




RESEARCH ARTICLE OPEN ACCESS

Defining the Critical Requisites for Accurate Simulation of SARS-CoV-2 Viral Dynamics: Patient Characteristics and Data Collection Protocol

Hoong Kai Chua¹  | Ananya Singh² | Yuqian Wang² | Yun Shan Goh³ | Conrad En Zuo Chan⁴  | Jean-Marc Chavatte⁴ | Raymond Valentine Tzer Pin Lin⁴ | Yvonne C. F. Su⁵ | Marco Ajelli⁶ | Po Ying Chia^{4,7} | Sean W. X. Ong^{4,7} | David Chien Lye^{2,4,7,8} | Barnaby E. Young^{2,4,7} | Keisuke Ejima^{2,4} 

¹School of Biological Sciences, Nanyang Technological University, Singapore, Singapore | ²Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore | ³A*STAR Infectious Diseases Labs (A*STAR ID Labs), Agency for Science, Technology and Research (A*STAR), Singapore, Singapore | ⁴National Centre for Infectious Diseases, Singapore, Singapore | ⁵Programme in Emerging Infectious Diseases, Duke-NUS Medical School, Singapore, Singapore | ⁶Laboratory for Computational Epidemiology and Public Health, Department of Epidemiology and Biostatistics, Indiana University School of Public Health-Bloomington, Bloomington, Indiana, USA | ⁷Department of Infectious Diseases, Tan Tock Seng Hospital, Singapore, Singapore | ⁸Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

Correspondence: Barnaby E. Young (barnaby_young@ncid.sg) | Keisuke Ejima (keisuke.ejima@ntu.edu.sg)

Received: 20 June 2024 | **Revised:** 5 November 2024 | **Accepted:** 5 January 2025

Funding: The study was supported by the Ministry of Education - Singapore (RLMOE100201900000001), Lee Kong Chian School of Medicine (LKCmedicine-SUG, #022487-00001), Nanyang Technological University, and Japan Science and Technology Agency (JPMJPR23R3).

Keywords: COVID-19 | isolation | mathematical model | policy guidance | SARS-CoV-2

ABSTRACT

Mathematical models of viral dynamics are crucial in understanding infection trajectories. However, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) viral load data often includes limited sparse observations with significant heterogeneity. This study aims to: (1) understand the impact of patient characteristics in shaping the temporal viral load trajectory and (2) establish a data collection protocol (DCP) to reliably reconstruct individual viral load trajectories. We collected longitudinal viral load data for SARS-CoV-2 Delta and Omicron variants from 243 patients in Singapore (2021–2022). A viral dynamics model was calibrated using patients' age, symptom presence, and vaccination status. We assessed associations between these patient characteristics and aspects of viral dynamics using linear regression models. We evaluated the accuracy of viral load trajectory estimation under different simulated DCPs by varying patient numbers, test frequencies, and test intervals. Older unvaccinated individuals had a longer viral shedding duration due to lower infection and cell death rates. Higher peak viral loads were found in older, symptomatic, and vaccinated individuals, with earlier peaks in younger vaccinated individuals. Symptom presence and vaccination resulted in a shorter time from infection to diagnosis. To accurately estimate viral dynamics, more frequent tests, longer test intervals, and larger patient samples are required. For 500 patients, a 21-day follow-up with measurements every 3 days and an 8-day follow-up with daily measurements was optimal for the Delta and Omicron variants, respectively. Patient characteristics significantly impacted viral dynamics. Our analytic approach and recommended DCPs can enhance preparedness and response to emerging pathogens beyond SARS-CoV-2.

Hoong Kai Chua and Ananya Singh are co-first authors.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2025 The Author(s). *Journal of Medical Virology* published by Wiley Periodicals LLC.

1 | Introduction

The spread of multiple severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants of concern during the coronavirus disease 2019 (COVID-19) pandemic has demonstrated the pivotal role of viral load data in pandemic preparedness. For example, viral load data is useful to inform contact tracing [1], develop isolation guidelines [2, 3], and monitor the epidemiological situation [4]. Furthermore, viral load data and its related measurements are used as clinical outcomes and as a tool to triage patients as its associations with disease severity, risk of requiring intubation, mortality [5–7], and transmissibility [8] have been suggested.

However, one potential issue is that viral load changes dynamically over the course of infection: it first increases exponentially, hits a peak, and eventually declines. Additionally, the time scale must be standardized across individuals before use in epidemiological and clinical practices. When using viral load data as an outcome, data collected at the same time point after symptom onset (or infection if known) can be fairly compared; however, aligning viral load measurements to the same time scale is challenging. For other viral load-related outcomes, observing the exact timing of negative conversion is technically impossible because it occurs sometime between positive and negative viral test results, and the peak viral load appears between observations. Moreover, measurement error [9, 10] may yield a positive retest following a negative result [11–13], making the definition of negative conversion subjective (some studies set the first or second consecutive negative result as negative conversion) [14, 15].

To explain the temporal viral load, one approach is to fit viral dynamics models to discretely observed viral load data [16–20]. While viral load cannot be measured continuously at every point within a certain time interval, these models extrapolate viral load over the course of infection. However, research focusing on viral dynamics remains limited. Despite numerous studies examining SARS-CoV-2 viral load (4003 PubMed results using search terms “viral load” AND “SARS-CoV-2” as of April 29, 2024), a smaller subset has specifically focused on the viral dynamics (223 PubMed results using search terms “viral load dynamics” AND “SARS-CoV-2” as of April 29, 2024). Indeed, a recent systematic review reported the association between single-timepoint viral load and disease severity to be inconclusive [21], possibly due to inconsistent timings of data collection. Conversely, by modeling viral dynamics, Néant et al. found temporal viral load and mortality to be significantly associated [18], especially when viral load at the late phase of infection was considered. These findings collectively imply that late-phase viral load may reflect immunological differences between patients, which is associated with mortality risk. Therefore, approaches to estimate viral load and viral load-related outcomes from limited sparse viral load observations are warranted.

Undoubtedly, having the full longitudinal viral load data over the entire infection time course is essential to correctly quantify the viral dynamics model. However, significant variability in viral dynamics among individuals requires that patient characteristics are considered to improve estimations. Furthermore,

it is currently unclear what data collection protocol (DCP) is suitable to accurately estimate viral dynamics. Under practical clinical settings, viral load data are often collected at discrete time points, inconsistent intervals, and usually after symptom onset, thus raising concerns about reliably estimating the viral load trajectory. For example, parameter identification becomes difficult if viral load data during the exponential phase following initial infection is completely absent [22].

Therefore, this study aims to: (1) understand the impact of patient characteristics in shaping the temporal viral load trajectory and (2) establish a DCP to reliably reconstruct individual viral load trajectories.

2 | Methods

2.1 | Study Data and Design

We analyzed two distinct clinical datasets comprising viral load observations from patients infected with SARS-CoV-2 Delta (B.1.617.2) or Omicron (B.1.1.529) variants, collected from April 1, 2021 to June 14, 2021 as per our previous study [23], and December 1, 2021 to January 31, 2022, respectively. Data were obtained from electronic medical records of patients admitted to the National Centre for Infectious Diseases, Singapore. As per prevailing guidance from the Ministry of Health, Singapore, admissions were initiated primarily for isolation purposes and were not necessarily of clinical need. Nasopharyngeal swabs were collected at specific intervals after diagnosis and subjected to reverse transcriptase-quantitative polymerase chain reaction (RT-qPCR) assays. Viral load was quantified from cycle threshold (C_t) values using the conversion formula from Zou et al. [24]: $\log_{10}(\text{viral load [copies/mL]}) = -0.32 \times C_t \text{ value [cycles]} + 14.11$. The lower limit of detection was set at $C_t = 50$ ($10^{-1.89}$ copies/mL) for Delta patients and $C_t = 45$ ($10^{-0.29}$ copies/mL) for Omicron patients. The infectiousness threshold was set as $C_t = 25$ ($10^{6.11}$ copies/mL), based on our earlier study [25]. Age, day of symptom onset, and vaccination status were also collected from each patient.

Note that the epidemiological situation when data from Delta and Omicron patients were collected was highly different. The COVID-19 vaccination campaign was ongoing when the Delta variant was dominant, where 40% of the eligible population was fully vaccinated, compared to 86% at the peak of the Omicron epidemic wave [26]. During the early phase of vaccine rollout, essential workers and older age groups were prioritized. We only included patients with a known first positive diagnosis, were naïvely infected, were either unvaccinated or fully vaccinated with at least two doses of a messenger RNA vaccine (Pfizer-BioNTech BNT162b2/Moderna mRNA-1273). Given the vaccination schedule in Singapore, all vaccinations would have been administered within the year before infection (exact timing of vaccination was not recorded). Further, we only included patients with at least three viral load measurements, as we need to capture the viral load dynamics (Supporting Information: Figure 1). To avoid confounding effects arising from incomplete/non-mRNA vaccination and hybrid immunity, we excluded from the analysis patients with a history of previous infections ($n = 3$), patients who received a non-mRNA vaccine ($n = 12$),

and patients who received a single dose of an mRNA vaccine ($n = 12$).

IRB review was exempted at Nanyang Technological University (IRB-2022-1041). Informed consent for retrospective data collection was waived (NHG-DSRB 2020/01122).

2.2 | SARS-CoV-2 Viral Dynamics Model

We employed the target cell-limited model, which has been mostly used to characterize SARS-CoV-2 viral dynamics [2, 9, 27–29]. We applied a nonlinear mixed-effects modeling approach considering both fixed effects common to the population and random effects that capture interindividual variability. Although the estimated parameter values from the best-fit model are informative to infer associations between patient characteristics (i.e., model covariates) and viral dynamics, such interpretation is challenging because the impact of parameters acting collectively on viral dynamics is complex (Figure 1). Therefore, we ran a series of multiple linear regressions for each SARS-CoV-2 variant using patient characteristics as independent variables and the following six quantitative metrics of viral dynamics as dependent variables: timing of peak viral load,

peak viral load, duration of viral shedding, time from infection to diagnosis, time from symptom onset to diagnosis, and incubation period (Figure 1). Significance of the regression coefficients was evaluated via the Wald test, while the Mann–Whitney U test and chi-square test of independence were used to test for group differences in continuous and categorical variables, respectively. All statistical tests were two-tailed, with $p < 0.05$ indicating statistical significance. See Supporting Information: Methods S8 for full descriptions of the model and fitting process.

2.3 | Simulating DCPs

We examined different values for parameters governing different DCPs, namely: the number of patients (100, 500, and 1000), number of tests per patient (3–8), and time interval between consecutive tests (1–4 days). We assumed the first test to be taken on the day of diagnosis, while patient characteristics were randomly sampled to resemble the distributions in our data set. Briefly, we first simulated longitudinal viral load data from our calibrated model for the Delta and Omicron variants separately. We then fitted the viral dynamics model to the simulated data and computed the root mean squared error (RMSE)

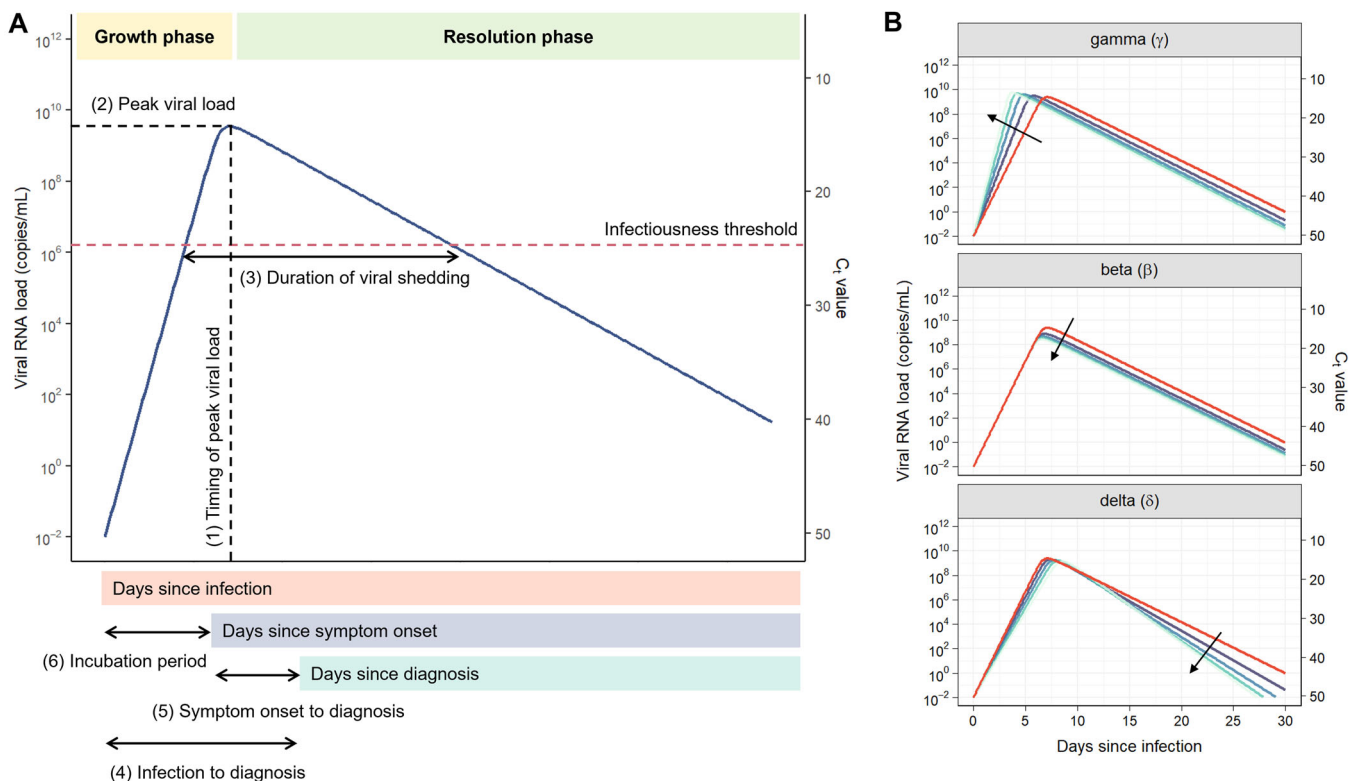


FIGURE 1 | Schematic illustration of the characteristics of the viral load trajectory. (A) The quantitative metrics of viral dynamics considered in this study were: (1) timing of peak viral load (days since infection), (2) peak viral load (copies/mL), (3) duration of viral shedding (days), (4) time from infection to diagnosis (days), (5) time from symptom onset to diagnosis (days), and (6) incubation period (days). Note that symptom onset may either precede or succeed diagnosis and may not necessarily occur at the timing of peak viral load. The horizontal red dotted line represents the infectiousness threshold, set at $10^{6.11}$ copies/mL ($C_t = 25$). (B) The changes in model behavior under the effect of varying the parameters gamma (γ), beta (β), and delta (δ) are shown in each respective panel. The red solid line represents the common reference trajectory using the same parameters $\gamma = 5.0 \text{ day}^{-1}$, $\beta = 1.0 \times 10^9 \text{ (copies/mL)}^{-1} \text{ day}^{-1}$, and $\delta = 1.0 \text{ day}^{-1}$ in all panels. The black arrows highlight the changes in viral load trajectory (colored on a gradient scale) due to the effect of an increasing parameter value of $\gamma = [6.0, 7.0, 8.0, 9.0] \text{ day}^{-1}$, $\beta = [3.0 \times 10^9, 5.0 \times 10^9, 7.0 \times 10^9, 9.0 \times 10^9] \text{ (copies/mL)}^{-1} \text{ day}^{-1}$, and $\delta = [1.2, 1.4, 1.6, 1.8] \text{ day}^{-1}$, respectively.

representing the deviation of the best-fit curves from the simulated longitudinal viral load data (Supporting Information: Figure 2). The best DCP was defined as the shortest follow-up period whilst maintaining the average margin of error below 1.25 units of \log_{10} -transformed viral load, as a longer follow-up period is a burden for both patients and health practitioners involved. See Supporting Information: Methods S8 for full descriptions of the simulation process.

3 | Results

3.1 | COVID-19 Clinical Data

We analyzed 243 (78.9%) patients, of which 154 (63.4%) and 89 (36.6%) patients were infected by the Delta and Omicron variants, respectively. The main characteristics of the analyzed sample are reported in Table 1 and are as follows: (1) only 41 (26.6%) Delta patients were fully vaccinated, whereas all Omicron patients were fully vaccinated; (2) unvaccinated Delta patients were significantly younger than vaccinated Delta patients (median age of 40 vs. 57 years, $p = 0.00235$); (3) a significantly smaller proportion of vaccinated Delta patients were symptomatic compared to unvaccinated Delta patients (78.1% vs. 90.3%, $p = 0.0462$); and (4) the proportion of vaccinated patients who were symptomatic were not significantly different between the Delta and Omicron variants (78.1% vs. 79.8%, $p = 0.822$). On average, for both SARS-CoV-2 variants, viral load was measured 4 to 7 times with an interval of 2–5 days between tests, while the absolute time interval from symptom onset to diagnosis was less than a day, suggesting that most patients were diagnosed soon after symptom onset, possibly due to public awareness and widely available mandatory free viral tests.

3.2 | Determinants of Viral Dynamics

We assessed the raw longitudinal viral load data for the Delta and Omicron variants stratified by vaccination status and/or symptom presence (Figure 2). Fitted curves for each stratified population group are shown in Figure 3, while individual-level viral load trajectories are available in Supporting Information: Figures 3 and 4. Distributions of the quantitative metrics of viral dynamics can be found in Supporting Information: Figure 5.

We examined associations between patient characteristics and viral dynamics using a series of linear regression analyses, incorporating the covariate impact on model parameters (Table 2) to provide a holistic interpretation. First, we focused on three metrics of viral dynamics related to viral kinetics which determine the shape of the viral load trajectory: timing of peak viral load, peak viral load, and viral shedding duration (Figure 4A). For both SARS-CoV-2 variants, older age was associated with a lower viral infection and replication rate, whereby a slow rise in viral load at the initial phase of infection resulted in a short delay in the timing of peak viral load, a small rise in peak viral load, and a slightly extended viral shedding duration. Symptom presence also resulted in a smaller viral infection rate, yielding similar outcomes as age. For Delta

patients only, vaccination was associated with a faster initial viral replication and an earlier timing of peak viral load by 1.95 days (95% confidence interval [CI] [1.46–2.43]). Nonetheless, vaccination reduced viral shedding significantly by 5.86 days (95% CI [3.56–8.17]), with a larger death rate of infected cells that expedited the decline in viral load following the peak. On average, Omicron patients had a lower viral load that peaks much later compared to Delta patients.

Second, we addressed three other metrics of viral dynamics related to disease progression following initial infection: time from infection to diagnosis, time from symptom onset to diagnosis, and incubation period (Figure 4B). The time from infection to diagnosis was longer in older Omicron patients by 0.09 days per year of age (95% CI [0.08–0.10]), mainly due to the longer incubation period by a similar magnitude of 0.08 days per year of age (95% CI [0.06–0.10]). Generally, symptom presence was associated with shorter time from infection to diagnosis, possibly since the onset of symptoms prompted patients to seek treatment. Compared to asymptomatic patients, symptomatic Delta patients were diagnosed significantly earlier by 4.79 days (95% CI [4.16–5.41]), whereas symptomatic Omicron patients were only diagnosed 0.28 days (95% CI [–0.007 to 0.57]) earlier. This was possibly due to the changing epidemiological landscape and patient profile, such as regular testing for asymptomatic cases. Yet, there were no significant associations between patient characteristics and the time from symptom onset to diagnosis. Among Delta patients, vaccination was associated with a reduced time from infection to diagnosis by 2.60 days (95% CI [2.11–3.09]) and a shorter incubation period by 1.87 days (95% CI [1.37–2.37]). Interestingly, while the incubation period was comparable between the SARS-CoV-2 variants, the peak viral load for the Omicron variant occurs much later than the Delta variant and increases with age, implying that symptom onset occurs much earlier than the peak viral load in Omicron patients.

3.3 | Defining DCPs

We simulated longitudinal viral load data collected under different DCPs (Supporting Information: Figure 6) and explored the accuracy of estimating individual viral load trajectories from the simulated data (Figure 5). Henceforth, we use the average RMSE as a proxy for overall model performance (i.e., a lower average RSME suggests model fits were more accurately estimated), but variations in RMSE across individual fits are found in Supporting Information: Figure 7.

Our results suggest that, for both SARS-CoV-2 variants, increasing the number of consecutive tests, interval between tests, and number of patients resulted in a lower average RMSE (Figure 5). For the Delta variant, a small sample size of 100 patients resulted in relatively worse model fits, but this was eventually alleviated at higher sample sizes of 500 and 1000 patients. Furthermore, model fits for the Omicron variant appear to be consistently better compared to the Delta variant. As the peak viral load for the Omicron variant occurs much later compared to the Delta variant, by the time patients present for testing, it is likely that viral load data collected would contain pre-peak observations (Supporting Information: Figure 6).

TABLE 1 | Summary of SARS-CoV-2 patient characteristics^a.

Variable	Delta		Omicron vaccinated (n = 89)	Delta vaccinated vs. Delta unvaccinated (p)	Omicron vaccinated vs. Delta vaccinated (p)
	unvaccinated (n = 113)	vaccinated (n = 41)			
Age range	[19–94]	[25–94]	[14–73]		
Mean age in years (SD)	45.6 (19.6)	55.1 (17.4)	34.3 (11.6)	0.00235**	< 0.001***
0–20 years old (%)	5 (4.42%)	0 (0%)	9 (10.1%)		
21–40 years old (%)	55 (48.7%)	12 (29.3%)	54 (60.7%)		
41–60 years old (%)	27 (23.9%)	13 (31.7%)	23 (25.8%)		
61–80 years old (%)	20 (17.7%)	12 (29.3%)	3 (3.37%)		
>80 years old (%)	6 (5.31%)	4 (9.76%)	0 (0%)		
Symptomatic cases (%)	102 (90.3%)	32 (78.1%)	71 (79.8%)	0.0462*	0.822
Mean number of viral load measurements (SD)	6.82 (4.79)	4.88 (1.86)	5.31 (1.34)	0.0382*	0.0752
Mean days between consecutive viral load measurements (SD)	4.33 (4.49)	5.42 (6.02)	2.23 (1.21)	0.895	< 0.001***
Mean days between symptom onset and diagnosis (SD) ^b	0.24 (0.77)	0.62 (1.21)	0.37 (0.96)	0.0230*	0.234

Abbreviations: SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; SD, standard deviation.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ by Mann-Whitney U test for continuous variables or chi-square test for categorical variables.

^aConsists only of the analyzed patients with three or more observations of viral load.

^bUsing only symptomatic patients; absolute value in days.

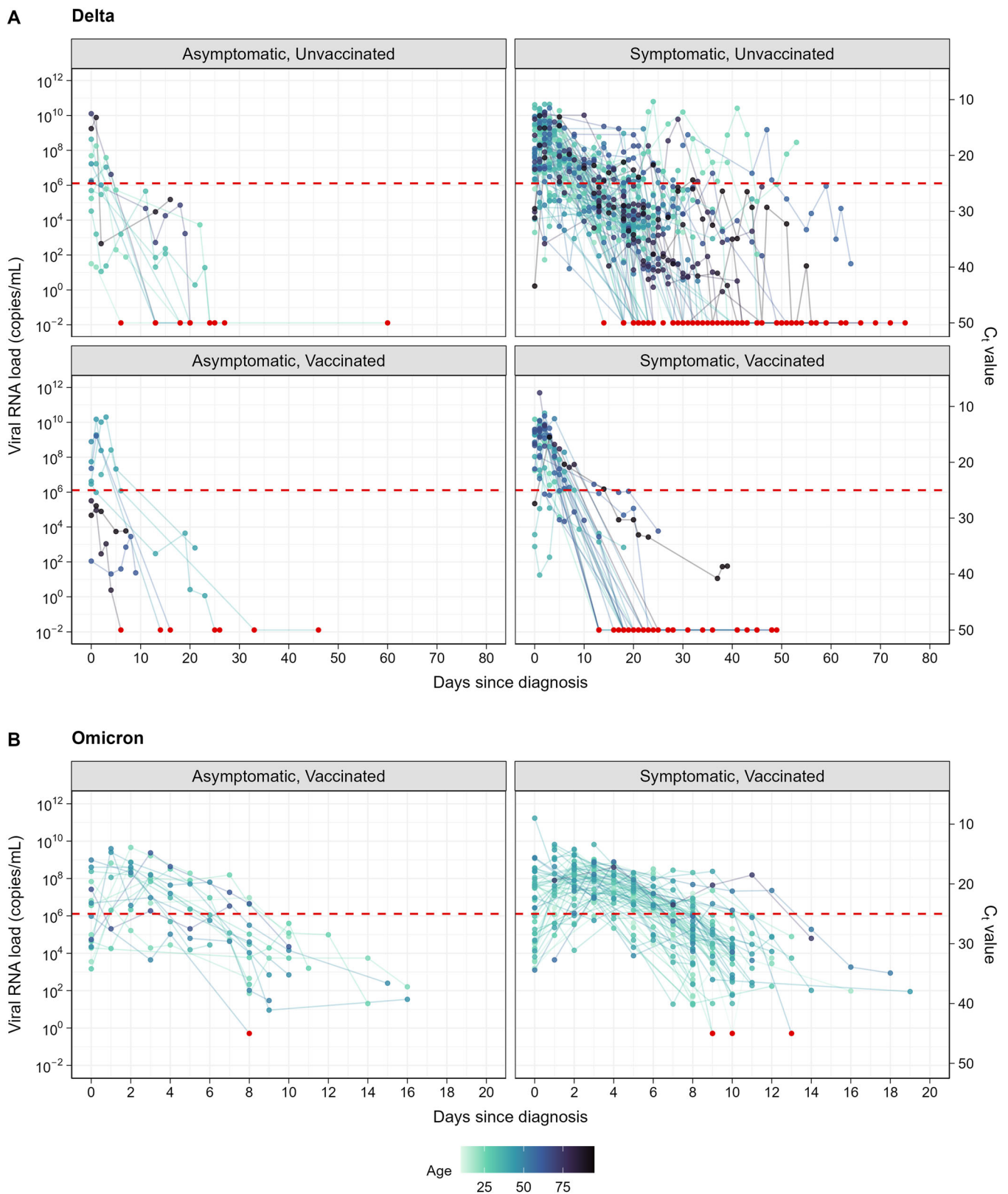


FIGURE 2 | Spaghetti plot of the raw viral load data of SARS-CoV-2 patients. (A) Delta patients stratified by vaccination status and symptom presence. The lower limit of detection was $10^{-1.89}$ copies/mL ($C_t = 50$). (B) Omicron patients stratified by symptom presence only (all were vaccinated). The lower limit of detection was $10^{-0.29}$ copies/mL ($C_t = 45$). Each point corresponds to the measured viral load, with the number of days since diagnosis as the time scale. Trajectories are depicted as lines connecting the points and are colored on a gradient scale based on continuous age (in years). Red points represent observations under the detection limit and are thus plotted at the detection limit value. The horizontal red dotted line represents the infectiousness threshold, set at $10^{6.11}$ copies/mL ($C_t = 25$). SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

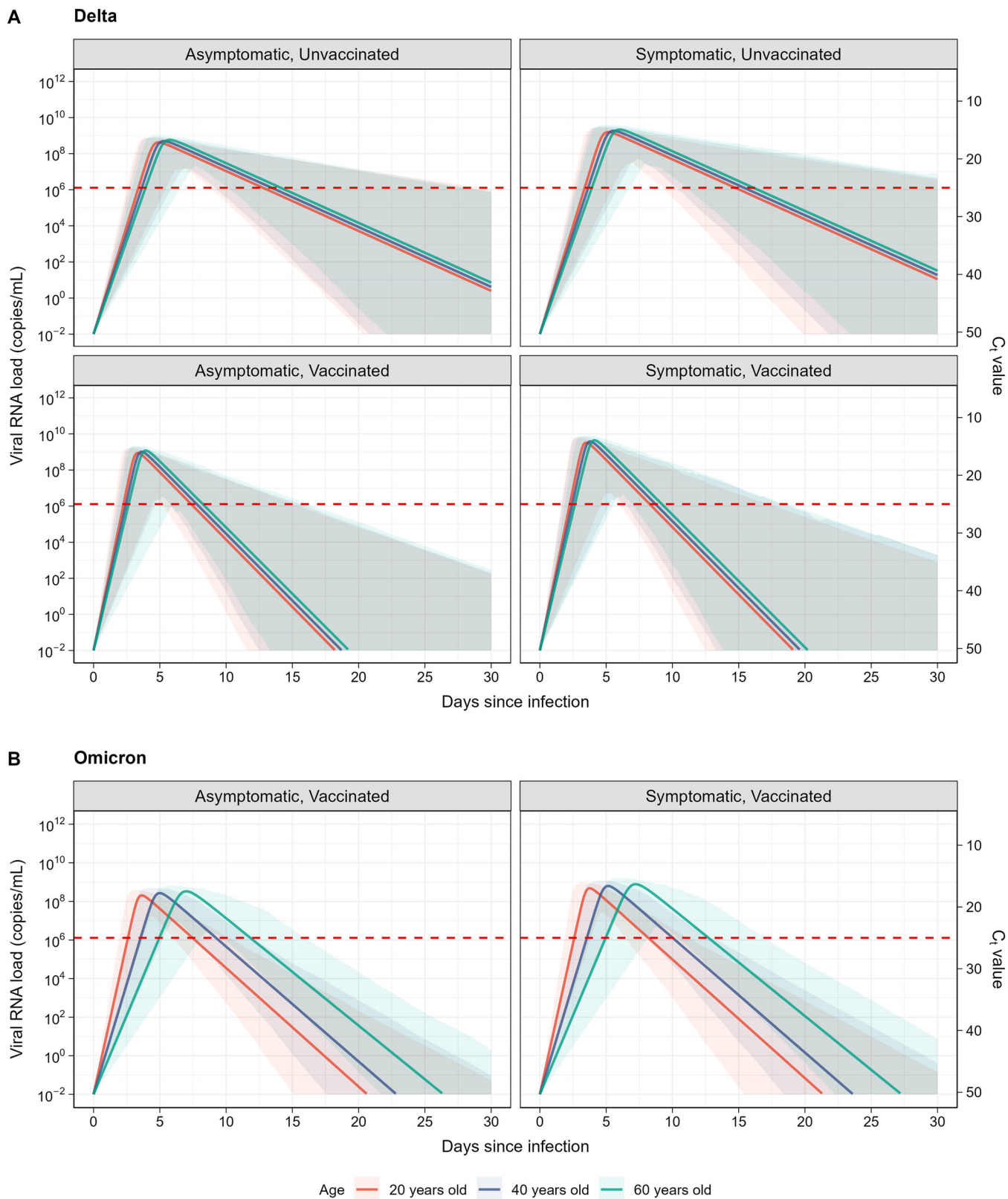


FIGURE 3 | Predicted viral load trajectories for SARS-CoV-2 patients. (A) Delta patients stratified by vaccination status and symptom presence. (B) Omicron patients stratified by symptom presence only (all were vaccinated). Three age groups: 20 years old (red), 40 years old (blue), and 60 years old (green), were chosen to depict the age effect on viral dynamics. Solid lines are the predicted viral load curves representing the population typical value, while shaded areas correspond to 95% prediction intervals representing the interindividual variability, computed using 1,000 simulated individual parameter sets. The horizontal red dotted line represents the infectiousness threshold, set at $10^{6.11}$ copies/mL ($C_t = 25$). SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

TABLE 2 | Estimated model parameters.

Parameter ^a	Maximum rate constant for viral replication		Rate constant for virus infection		Death rate of infected cells		Time from infection to diagnosis	
	Symbol	γ	β	δ	τ	Days	Days	Days
SARS-CoV-2 variant		Day ⁻¹	(copies/mL) ⁻¹ Day ⁻¹	Day ⁻¹	Days	Days	Days	Days
Median of fixed effects ^b		Delta	Delta	Delta	Omicron	Delta	Delta	Omicron
		Omicron	Omicron	Omicron	Omicron	Omicron	Omicron	Omicron
		9.33 (0.67)	1.19 × 10 ⁻⁹ (2.76 × 10 ⁻¹⁰)	1.75 (0.20)	4.54 × 10 ⁻⁹ (1.22 × 10 ⁻¹⁰)	1.44 (0.082)	3.41 (0.42)	4.07 (0.48)
Covariate effects	Age ^c	-0.0032 (0.0022)	-0.013*** (0.0038)	—	-0.034** (0.012)	—	—	0.024*** (0.0062)
	Symptom presence	—	—	—	Reference	—	—	—
	Symptomatic	—	—	—	Reference	—	—	—
	Asymptomatic	—	—	—	0.85* (0.39)	—	0.67*** (0.14)	—
	Vaccination status	—	—	—	Reference	—	—	—
	Vaccinated	—	—	—	Reference	—	—	—
	Unvaccinated	-0.47*** (0.11)	0.45 (0.26)	—	—	-0.83*** (0.13)	0.50** (0.18)	—
Standard deviation of random effects ^d		0.14 (0.045)	0.26 (0.084)	0.40 (0.16)	0.55 (0.046)	0.26 (0.039)	0.33 (0.072)	0.21 (0.075)
Standard deviation of the residual error ^d			1.83 (0.056)	1.16 (0.047)	1.16 (0.047)	1.16 (0.047)	1.16 (0.047)	1.16 (0.047)

Abbreviation: SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.
^a $p < 0.05$; $**p < 0.01$; $***p < 0.001$ by the Wald test (only for covariate effects).
^bNumbers in parentheses are the standard error.
^cLognormal distributions were assumed.
^dWith reference to the weighted mean age of 42.
^eNormal distributions were assumed with mean 0.

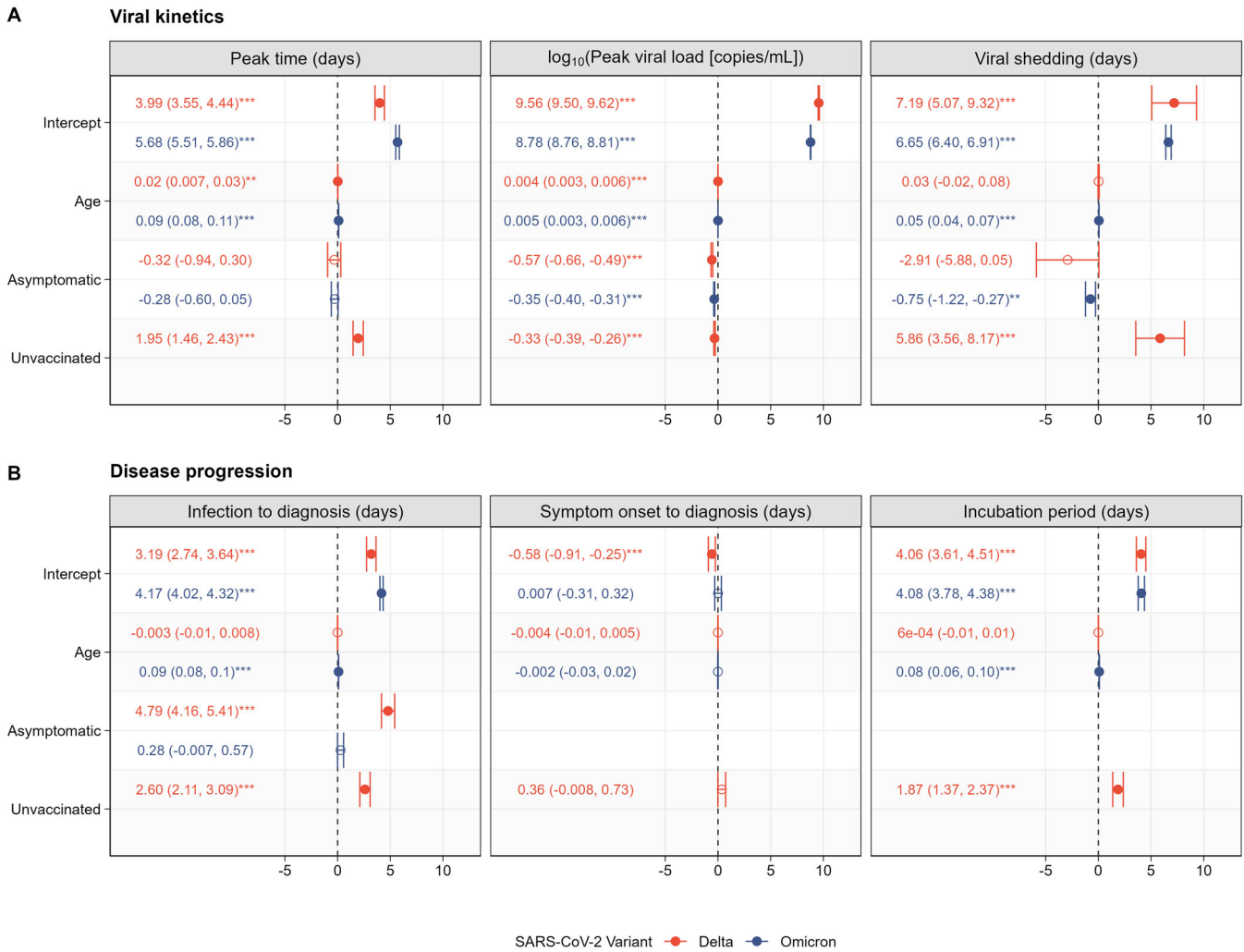


FIGURE 4 | Forest plots of the multivariate multiple linear regression results. Quantitative metrics of the viral dynamics related to (A) viral kinetics in terms of the timing of peak viral load (days), peak viral load (\log_{10} [copies/mL]) and duration of viral shedding (days), as well as (B) disease progression in terms of the time from infection to diagnosis (days), time from symptom onset to diagnosis (days), and incubation period (days) are shown as the dependent variables in each panel. All covariates considered in the analysis such as continuous age, symptom presence, and vaccination status are shown as independent variables on the y-axis. The reference group was set to an individual who is 42 years old, symptomatic, and vaccinated. Regression coefficients are reported as text and statistical significance was evaluated by the Wald test. Each regression coefficient is shown as point estimates (filled: significant, empty: nonsignificant) along the x-axis together with their corresponding 95% confidence intervals for both the SARS-CoV-2 Delta (red) and Omicron (blue) variants. Note that the time from symptom onset to diagnosis and incubation period are computed only for symptomatic patients by definition. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

Furthermore, model parameters for the Omicron variant were generally more precisely estimated because we assumed all Omicron patients to be fully vaccinated in our simulations, given the high vaccination rate and our model was calibrated without vaccination status as a covariate. This reduced the additional level of heterogeneity in the immune life history of patients and minimized the possible uncertainty during model fitting. However, post-peak observations are equally important in determining the overall model fit. For example, if the observed data is only restricted to three tests spaced only 1 day apart (i.e., bottom-left square of heatmaps in Figure 5), poorer model fits are obtained.

When aiming for an acceptable average error within 1.25 units of \log_{10} -transformed viral load while minimizing the follow-up period, we recommend a DCP consisting of 7 consecutive

measurements with a 3-day interval (21 days of follow-up) for 500 Delta patients. Conversely, for 500 Omicron patients, eight consecutive measurements with a 1-day interval (8 days of follow-up) are recommended. Furthermore, the necessary follow-up period decreases with larger patient cohorts.

4 | Discussion

In this study, we used a viral dynamics model to study population heterogeneity in viral dynamics. We found that the temporal viral load trajectory was influenced by age, symptom presence, and vaccination status. Specifically, older and unvaccinated individuals showed a later peak timing and a longer viral shedding duration. A higher peak viral load was observed for older, symptomatic, and vaccinated individuals.

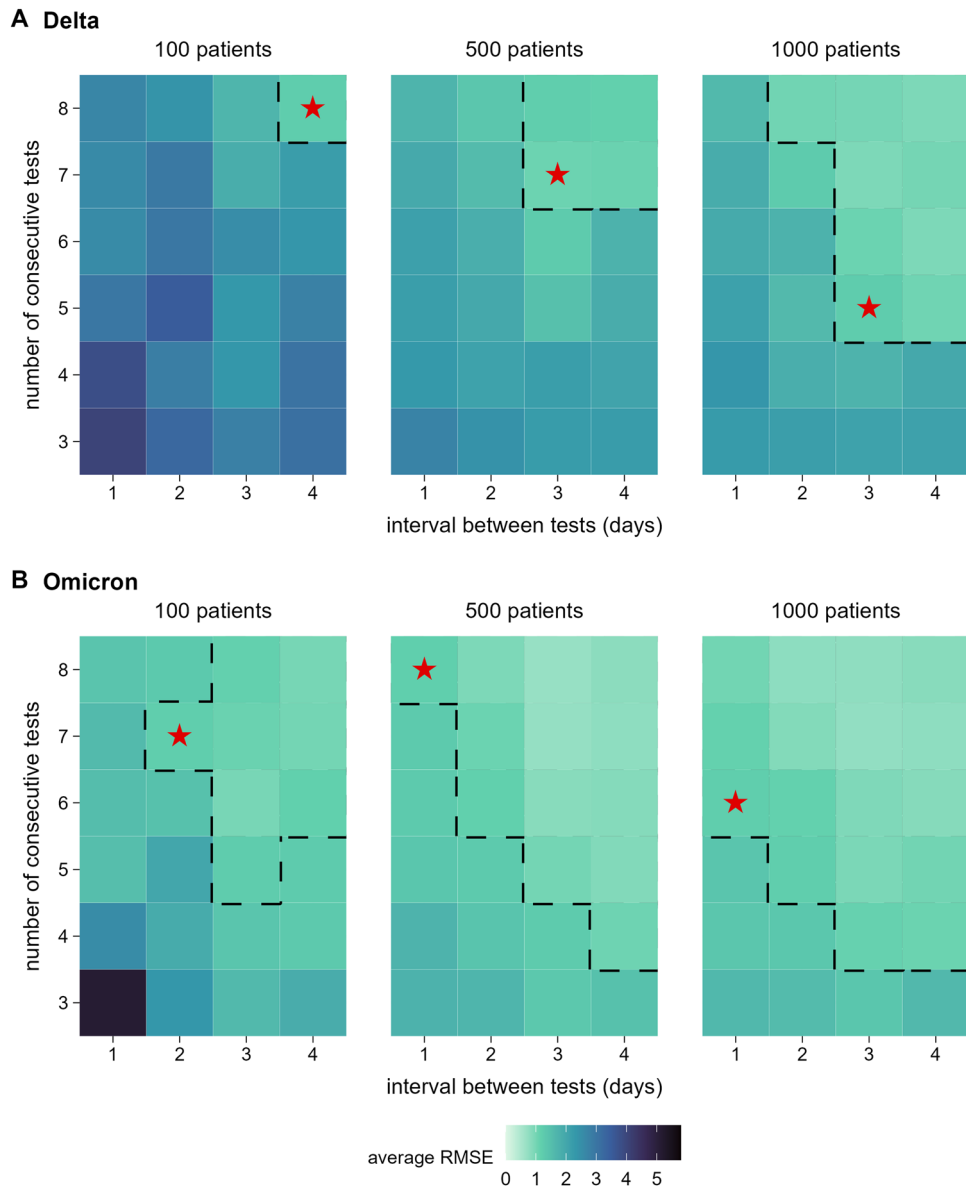


FIGURE 5 | Accuracy of estimating the viral load trajectories under different data collection protocols. Variations in the root mean squared error (RMSE) averaged over all individual fits for both the (A) Delta variant and (B) Omicron variant are shown as heatmaps. The RMSE represents the root mean squared deviation of the estimated best-fit curves from the true longitudinal viral load data determined under the calibrated model, and this varies with the number of consecutive tests and the interval between tests. The number of simulated virtual patients differ from left to right in each panel. The areas surrounded by dotted lines correspond to combinations that realize an RMSE value below 1.25 units of \log_{10} -transformed viral load, while the best combination that realizes the shortest duration of follow-up (i.e., the product of the number of consecutive tests and the interval between tests) is depicted by a red star.

However, both symptom presence and vaccination were associated with a shorter time from infection to diagnosis. Furthermore, we examined the suitability of different DCPs to accurately estimate viral dynamics. Our analysis reveals that to reconstruct individual viral dynamics and minimize the number of viral tests required for 500 Delta patients, a 21-day follow-up period with measurements every 3 days is recommended. For 500 Omicron patients, an 8-day follow-up period with consecutive daily measurements is recommended. Additionally, the required follow-up period decreases with increasing sample size. Our findings have profound multifaceted implications ranging from disease diagnosis to virus characteristics, and to public health policy.

From the biological point of view, viral dynamics is influenced by multiple factors, particularly immunological differences between patients. We observed a lower viral infection and replication rate with age, which is likely related to immunosenescence. The abrupt appearance of symptoms is thought to be caused either by viral cytopathic effects (as a consequence of the viral infection and replication rate), or the host immune response to infection mediated by cytokine signaling [30]. In line with earlier SARS-CoV-2 studies, we found that older patients are more likely to have a blunted immune response towards infection, leading to a delay in symptom onset and, in turn, a longer incubation period [31–33]. Moreover, one of the hallmarks of COVID-19 disease progression is immune

dysregulation, which may be more pronounced in older patients and likely contributes to severe COVID-19 outcomes. With age, there is a marked increase in the proportion of highly differentiated effector and memory T cells due to exposure to a variety of pathogens over time [34]. This accumulated antigenic burden leads to immunosenescence through continuous reshaping and eventual shrinkage of the immune repertoire [35], thus shifting the immune system towards an inflammatory Th2 cell profile. Our findings demonstrate that vaccination increases the death rate of infected cells because vaccination stimulates the immunological response to infection, although we did not detect any age-specific differences in vaccine responses. In addition, symptomatic patients had higher peak viral loads and a longer viral shedding duration compared to asymptomatic patients. The more viral stimuli exposed to over a longer period of time is likely why symptomatic patients have higher neutralizing antibody levels, although other studies have reported a shorter viral shedding duration for symptomatic patients [36]. Furthermore, consistent with a previous study [37], the Omicron variant induced a higher viral infection rate with lower peak viral loads, which might be due to immune escape of the virus.

From the public health perspective, vaccination is associated with a shorter viral shedding duration, implying that the vaccine is not only effective in providing individual immunological protection but also in mitigating forward transmission. Although our analysis shows that vaccinated individuals have a slightly higher peak viral load compared to unvaccinated individuals, the overall transmissibility may still be lower with a significant reduction in viral shedding duration (i.e., smaller area under the viral load curve above the infectiousness threshold). Moreover, among Delta patients, symptomatic cases were diagnosed significantly earlier than asymptomatic cases on average, highlighting the importance of identifying asymptomatic cases through viral tests regardless of symptom presence, especially since the peak viral load of asymptomatic cases is comparable to that of symptomatic cases [14]. For the Omicron variant, older age was associated with a longer viral shedding duration, suggesting that the isolation period could be shortened for younger patients.

The simulation investigating DCPs is beneficial in alleviating the burden on both patients and healthcare practitioners. Especially during a pandemic, prolonged follow-up periods pose additional stress on both parties. While collecting longitudinal viral load data was feasible in Singapore due to stringent isolation policies, implementing such measures might prove challenging during periods of surging cases. Therefore, our primary aim was to minimize the follow-up period while maintaining the error within a predetermined threshold, balancing the need for accurate DCPs with practical considerations.

We outline the main strengths of this study below. First, we incorporated viral load data collection into clinical routine to obtain consistent longitudinal data from the general population (i.e., individuals hospitalized for isolation purposes and not necessarily of clinical need) for more accurate estimates of the viral dynamics. Second, as viral load changes over the course of infection, we used a mechanistic modeling approach to fully

recover the viral load trajectory and were able to characterize quantitative aspects of viral dynamics that corroborated with the findings of earlier studies [8, 31, 37], given that most of the observed data occurred after the viral load peak. Third, we used a nonlinear mixed-effects model to capture both population heterogeneities in viral dynamics as well as interindividual variability which cannot be entirely explained by the covariates included in the analysis. Incorporating this additional variability enabled us to simulate synthetic longitudinal data that was representative of discrete observations at the population level, which gave us a unique opportunity to evaluate the effectiveness of different DCPs in accurately recapitulating the viral dynamics.

There are a few limitations in our analyses. First, although the target cell-limited model has been widely used to characterize SARS-CoV-2 viral dynamics, misspecifications might be present and additional factors including the immune response to viral infection may need to be incorporated for more biologically reasonable models. Nonetheless, as demonstrated in our previous study where we compared several models considering the effect of interferons or the eclipse phase of infection, the model used in this study holds an advantage in its simplicity and goodness-of-fit [2, 28]. Second, previous studies have shown possible associations between viral dynamics and the number of antigenic exposures, prior infection, immunological status, sex, use of antivirals, vaccination history (i.e., doses, types of vaccine and timing of vaccination), and underlying health conditions [8, 37, 38]. However, we could not investigate the effect of these variables due to limitations in the sample size or lack of information. In particular, we do not have access to antiviral treatment records for the patients analyzed in this study. During the study period, intravenous remdesivir was the only antiviral available in Singapore and it was accessible at no cost to patients. It is thus possible that some of the patients analyzed in this study had received remdesivir. Although previous literature showed a limited effect of remdesivir on SARS-CoV-2 viral kinetics [39, 40], it is hard to assess this impact on our results. Third, as symptom scores were self-reported, the estimated incubation period may not be truly accurate. Fourth, it is imperative to consider the costs and logistical factors involved in DCPs. It may be essential to establish dedicated facilities for patient follow-up. If patients are isolated at home, it becomes necessary to facilitate self-collection of samples and utilize courier services for the transport of samples to laboratories. While nasopharyngeal swabs have traditionally been the gold standard, recent evidence suggests that saliva samples are equally sensitive and easier to obtain [41, 42], and thus, may be a viable option for future data collection efforts.

In conclusion, we identified multiple factors influencing viral dynamics and appropriate DCPs to accurately estimate viral dynamics. These findings would deepen our understanding of the biology of SARS-CoV-2 infection and, at the same time, have strong public health implications, including the definition of isolation guidelines and the effect of vaccination in hampering forward transmission. Our analytic pipeline based on the interpretation of limited sparse viral load observations with a viral dynamics model is applicable in contexts beyond SARS-CoV-2 and will be instrumental to improve preparedness for the next pathogen with epidemic/pandemic potential.

Author Contributions

Conceived and designed the study: Keisuke Ejima, David Chien Lye, and Barnaby E. Young. Obtained and analyzed the data: Hoong Kai Chua, Ananya Singh, Yuqian Wang, Keisuke Ejima, David Chien Lye, and Barnaby E. Young. Wrote the paper: Hoong Kai Chua, Ananya Singh, Yuqian Wang, Yun Shan Goh, Marco Ajelli, Keisuke Ejima, and Barnaby E. Young. All authors read and approved the final manuscript.

Acknowledgments

This study was supported in part by the Ministry of Education, Singapore, under the Academic Research Fund Tier 1 Seed Award (RLMOE100201900000001), a Lee Kong Chian School of Medicine startup grant (LKCmedicine-SUG, #022487-00001), and JST, PRESTO (JPMJPR23R3) (to KE). The study does not necessarily represent the views of the funding agencies listed above.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data that support the findings of this study are available from the corresponding authors upon request. All analyses were performed with the statistical computing software R (version 4.3.3) (<https://www.r-project.org/>). The analysis using nonlinear mixed-effects modeling was performed on Monolix 2023R1 (<https://www.lixoft.com/>) and the corresponding R functions in the lixoftConnectors package. The study's supporting codes can be found on Zenodo at <https://doi.org/10.5281/zenodo.11481803>.

References

1. K. Ejima, K. S. Kim, A. I. Bento, et al., "Estimation of Timing of Infection From Longitudinal SARS-CoV-2 Viral Load Data: Mathematical Modelling Study," *BMC Infectious Diseases* 22, no. 1 (2022): 656, <https://doi.org/10.1186/s12879-022-07646-2>.
2. Y. D. Jeong, K. Ejima, K. S. Kim, et al., "Revisiting the Guidelines for Ending Isolation for COVID-19 Patients," *eLife* 10 (2021): e69340, <https://doi.org/10.7554/eLife.69340>.
3. Y. D. Jeong, K. Ejima, K. S. Kim, et al., "Designing Isolation Guidelines for COVID-19 Patients With Rapid Antigen Tests," *Nature Communications* 13, no. 1 (2022): 4910, <https://doi.org/10.1038/s41467-022-32663-9>.
4. J. A. Hay, L. Kennedy-Shaffer, S. Kanjilal, et al., "Estimating Epidemiologic Dynamics From Cross-Sectional Viral Load Distributions," *Science* 373, no. 6552 (2021): eabh0635, <https://doi.org/10.1126/science.abh0635>.
5. R. Magleby, L. F. Westblade, A. Trzebecki, et al., "Impact of Severe Acute Respiratory Syndrome Coronavirus 2 Viral Load on Risk of Intubation and Mortality Among Hospitalized Patients With Coronavirus Disease 2019," *Clinical Infectious Diseases* 73, no. 11 (2021): e4197–e4205, <https://doi.org/10.1093/cid/ciaa851>.
6. Y. Liu, L. M. Yan, L. Wan, et al., "Viral Dynamics in Mild and Severe Cases of COVID-19," *Lancet Infectious Diseases* 20, no. 6 (2020): 656–657, [https://doi.org/10.1016/s1473-3099\(20\)30232-2](https://doi.org/10.1016/s1473-3099(20)30232-2).
7. S. Zheng, J. Fan, F. Yu, et al., "Viral Load Dynamics and Disease Severity in Patients Infected With SARS-CoV-2 in Zhejiang Province, China, January–March 2020: Retrospective Cohort Study," *BMJ* 369 (2020): m1443, <https://doi.org/10.1136/bmj.m1443>.
8. O. Puhach, B. Meyer, and I. Eckerle, "SARS-CoV-2 Viral Load and Shedding Kinetics," *Nature Reviews Microbiology* 21, no. 3 (2023): 147–161, <https://doi.org/10.1038/s41579-022-00822-w>.
9. K. Ejima, K. S. Kim, S. Iwanami, et al., "Time Variation in the Probability of Failing to Detect a Case of Polymerase Chain Reaction Testing for SARS-CoV-2 as Estimated From a Viral Dynamics Model," *Journal of The Royal Society Interface* 18, no. 177 (2021): 20200947, <https://doi.org/10.1098/rsif.2020.0947>.
10. L. M. Kucirka, S. A. Lauer, O. Laeyendecker, D. Boon, and J. Lessler, "Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time Since Exposure," *Annals of Internal Medicine* 173, no. 4 (2020): 262–267, <https://doi.org/10.7326/m20-1495>.
11. H. Kang, Y. Wang, Z. Tong, and X. Liu, "Retest Positive for SARS-CoV-2 RNA of 'Recovered' Patients With COVID-19: Persistence, Sampling Issues, or Re-Infection," *Journal of Medical Virology* 92, no. 11 (2020): 2263–2265, <https://doi.org/10.1002/jmv.26114>.
12. P. Mancuso, F. Venturelli, M. Vicentini, et al., "Temporal Profile and Determinants of Viral Shedding and of Viral Clearance Confirmation on Nasopharyngeal Swabs From SARS-CoV-2-Positive Subjects: A Population-Based Prospective Cohort Study in Reggio Emilia, Italy," *BMJ Open* 10, no. 8 (2020): e040380, <https://doi.org/10.1136/bmjopen-2020-040380>.
13. H. Ye, C. Zhao, L. Yang, et al., "Twelve Out of 117 Recovered COVID-19 Patients Retest Positive in a Single-Center Study of China," *eClinicalMedicine* 26 (2020): 100492, <https://doi.org/10.1016/j.eclinm.2020.100492>.
14. S. Lee, T. Kim, E. Lee, et al., "Clinical Course and Molecular Viral Shedding Among Asymptomatic and Symptomatic Patients With SARS-CoV-2 Infection in a Community Treatment Center in the Republic of Korea," *JAMA Internal Medicine* 180, no. 11 (2020): 1447–1452, <https://doi.org/10.1001/jamainternmed.2020.3862>.
15. C. Bennisrallah, I. Zemni, W. Dhoubi, et al., "Factors Associated With a Prolonged Negative Conversion of Viral RNA in Patients With COVID-19," *International Journal of Infectious Diseases* 105 (2021): 463–469, <https://doi.org/10.1016/j.ijid.2021.02.089>.
16. K. S. Kim, S. Iwanami, T. Oda, et al., "Incomplete Antiviral Treatment May Induce Longer Durations of Viral Shedding During SARS-CoV-2 Infection," *Life Science Alliance* 4, no. 10 (2021): e202101049, <https://doi.org/10.26508/lsa.202101049>.
17. S. M. Kissler, J. R. Fauver, C. Mack, et al., "Viral Dynamics of Acute SARS-CoV-2 Infection and Applications to Diagnostic and Public Health Strategies," *PLOS Biology* 19, no. 7 (2021): e3001333, <https://doi.org/10.1371/journal.pbio.3001333>.
18. N. Néant, G. Lingas, Q. Le Hingrat, et al., "Modeling SARS-CoV-2 Viral Kinetics and Association With Mortality in Hospitalized Patients From the French Covid Cohort," *Proceedings of the National Academy of Sciences of the United States of America* 118, no. 8 (2021): e2017962118, <https://doi.org/10.1073/pnas.2017962118>.
19. R. Ke, P. P. Martinez, R. L. Smith, et al., "Daily Longitudinal Sampling of SARS-CoV-2 Infection Reveals Substantial Heterogeneity in Infectiousness," *Nature Microbiology* 7, no. 5 (2022): 640–652, <https://doi.org/10.1038/s41564-022-01105-z>.
20. R. Ke, C. Zitzmann, D. D. Ho, R. M. Ribeiro, and A. S. Perelson, "In Vivo Kinetics of SARS-CoV-2 Infection and Its Relationship With a Person's Infectiousness," *Proceedings of the National Academy of Sciences of the United States of America* 118, no. 49 (2021): e2111477118, <https://doi.org/10.1073/pnas.2111477118>.
21. O. Dadras, A. M. Afsahi, Z. Pashaei, et al., "The Relationship Between COVID-19 Viral Load and Disease Severity: A Systematic Review," *Immunity, Inflammation and Disease* 10, no. 3 (2022): e580, <https://doi.org/10.1002/iid3.580>.
22. C. Zitzmann, R. Ke, R. M. Ribeiro, and A. S. Perelson, "How Robust Are Estimates of Key Parameters in Standard Viral Dynamic Models?," *PLOS Computational Biology* 20, no. 4 (2024): e1011437, <https://doi.org/10.1371/journal.pcbi.1011437>.

23. P. Y. Chia, S. W. X. Ong, C. J. Chiew, et al., “Virological and Serological Kinetics of SARS-CoV-2 Delta Variant Vaccine Break-through Infections: A Multicentre Cohort Study,” *Clinical Microbiology and Infection* 28, no. 4 (2022): 612.e1–612.e7, <https://doi.org/10.1016/j.cmi.2021.11.010>.
24. L. Zou, F. Ruan, M. Huang, et al., “SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients,” *New England Journal of Medicine* 382, no. 12 (2020): 1177–1179, <https://doi.org/10.1056/NEJMc2001737>.
25. B. E. Young, S. W. X. Ong, L. F. P. Ng, et al., “Viral Dynamics and Immune Correlates of Coronavirus Disease 2019 (COVID-19) Severity,” *Clinical Infectious Diseases* 73, no. 9 (2020): e2932–e2942, <https://doi.org/10.1093/cid/ciaa1280>.
26. Ministry of Health Singapore, “Progress of COVID-19 Vaccination,” accessed April 30, 2024, <https://www.beta.data.gov.sg/datasets>.
27. K. Ejima, K. S. Kim, C. Ludema, et al., “Estimation of the Incubation Period of COVID-19 Using Viral Load Data,” *Epidemics* 35 (2021): 100454, <https://doi.org/10.1016/j.epidem.2021.100454>.
28. S. Iwanami, K. Ejima, K. S. Kim, et al., “Detection of Significant Antiviral Drug Effects on COVID-19 With Reasonable Sample Sizes in Randomized Controlled Trials: A Modeling Study,” *PLOS Medicine* 18, no. 7 (2021): e1003660, <https://doi.org/10.1371/journal.pmed.1003660>.
29. K. S. Kim, K. Ejima, S. Iwanami, et al., “A Quantitative Model Used to Compare Within-Host SARS-CoV-2, MERS-CoV, and SARS-CoV Dynamics Provides Insights Into the Pathogenesis and Treatment of SARS-CoV-2,” *PLOS Biology* 19, no. 3 (2021): e3001128, <https://doi.org/10.1371/journal.pbio.3001128>.
30. T. Hermesh, B. Moltedo, C. B. López, and T. M. Moran, “Buying Time—The Immune System Determinants of the Incubation Period to Respiratory Viruses,” *Viruses* 2, no. 11 (2010): 2541–2558, <https://doi.org/10.3390/v2112541>.
31. S. Galmiche, T. Cortier, T. Charmet, et al., “SARS-CoV-2 Incubation Period Across Variants of Concern, Individual Factors, and Circumstances of Infection in France: A Case Series Analysis From the Com-Cor Study,” *Lancet Microbe* 4, no. 6 (2023): e409–e417, [https://doi.org/10.1016/S2666-5247\(23\)00005-8](https://doi.org/10.1016/S2666-5247(23)00005-8).
32. W. Y. T. Tan, L. Y. Wong, Y. S. Leo, and M. P. H. S. Toh, “Does Incubation Period of COVID-19 Vary With Age? A Study of Epidemiologically Linked Cases in Singapore,” *Epidemiology and Infection* 148 (2020): e197, <https://doi.org/10.1017/S0950268820001995>.
33. T. Kong, “Longer Incubation Period of Coronavirus Disease 2019 (COVID-19) in Older Adults,” *Aging Medicine* 3, no. 2 (2020): 102–109, <https://doi.org/10.1002/agm2.12114>.
34. R. Alonso-Arias, M. A. Moro-García, A. López-Vázquez, et al., “NKG2D Expression in CD4+ T Lymphocytes as a Marker of Senescence in the Aged Immune System,” *Age* 33, no. 4 (2011): 591–605, <https://doi.org/10.1007/s11357-010-9200-6>.
35. F. Hassouneh, N. Lopez-Sejas, C. Campos, et al., “Differential Effect of Cytomegalovirus Infection With Age on the Expression of CD57, CD300a, and CD161 on T-Cell Subpopulations,” *Frontiers in Immunology* 8 (2017): 649, <https://doi.org/10.3389/fimmu.2017.00649>.
36. Q.-X. Long, X.-J. Tang, Q.-L. Shi, et al., “Clinical and Immunological Assessment of Asymptomatic SARS-CoV-2 Infections,” *Nature Medicine* 26, no. 8 (2020): 1200–1204, <https://doi.org/10.1038/s41591-020-0965-6>.
37. Y. Yang, L. Guo, J. Yuan, et al., “Viral and Antibody Dynamics of Acute Infection With SARS-CoV-2 Omicron Variant (B.1.1.529): A Prospective Cohort Study From Shenzhen, China,” *Lancet Microbe* 4, no. 8 (2023): e632–e641, [https://doi.org/10.1016/S2666-5247\(23\)00139-8](https://doi.org/10.1016/S2666-5247(23)00139-8).
38. S. M. Kissler, J. A. Hay, J. R. Fauver, et al., “Viral Kinetics of Sequential SARS-CoV-2 Infections,” *Nature Communications* 14, no. 1 (2023): 6206, <https://doi.org/10.1038/s41467-023-41941-z>.
39. K. Hagman, M. Hedenstierna, J. Widaeus, et al., “Effects of Remdesivir on SARS-CoV-2 Viral Dynamics and Mortality in Viraemic Patients Hospitalized for COVID-19,” *Journal of Antimicrobial Chemotherapy* 78, no. 11 (2023): 2735–2742, <https://doi.org/10.1093/jac/dkad295>.
40. I. Faghihi and V. C. Yan, “Remdesivir Treatment Does Not Reduce Viral Titers in Patients With COVID-19,” *Antimicrobial Agents and Chemotherapy* 68, no. 8 (2024): e00856–24, <https://doi.org/10.1128/aac.00856-24>.
41. L. Azzi, G. Carcano, F. Gianfagna, et al., “Saliva Is a Reliable Tool to Detect SARS-CoV-2,” *Journal of Infection* 81, no. 1 (2020): e45–e50, <https://doi.org/10.1016/j.jinf.2020.04.005>.
42. A. L. Wyllie, J. Fournier, A. Casanovas-Massana, et al., “Saliva or Nasopharyngeal Swab Specimens for Detection of SARS-CoV-2,” *New England Journal of Medicine* 383, no. 13 (2020): 1283–1286, <https://doi.org/10.1056/NEJMc2016359>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.