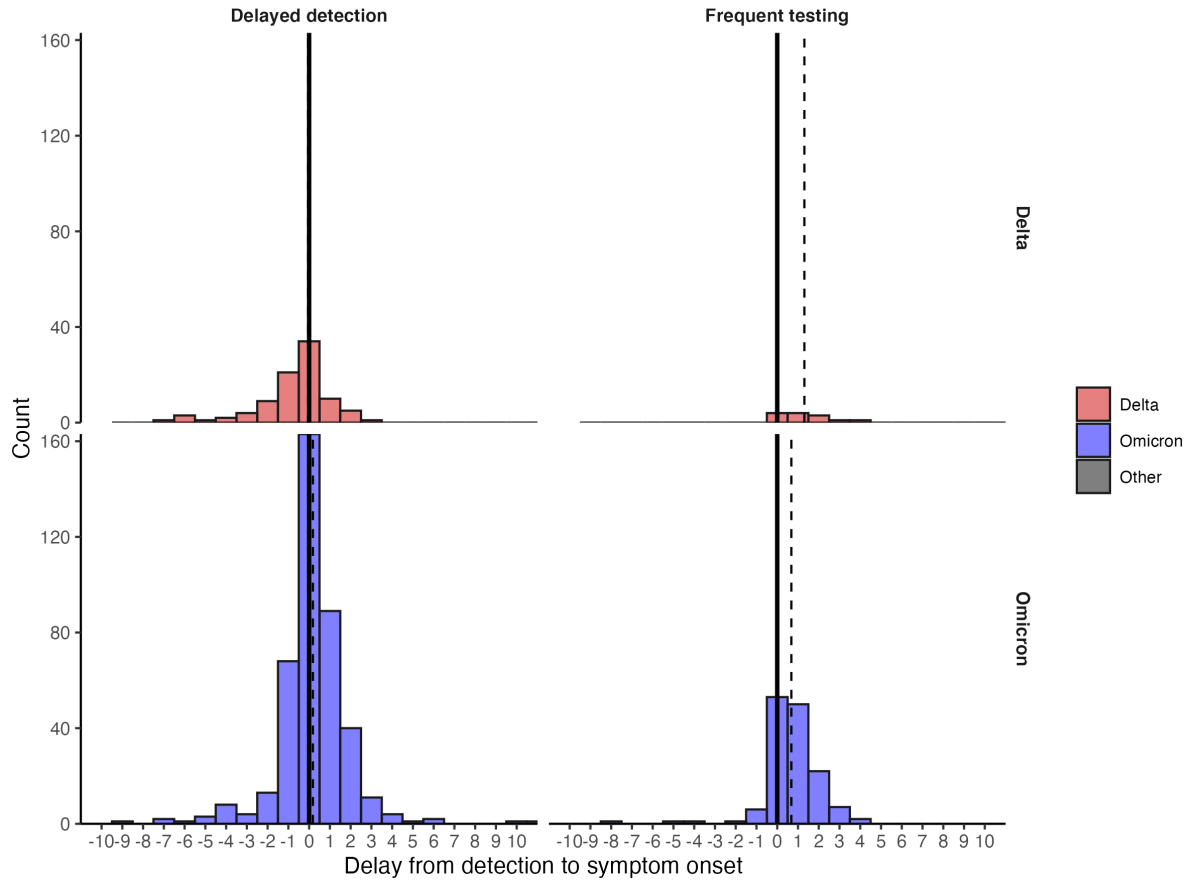


Figure S2. Distribution of delays from detection to symptom onset among individuals with known symptom status. Dashed lines mark the median delay between detection and symptom onset. Solid lines mark the day of detection (0).

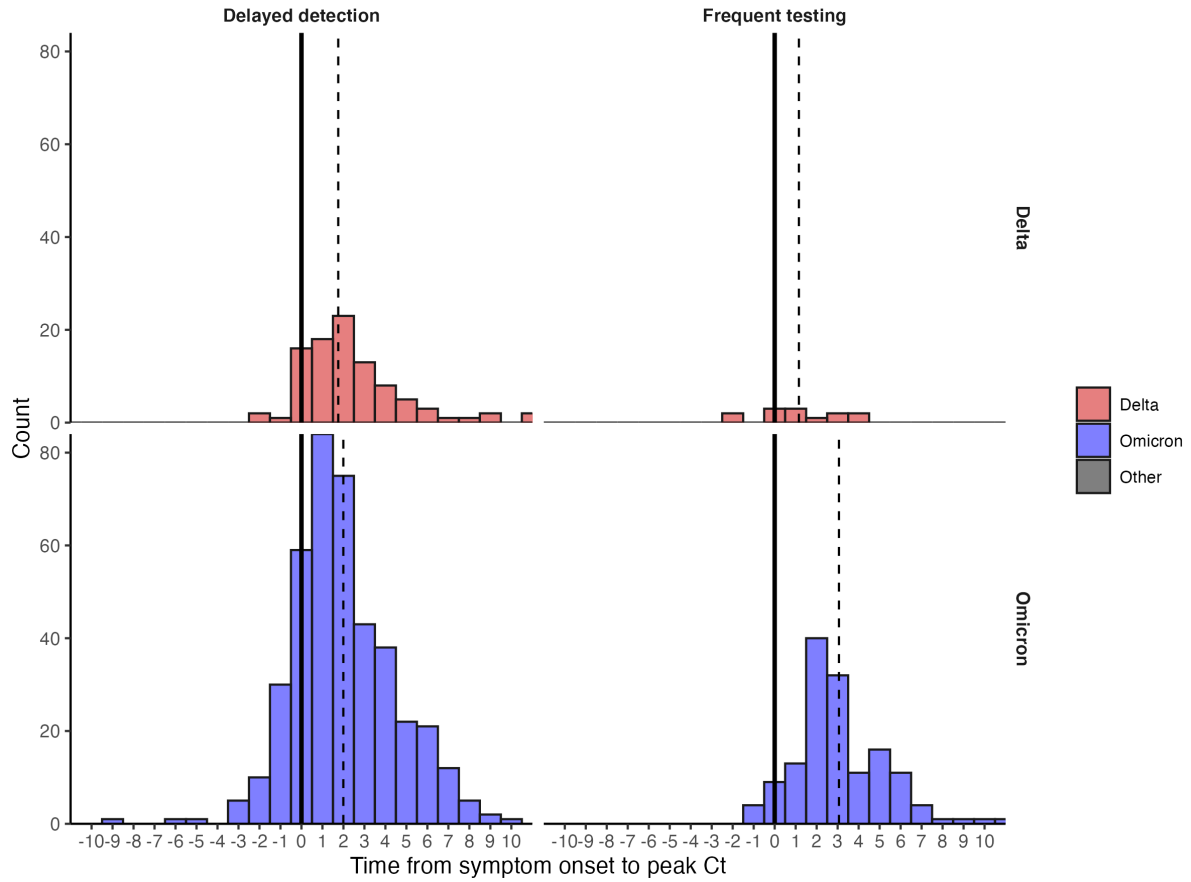


Figure S3. Distribution of delays from symptom onset to peak Ct values among individuals with known symptom status. Dashed lines mark the median delay between detection and symptom onset. Solid lines mark the day of detection (0).

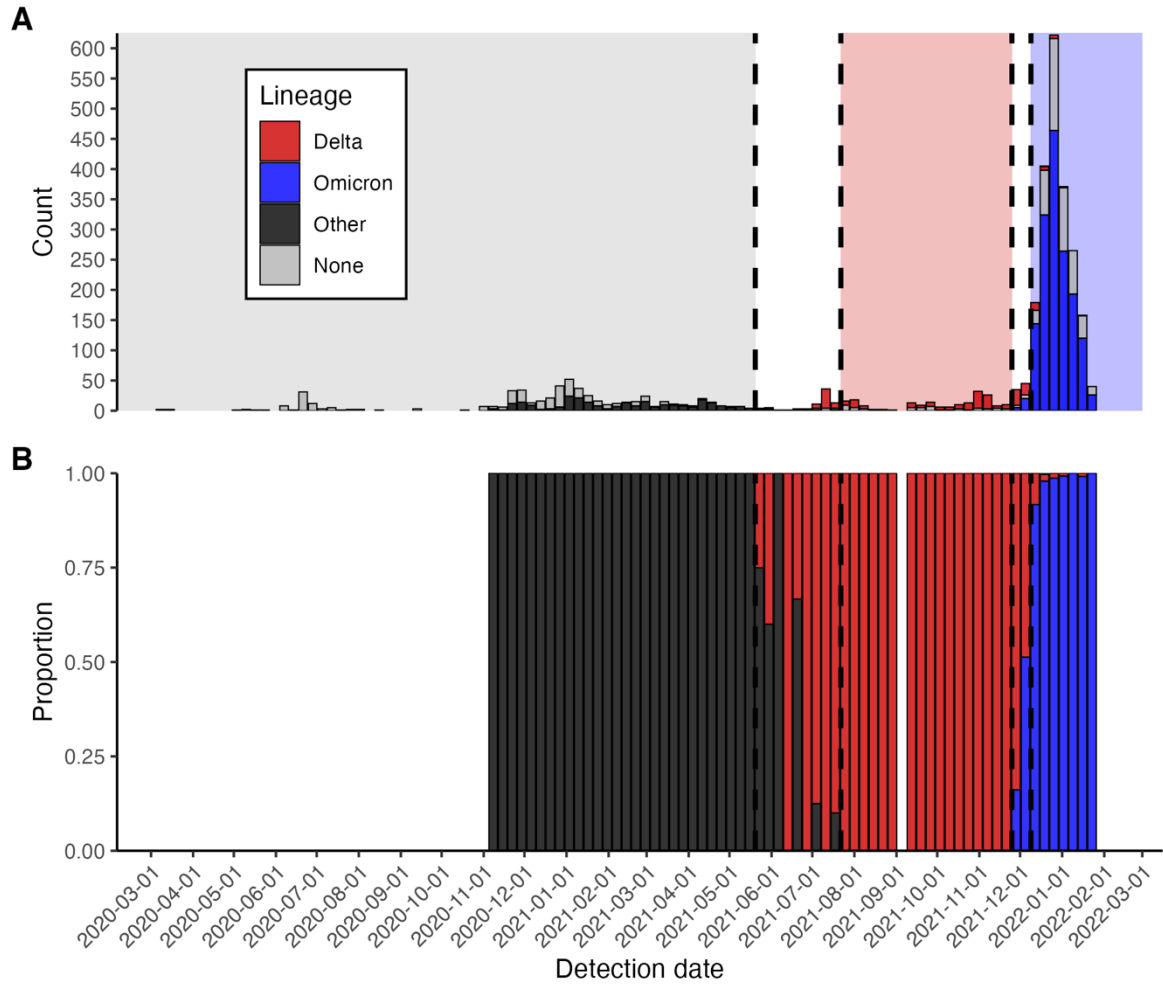


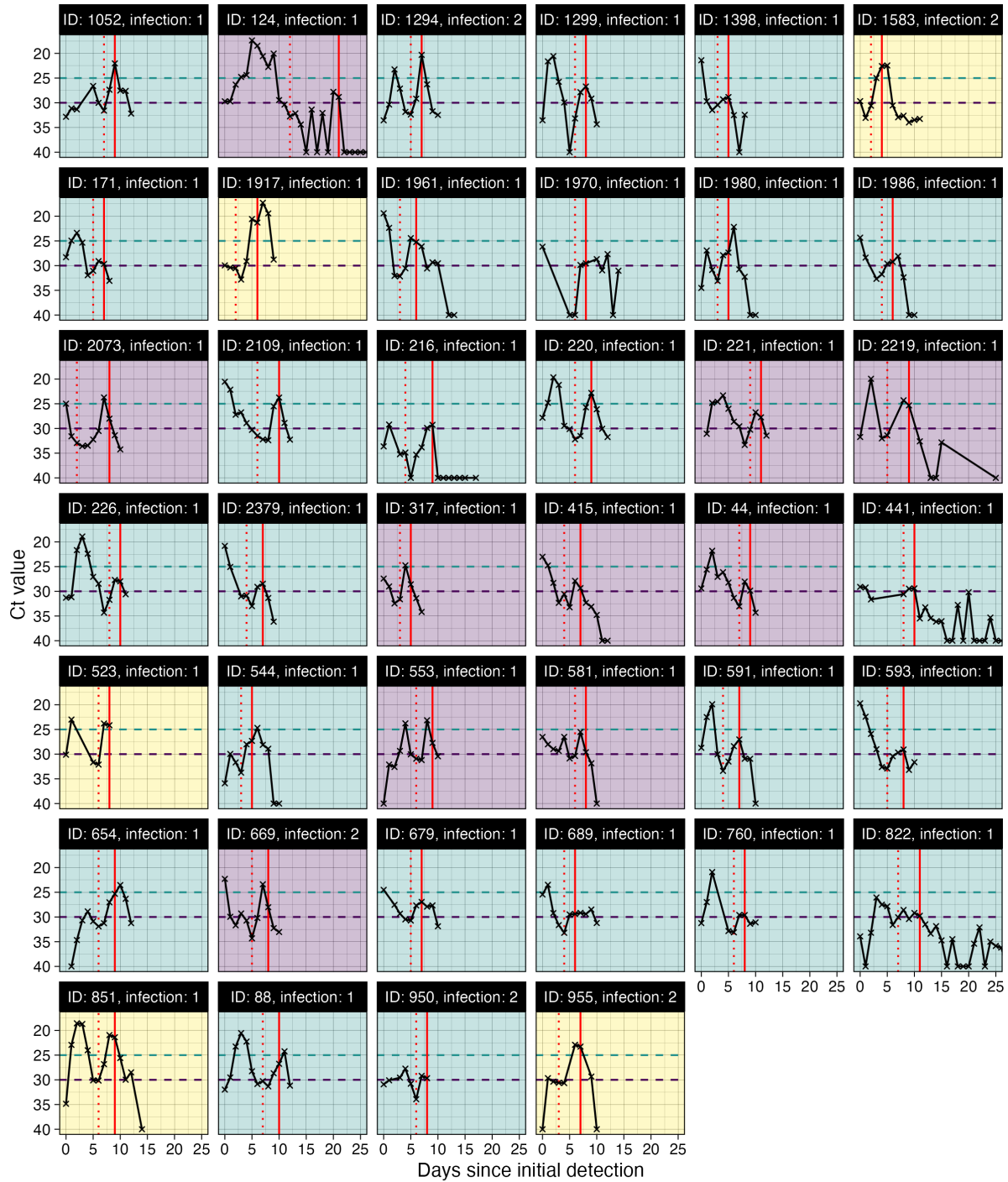
Figure S4. (A) Frequency of sequenced and unsequenced detected infections over time by week. Vertical dashed lines and shaded backgrounds demarcate periods of variant dominance. (B) Proportion of sequenced infections attributable to Delta, Omicron or other lineages.

Table S1. Number of identified rebounds stratified by variant, either confirmed through sequencing or assumed based on detection date. Rebounds are defined here as any trajectory with an initial Ct value <30, followed by a sequence of two or more consecutive negative tests or tests with Ct value ≥ 30 , and subsequently followed by two or more consecutive tests with Ct value <30.

Lineage	Rebounds	Total infections	Percentage rebounded
Omicron	36	877	4.10%
Delta	1	190	0.526%
Other	3	293	1.02 %

Table S2. Number of identified rebounds stratified by vaccination status. Rebounds are defined here as any trajectory with an initial Ct value <30, followed by a sequence of two or more consecutive negative tests or tests with Ct value ≥ 30 , and subsequently followed by two or more consecutive tests with Ct value <30.

Vaccination status	Rebounds	Total infections	Percentage rebounded
Boosted	32	494	6.48 %
No record	5	420	1.19%
Second dose	3	325	0.923%
Unvaccinated	2	161	1.24%



Supplementary Figure S5. All viral trajectories classified as rebound shown in Figure 1B. Subplots are colored by the most stringent definition for rebound. To be included here, individuals must have 2+ consecutive days of $Ct \geq 30$ after an initial $Ct < 30$. The vertical red dotted line marks this initial clearance time. Trajectories are then classified as rebounds following either two consecutive tests with $Ct < 30$ (purple), two consecutive tests with $Ct < 30$ but

with at least a 2 Ct decrease (green), or two consecutive tests with $Ct < 25$ (yellow). The vertical red line marks the timing of rebound detection. The horizontal dashed lines show the different Ct value thresholds for rebound classification. Panels are labeled by arbitrary person ID and infection number.

Table S3. Comparison of linear logistic regression models predicting probability of Ct<30 on each day since detection among individuals in the frequent testing group. Models are ranked based on their expected log pointwise predictive density (ELPD), where a lower ELPD implies better prediction accuracy. Model weight refers to the weight of each model in a Bayesian Model Averaging analysis, where a higher value implies a greater contribution to model prediction when combining multiple models. AUC = area under the curve.

Model	ELPD difference	SE difference	Model weight	AUC	Classification accuracy	Accuracy (>=30)	Accuracy (<30)
Cumulative number of exposures and lineage	0.000	0.000	0.189	0.900	0.829	0.894	0.637
Days since previous exposure and lineage	-6.053	11.613	0.281	0.899	0.826	0.890	0.639
Vaccination status and lineage	-6.907	11.836	0.292	0.899	0.826	0.887	0.649
Cumulative number of exposures	-24.622	8.645	0.237	0.896	0.825	0.898	0.613
Vaccination status	-43.428	13.353	0	0.892	0.821	0.888	0.626
Days since previous exposure	-44.910	13.731	0.001	0.893	0.820	0.890	0.616
Lineage	-74.343	12.956	0	0.888	0.818	0.891	0.605
Baseline	-80.956	14.080	0	0.887	0.817	0.894	0.592

Table S4. Comparison of linear logistic regression models predicting probability of Ct<30 on each day since detection among individuals in the delayed testing group. Models are ranked based on their expected log pointwise predictive density (ELPD), where a lower ELPD implies better prediction accuracy. Model weight refers to the weight of each model in a Bayesian Model Averaging analysis, where a higher value implies a greater contribution to model prediction when combining multiple models. AUC = area under the curve.

Model	ELPD difference	SE difference	Model weight	AUC	Classification accuracy	Accuracy (>=30)	Accuracy (<30)
Cumulative number of exposures and lineage	0.000	0.000	0.314	0.913	0.848	0.896	0.696
Cumulative number of exposures	-3.543	6.862	0.282	0.911	0.847	0.895	0.694
Vaccination status and lineage	-19.169	12.714	0.404	0.91	0.846	0.890	0.705
Vaccination status	-31.913	13.496	0	0.909	0.845	0.890	0.703
Days since previous exposure and lineage	-44.479	12.735	0	0.91	0.846	0.891	0.705
Lineage	-61.323	13.566	0	0.908	0.844	0.883	0.718
Days since previous exposure	-87.874	15.524	0	0.906	0.843	0.882	0.720
Baseline	-117.764	17.502	0	0.904	0.841	0.877	0.725

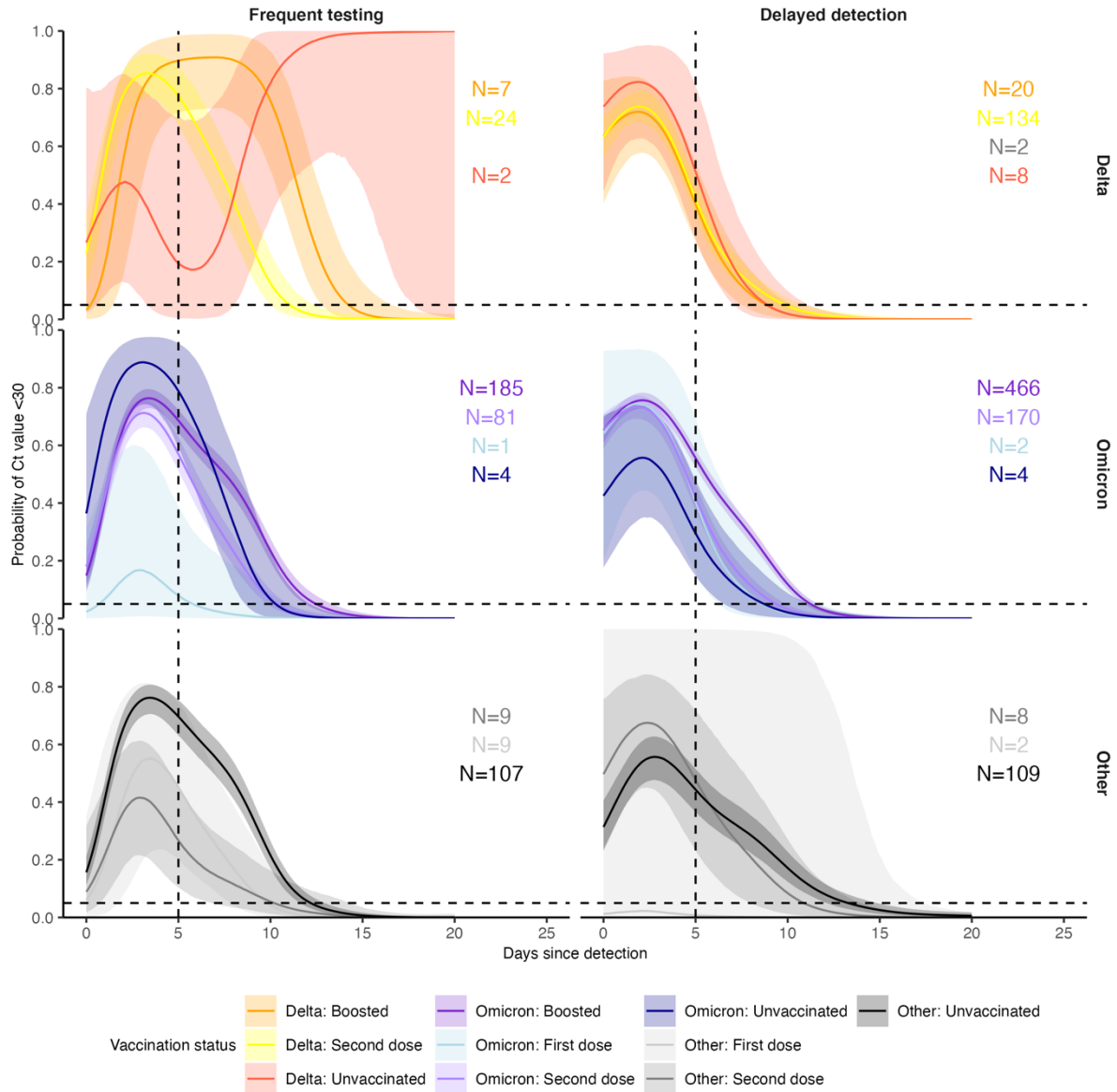


Figure S6. Proportion of infections with Ct value <30 on each day post detection by confirmed or suspected variant, vaccination status and detection group. Solid colored lines and shaded ribbons are posterior estimates from a generalized linear model predicting probability of Ct value <30 as a function of days since detection and vaccination status, showing the posterior mean (solid line) and 95% credible intervals (shaded ribbon) of each conditional effect. Dotted horizontal and vertical lines show 5% probability and day 5 post detection respectively.

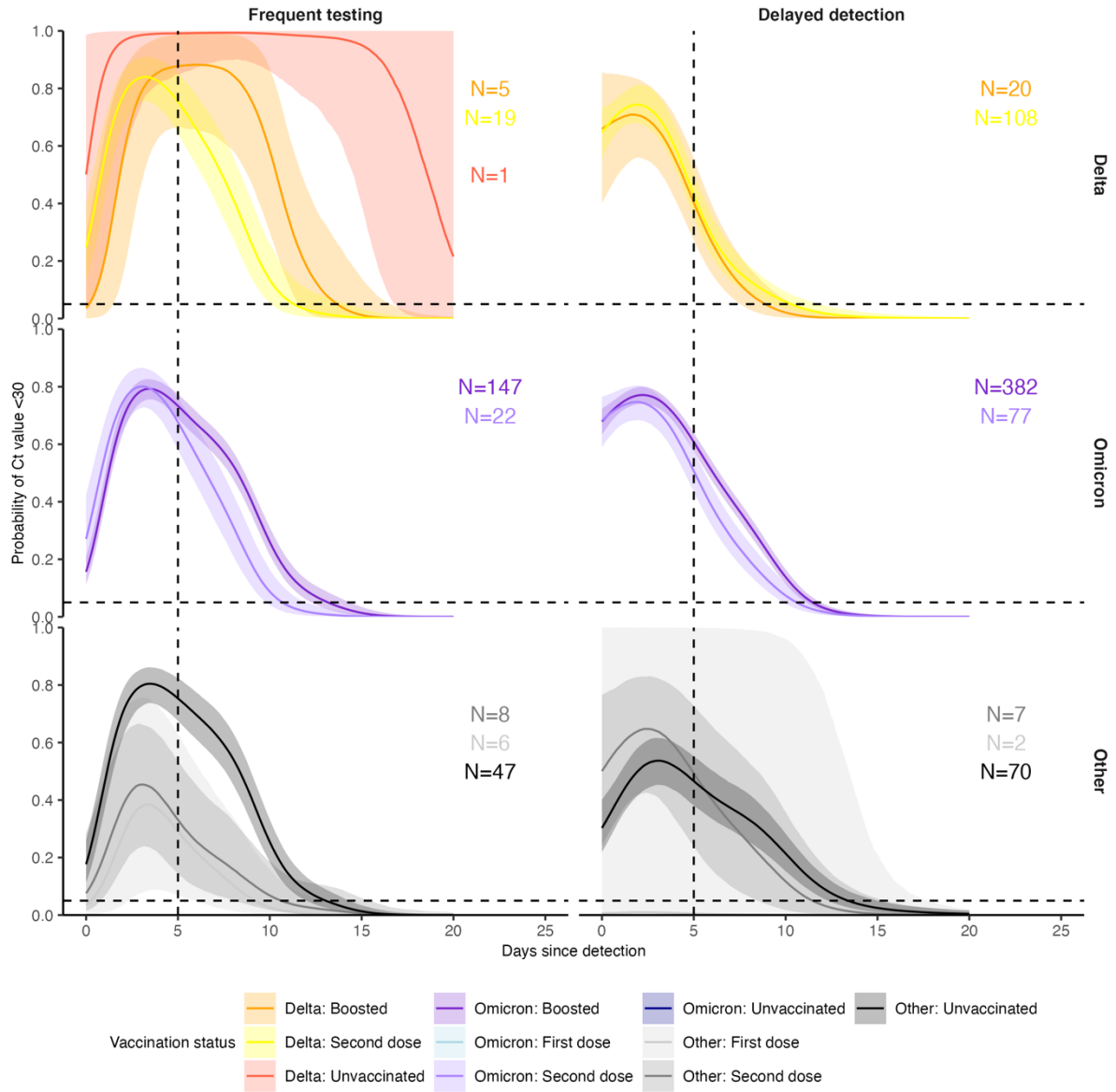


Figure S7. Identical to Supplementary Figure 6, but after excluding data from all players.

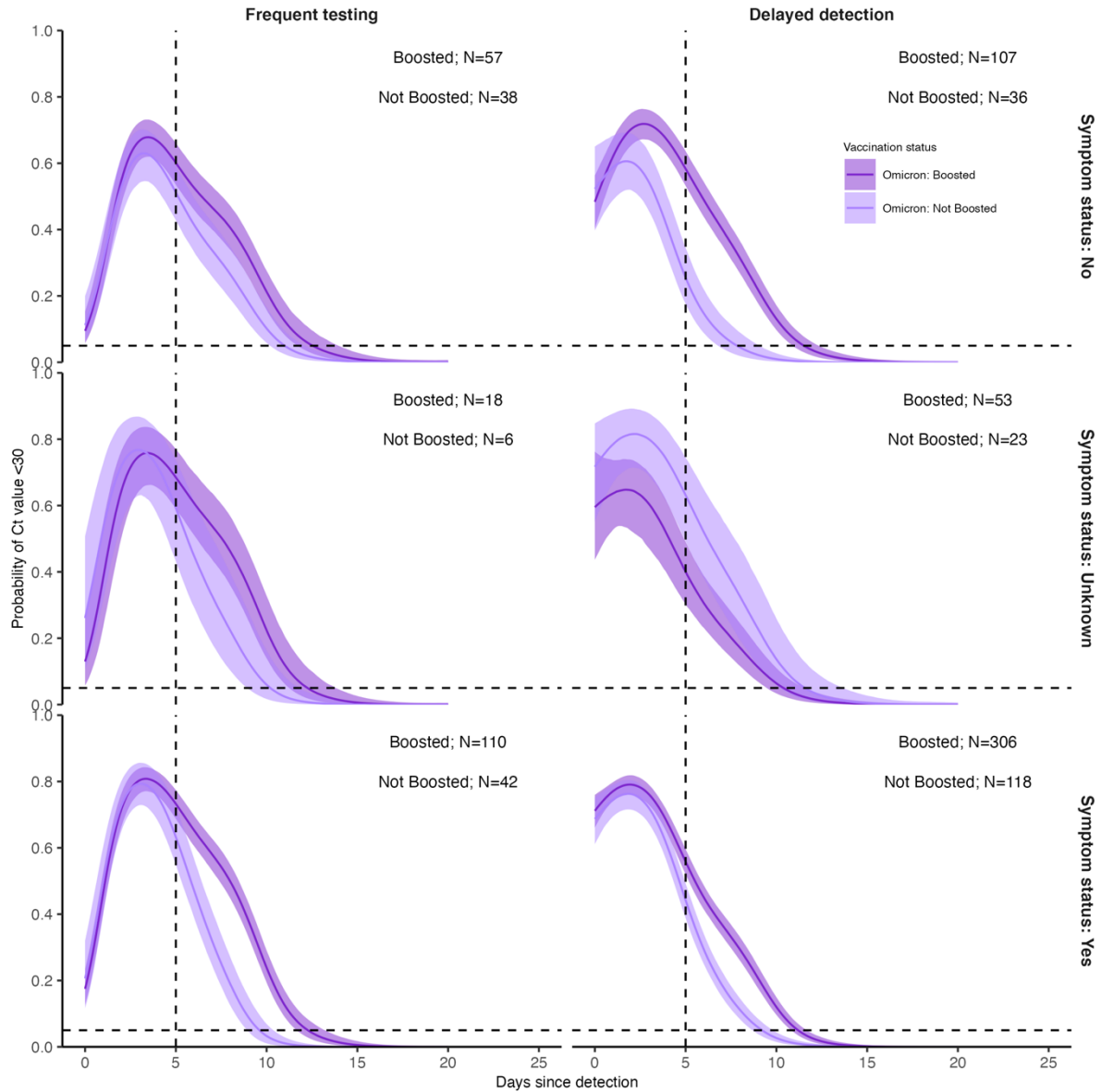


Figure S8. Proportion of Omicron infections, stratified by symptom status, with Ct value <30 on each day post detection by booster status and detection group. Solid colored lines and shaded ribbons are posterior estimates from a generalized linear model predicting probability of Ct value <30 as a function of days since detection and vaccination status, showing the posterior mean (solid line) and 95% credible intervals (shaded ribbon) of each conditional effect.

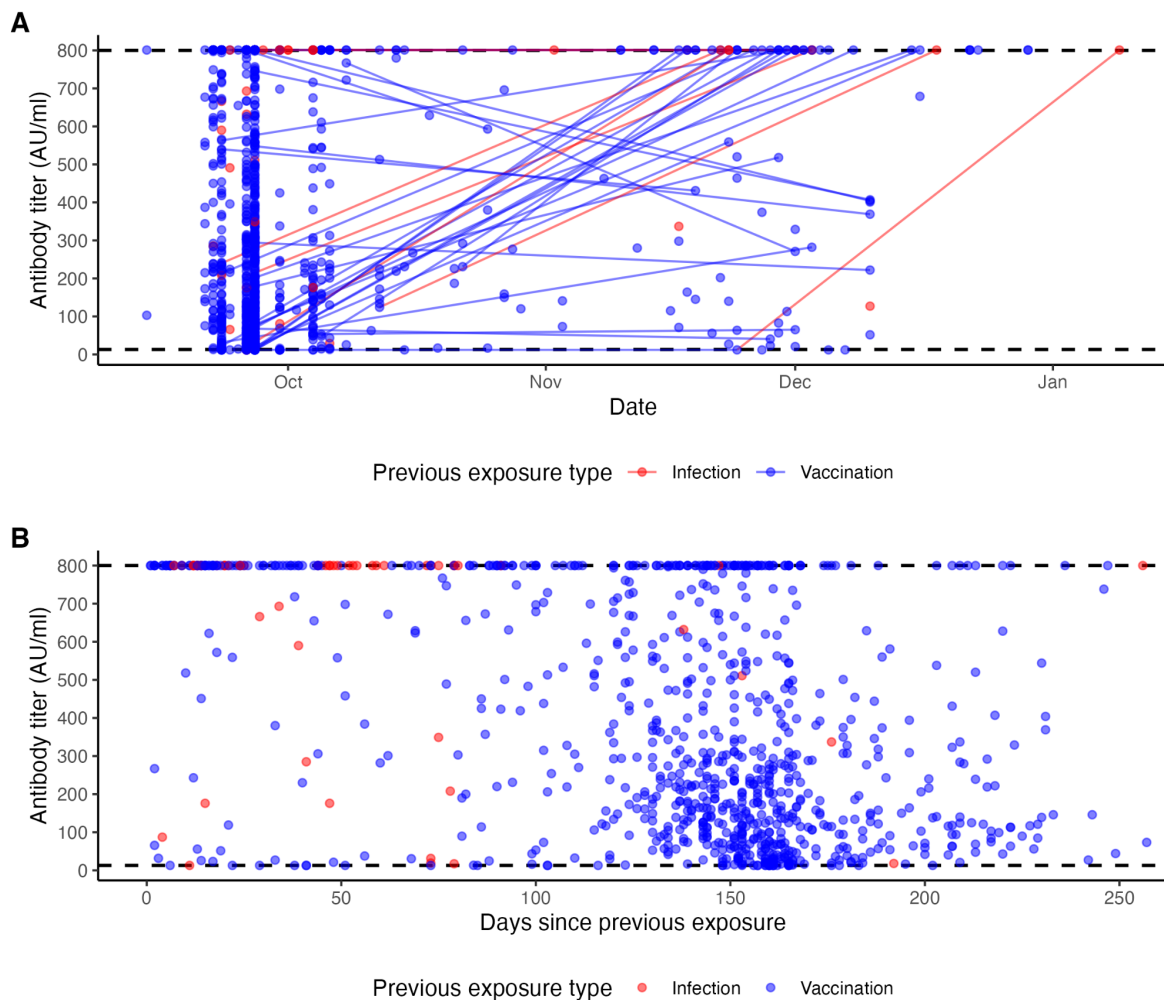


Figure S9. (A) Measured antibody titers by date of sample collection. Lines show longitudinal samples from the same individual, colored by the most recent exposure at the time of sample collection. Lines going up therefore represent antibody boosting events, and lines going down represent waning. **(B)** Measured antibody titers by days since previous exposure at time of sample collection.

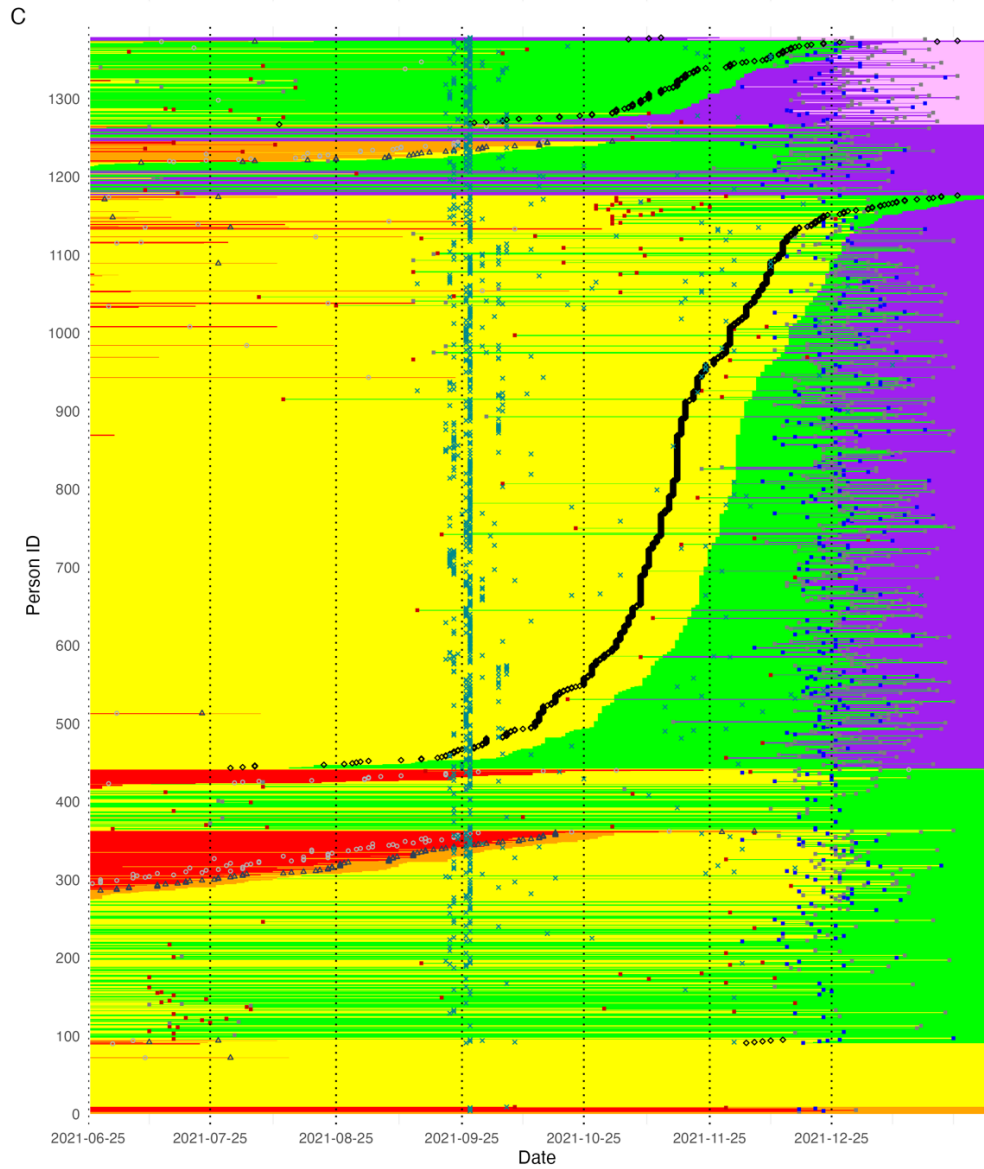
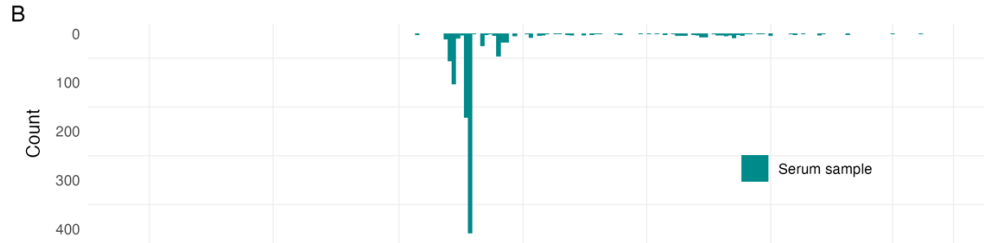
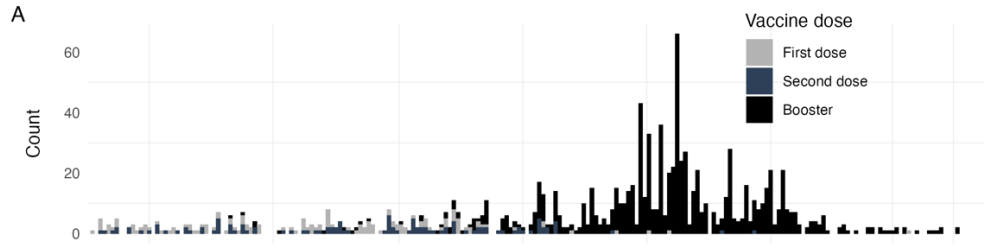


Figure S10. (A) Distribution of vaccination dates (note that most first doses were administered prior to 2021-06-25). (B) Distribution of serum sample times. (C) Heatmap of individual exposure status over time. Rows represent individuals and columns represent date. Each cell is shaded by the number of prior exposures at that date. Points show detected infections, recorded vaccinations, and serum samples.

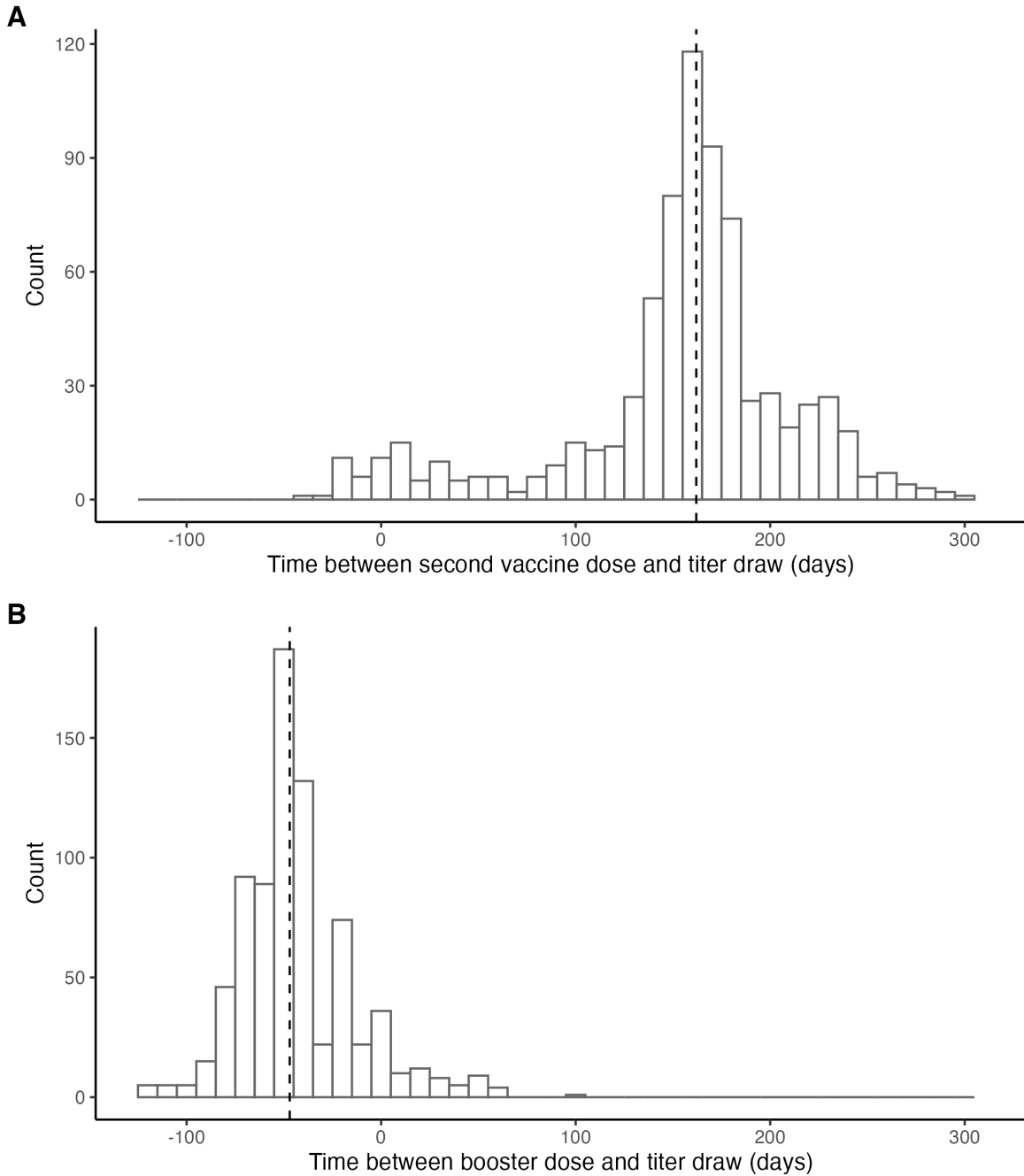


Figure S11. Histogram of time between (A) second vaccine dose and antibody titer measurement and (B) booster dose and antibody titer measurement. Dashed line marks the median lag (162 days). 1 individual was infected between receiving their second vaccine dose and having a titer measurement taken (Delta infection). 42 individuals were infected between having their titer measurement taken and receiving their booster vaccine dose (32 Delta; 9 unsequenced; 1 confirmed Omicron BA.1).

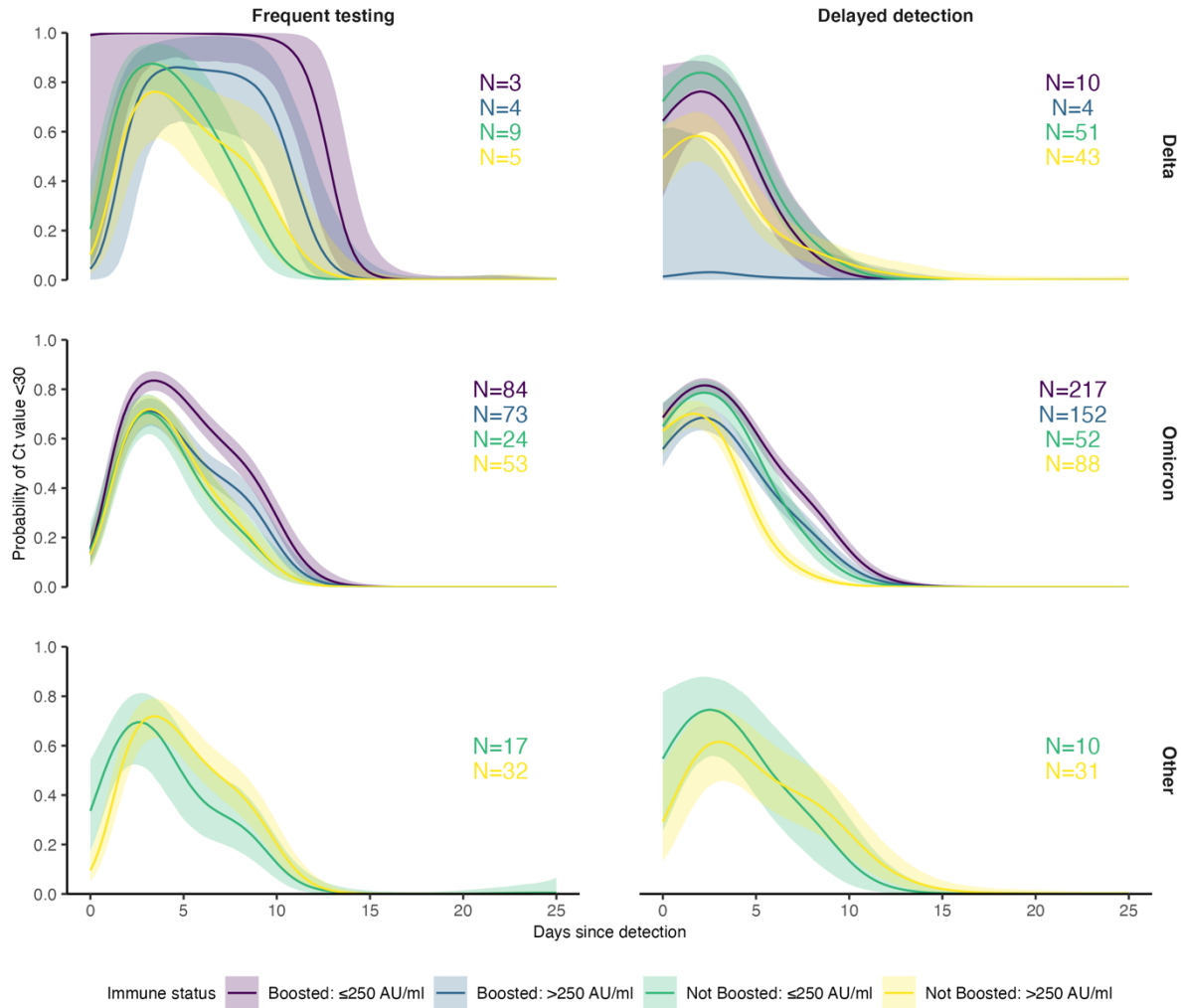


Figure S12. Proportion of infections with Ct value <30 on each day post detection stratified by single point-in-time anti-spike antibody titer against ancestral SARS-CoV-2 measured by the Diasorin Trimeric Assay. Solid colored lines and shaded ribbons are posterior estimates from a generalized linear model predicting probability of Ct value <30 as a function of days since detection and titer/vaccination status category, showing the posterior mean (solid line) and 95% credible intervals (shaded ribbon) of each conditional effect.

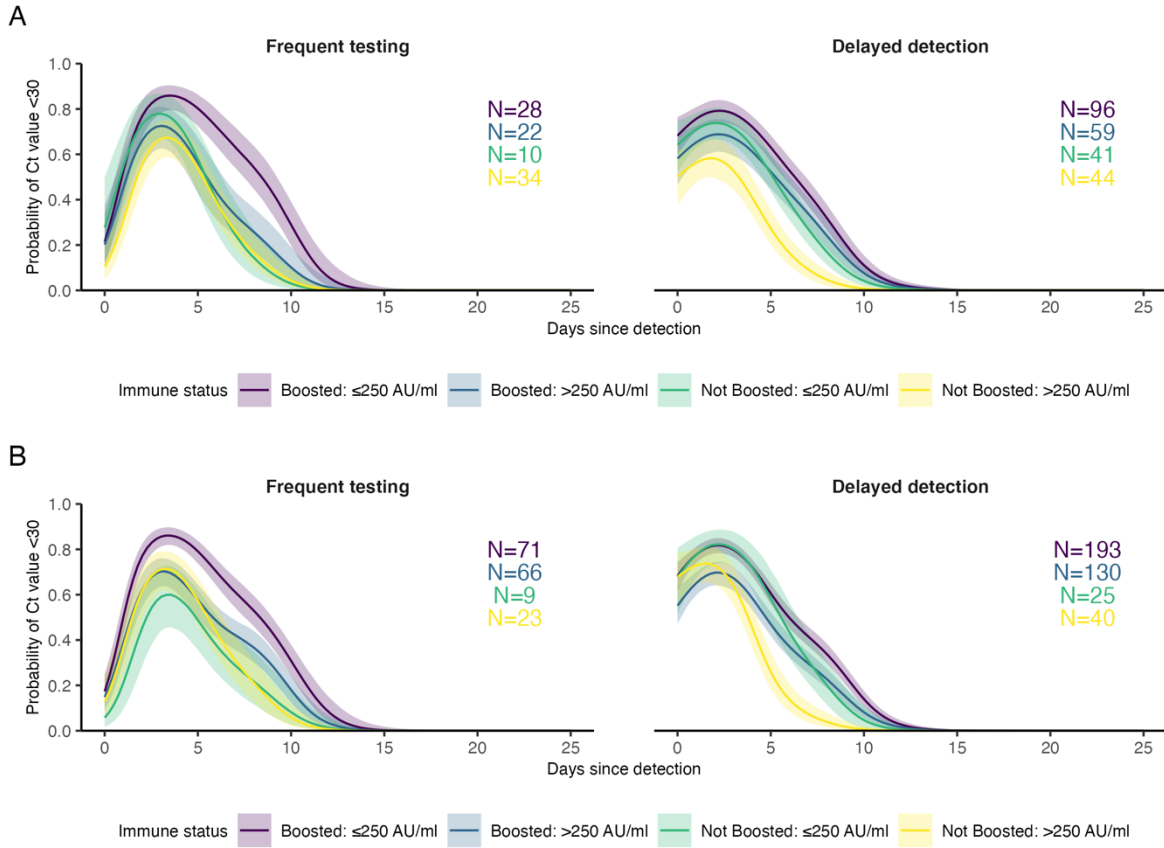


Figure S13. Identical to Omicron plots shown in Supplementary Figure 11, but only including individuals who had antibody titers measured between 100 and 200 days following a known previous infection of vaccination (A) or including infections between 60 and 90 days after an antibody titer measurement (B).

Table S5. Posterior estimates of viral trajectory attributes by variant and vaccination status. Estimates are posterior means with 95% credible intervals.

Trajectory value	Variant/vaccination status	Estimate
Peak viral load (Ct)	Other: Unvaccinated	25.7 (24.9, 26.5)
	Delta: 1-2 doses	22.8 (21.8, 23.9)
	Omicron: 1-2 doses	26 (25.4, 26.6)
	Omicron: Boosted	26.2 (25.8, 26.6)
Peak viral load (log ₁₀ copies/ml)	Other: Unvaccinated	6.6 (6.4, 6.8)
	Delta: 1-2 doses	7.4 (7.1, 7.7)
	Omicron: 1-2 doses	6.5 (6.4, 6.7)
	Omicron: Boosted	6.5 (6.4, 6.6)
Proliferation time (days)	Other: Unvaccinated	5.2 (4.6, 6)
	Delta: 1-2 doses	4.7 (3.9, 5.7)
	Omicron: 1-2 doses	5.2 (4.6, 5.9)
	Omicron: Boosted	6 (5.5, 6.5)
Clearance time (days)	Other: Unvaccinated	9.6 (8.9, 10.3)
	Delta: 1-2 doses	7.5 (6.8, 8.2)
	Omicron: 1-2 doses	6.1 (5.7, 6.5)
	Omicron: Boosted	8.1 (7.8, 8.5)

Table S6. Posterior estimates of viral trajectory attributes by variant and antibody titer.
 Estimates are posterior means with 95% credible intervals.

Trajectory attribute	Variant/immune status	Estimate
Peak viral load (Ct)	Other: Unexposed	25.4 (24.6, 26.2)
	Delta: Exposed, ≤ 250 AU	22.4 (21, 23.7)
	Delta: Exposed, > 250 AU	24.7 (22.9, 26.2)
	Omicron: Exposed, ≤ 250 AU	25.6 (25.1, 26)
	Omicron: Exposed, > 250 AU	26.7 (26.2, 27.1)
Peak viral load (\log_{10} copies/ml)	Other: Unexposed	6.7 (6.5, 6.9)
	Delta: Exposed, ≤ 250 AU	7.5 (7.2, 7.9)
	Delta: Exposed, > 250 AU	6.9 (6.5, 7.4)
	Omicron: Exposed, ≤ 250 AU	6.7 (6.5, 6.8)
	Omicron: Exposed, > 250 AU	6.4 (6.2, 6.5)
Proliferation time (days)	Other: Unexposed	5.2 (4.5, 6)
	Delta: Exposed, ≤ 250 AU	4.4 (3.4, 5.8)
	Delta: Exposed, > 250 AU	5.4 (4.1, 7)
	Omicron: Exposed, ≤ 250 AU	5.9 (5.3, 6.5)
	Omicron: Exposed, > 250 AU	5.5 (4.9, 6.1)
Clearance time (days)	Other: Unexposed	9.4 (8.8, 10.2)
	Delta: Exposed, ≤ 250 AU	7.6 (6.8, 8.5)
	Delta: Exposed, > 250 AU	7.5 (6.3, 8.8)
	Omicron: Exposed, ≤ 250 AU	8.1 (7.7, 8.6)
	Omicron: Exposed, > 250 AU	6.7 (6.3, 7.1)

Table S7. Posterior estimates of Omicron BA.1 viral trajectory attributes by symptom and vaccination status. Estimates are posterior means with 95% credible intervals.

Trajectory attribute	Variant/immune status	Estimate
Peak viral load (Ct)	Omicron: 1-2 doses, no symptoms	27 (26, 28)
	Omicron: 1-2 doses, symptoms	25.8 (25, 26.5)
	Omicron: boosted, no symptoms	26.8 (26.1, 27.4)
	Omicron: boosted, symptoms	25.8 (25.4, 26.3)
Peak viral load (log ₁₀ copies/ml)	Omicron: 1-2 doses, no symptoms	6.3 (6, 6.5)
	Omicron: 1-2 doses, symptoms	6.6 (6.4, 6.8)
	Omicron: boosted, no symptoms	6.3 (6.2, 6.5)
	Omicron: boosted, symptoms	6.6 (6.5, 6.7)
Proliferation time (days)	Omicron: 1-2 doses, no symptoms	5.1 (4.3, 5.9)
	Omicron: 1-2 doses, symptoms	5.2 (4.4, 6.1)
	Omicron: boosted, no symptoms	5.9 (5.2, 6.9)
	Omicron: boosted, symptoms	5.9 (5.3, 6.5)
Clearance time (days)	Omicron: 1-2 doses, no symptoms	5.9 (5.2, 6.6)
	Omicron: 1-2 doses, symptoms	6.2 (5.7, 6.7)
	Omicron: boosted, no symptoms	7.2 (6.5, 7.8)
	Omicron: boosted, symptoms	8.4 (8, 8.8)

Table S8. Posterior estimates of Omicron BA.1 viral trajectory attributes by symptom and antibody titer. Estimates are posterior means with 95% credible intervals. Low titer is ≤ 250 AU, high titer is >250 AU.

Trajectory attribute	Variant/immune status	Estimate
Peak viral load (Ct)	Omicron: Low titer, no symptoms	26.4 (25.6, 27.3)
	Omicron: Low titer, symptoms	25.2 (24.7, 25.7)
	Omicron: High titer, no symptoms	27.1 (26.3, 27.9)
	Omicron: High titer, symptoms	26.4 (25.9, 26.9)
Peak viral load (\log_{10} copies/ml)	Omicron: Low titer, no symptoms	6.4 (6.2, 6.7)
	Omicron: Low titer, symptoms	6.8 (6.6, 6.9)
	Omicron: High titer, no symptoms	6.2 (6, 6.4)
	Omicron: High titer, symptoms	6.4 (6.3, 6.6)
Proliferation time (days)	Omicron: Low titer, no symptoms	6 (5.2, 6.9)
	Omicron: Low titer, symptoms	5.6 (5, 6.3)
	Omicron: High titer, no symptoms	5.2 (4.4, 6)
	Omicron: High titer, symptoms	5.7 (5, 6.5)
Clearance time (days)	Omicron: Low titer, no symptoms	7.2 (6.6, 8)
	Omicron: Low titer, symptoms	8.5 (8, 9.1)
	Omicron: High titer, no symptoms	6.3 (5.7, 6.9)
	Omicron: High titer, symptoms	6.8 (6.3, 7.2)

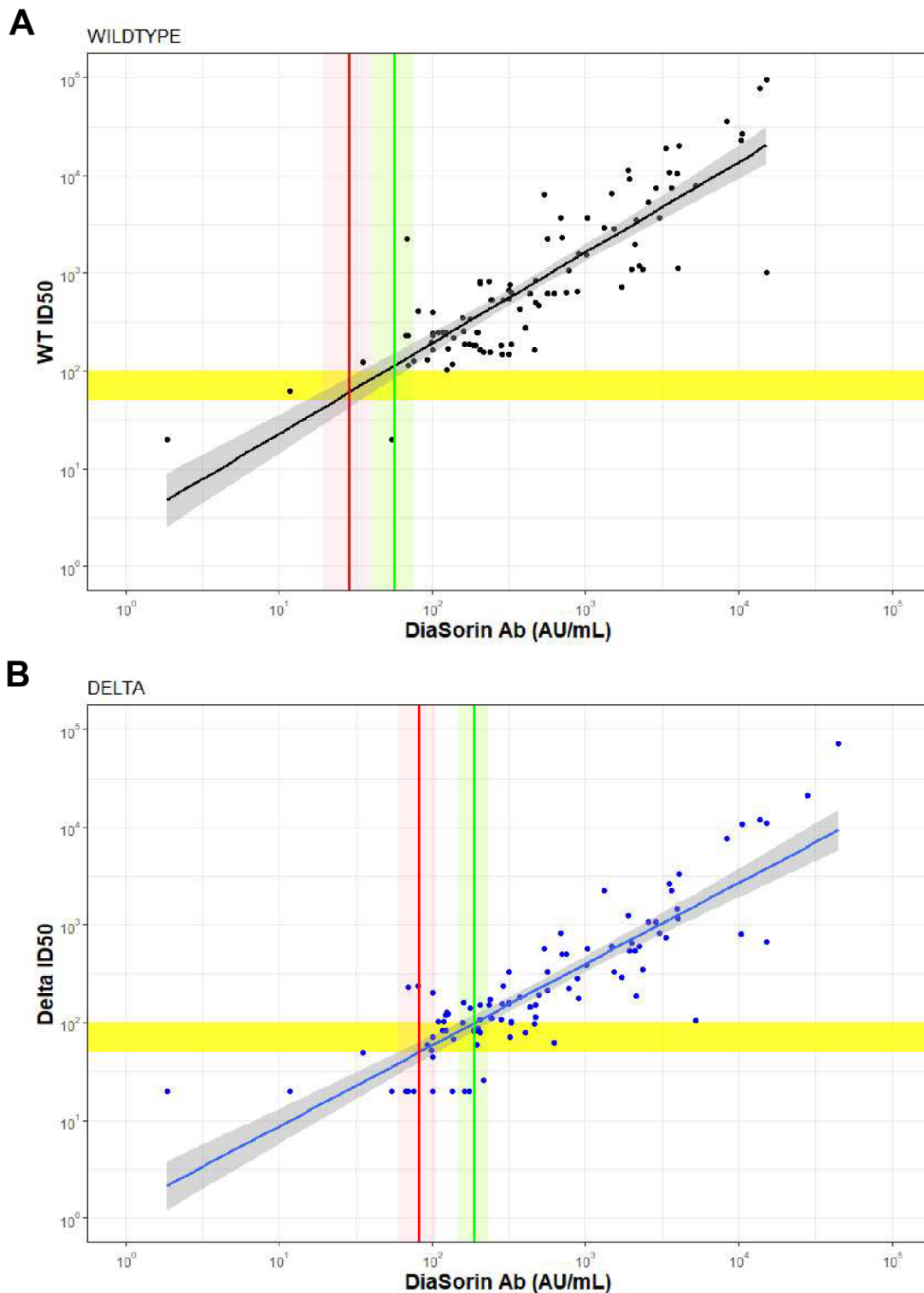


Figure S14. Correlation between authentic virus neutralization assay (ID50) and the Diasorin antibody titer against (A) wildtype and (B) Delta. Horizontal yellow bar shows an ID50 titer of 50 and 100 respectively, Diagonal lines and shaded regions show mean and 95% confidence intervals (CI) for a linear regression between the Diasorin antibody titer and ID50 titer.

line and shaded regions show point estimate and 95% CI for the Diasorin antibody titer corresponding to an ID50 titer of 50 (red) and 100 (green).

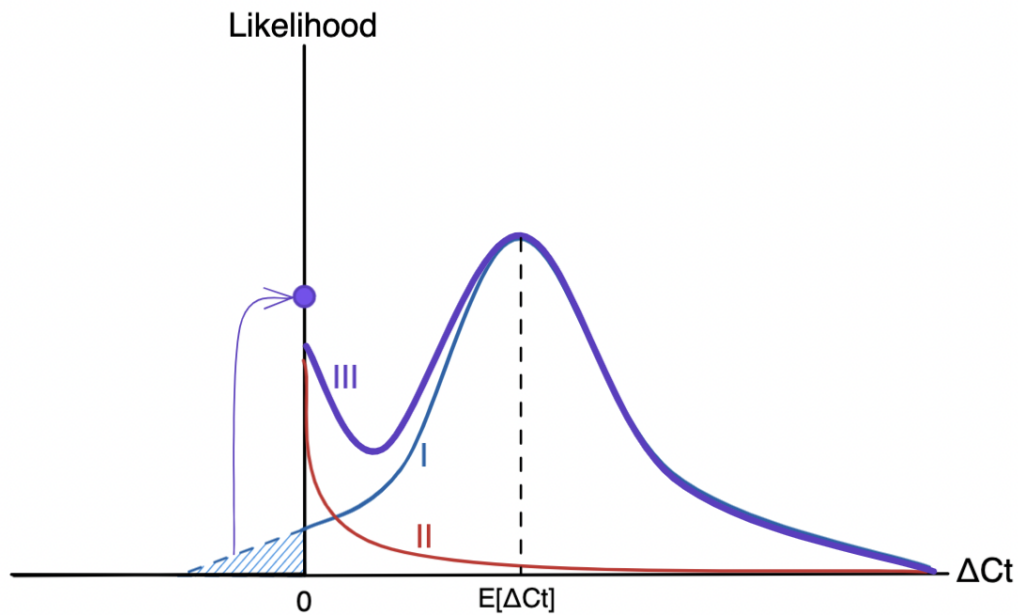


Figure S15. Schematic diagram of the likelihood function for viral kinetic inference. The plot depicts the likelihood as a function of ΔC_t , the difference between the observed C_t value and the limit of detection, so that $\Delta C_t = 0$ (the origin) represents observations at the limit of detection, with viral load increasing toward the right-hand side of the plot. The likelihood function (III, purple) is made up of two fundamental components: the process likelihood (I, blue) and the false negative distribution (II, red). The main component of the process likelihood (I) is defined by a normal distribution with mean $E[\Delta C_t]$, a function of the estimated viral kinetic parameters as defined by the viral kinetic model. Any mass of the process likelihood that extends below the limit of detection (blue hatched region) is instead added to a point probability mass at the origin, since viral loads below the limit of detection register at the limit of detection. The false negative distribution (II) is an exponential distribution with fixed rate to account for a small amount of noise near the limit of detection. Summing the process likelihood (I) and the false negative likelihood (I) using the mixing probability λ ($= 1 - \text{sensitivity}$) yields the overall likelihood (III).