

Daily sampling of early SARS-CoV-2 infection reveals substantial heterogeneity in infectiousness

Ruian Ke¹, Pamela P. Martinez^{2,3,4}, Rebecca L. Smith^{4,5,6}, Laura L. Gibson⁷, Agha Mirza⁸, Madison Conte⁸, Nicholas Gallagher⁹, Chun Huai Luo⁹, Junko Jarrett⁹, Abigail Conte¹⁰, Tongyu Liu², Mireille Farjo², Kimberly K.O. Walden¹¹, Gloria Rendon¹¹, Christopher J. Fields¹¹, Leyi Wang¹², Richard Fredrickson¹², Darci C. Edmonson⁴, Melinda E. Baughman⁴, Karen K. Chiu⁴, Hannah Choi⁴, Kevin R. Scardina⁴, Shannon Bradley⁴, Stacy L. Gloss⁴, Crystal Reinhart⁴, Jagadeesh Yedetore⁴, Jessica Quicksall⁴, Alyssa N. Owens¹³, John Broach^{14,15}, Bruce Barton^{16,17}, Peter Lazar¹⁶, William J. Heetderks¹⁸, Matthew L. Robinson⁸, Heba H. Mostafa⁹, Yukari C. Manabe^{8,10}, Andrew Pekosz¹⁰, David D. McManus¹⁹, Christopher B. Brooke^{2,4*}

¹ T-6, Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

² Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

³ Department of Statistics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

⁴ Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

⁵ Department of Pathobiology, University of Illinois at Urbana-Champaign, Urbana, IL 61802, USA

⁶ Carle Illinois College of Medicine, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

⁷ Division of Infectious Diseases and Immunology, Departments of Medicine and Pediatrics, University of Massachusetts Medical School, Worcester, MA 01655, USA

⁸ Division of Infectious Diseases, Department of Medicine, Johns Hopkins School of Medicine, Baltimore, MD 21287, USA

⁹ Division of Medical Microbiology, Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA

¹⁰ W. Harry Feinstone Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA

¹¹ High-Performance Biological Computing at the Roy J. Carver Biotechnology Center, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

¹² Veterinary Diagnostic Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61802, USA

¹³ Center for Clinical and Translational Research, University of Massachusetts Medical School, Worcester, MA 01655, USA

¹⁴ UMass Memorial Medical Center, Worcester, MA 01655, USA

¹⁵ Department of Emergency Medicine, University of Massachusetts Medical School, Worcester, MA 01655, USA

¹⁶ Division of Biostatistics and Health Services Research, University of Massachusetts Medical School, Worcester, MA 01655, USA

¹⁷ Department of Population and Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA 01655, USA

¹⁸ National Institute for Biomedical Imaging and Bioengineering, Bethesda, MD 20892, USA

¹⁹ Division of Cardiology, University of Massachusetts Medical School, Worcester, MA 01655, USA

* Corresponding author (cbrooke@illinois.edu)

ABSTRACT:

The dynamics of SARS-CoV-2 replication and shedding in humans remain poorly understood. We captured the dynamics of infectious virus and viral RNA shedding during acute infection through daily longitudinal sampling of 60 individuals for up to 14 days. By fitting mechanistic models, we directly estimate viral reproduction and clearance rates, and overall infectiousness for each individual. Significant person-to-person variation in infectious virus shedding suggests that individual-level heterogeneity in viral dynamics contributes to superspreading. Viral genome load often peaked days earlier in saliva than in nasal swabs, indicating strong compartmentalization and suggesting that saliva may serve as a superior sampling site for early detection of infection. Viral loads and clearance kinetics of B.1.1.7 and non-B.1.1.7 viruses in

nasal swabs were indistinguishable, however B.1.1.7 exhibited a significantly slower pre-peak growth rate in saliva. These results provide a high-resolution portrait of SARS-CoV-2 infection dynamics and implicate individual-level heterogeneity in infectiousness in superspreading.

MAIN TEXT:

Transmission of SARS-CoV-2 by pre-symptomatic and asymptomatic individuals has been a major contributor to the explosive spread of this virus (1–5). Recent epidemiological investigations of community outbreaks have indicated that transmission of SARS-CoV-2 is highly heterogeneous, with a small fraction of infected individuals (often referred to as “superspreaders”) contributing a disproportionate share of forward transmission (6–8). Transmission heterogeneity has also been implicated in the epidemic spread of several other important viral pathogens, including measles and smallpox (9). Numerous behavioral and environmental explanations have been offered to explain transmission heterogeneity, but the extent to which the underlying features of the infection process within individual hosts contribute towards the superspreading phenomenon remains unclear. Addressing this gap in knowledge will inform the design of more targeted and effective strategies for controlling community spread.

Viral infection is a highly complex process in which viral replication and shedding dynamics are shaped by the complex interplay between host and viral factors. Recent studies have suggested that the magnitude and/or duration of viral shedding in both nasal and saliva samples correlate with disease severity, highlighting the potential importance of viral dynamics in influencing infection outcomes (10–13). Variation in viral load has also been suggested to correlate with transmission risk (14). In addition to implications for pathogenesis and transmission, defining the contours of viral shedding dynamics is also critical for designing effective surveillance, screening, and testing strategies (15). To date, studies aimed at describing the longitudinal dynamics of SARS-CoV-2 shedding have been limited by (a) sparse sampling frequency, (b) failure to capture the early stages of infection when transmission is most likely, (c) absence of individual-level data on infectious virus shedding kinetics, and (d) biasing towards the most severe clinical outcomes (16–21). This is also true for viruses beyond SARS-CoV-2, as the dynamics of natural infection in humans have not been described in detail for any acute viral pathogen.

Here, we capture the early longitudinal viral dynamics of mild and asymptomatic acute SARS-CoV-2 infection in 60 people by recording daily measurements of both viral RNA shedding (from mid-turbinate nasal swabs and saliva samples) and infectious virus shedding (from mid-turbinate nasal swabs) for up to 14 days. We reveal a striking degree of individual-level heterogeneity in infectious virus shedding between individuals, thus providing a partial explanation for the central role of superspreaders in community transmission of SARS-CoV-2. We also directly compare the shedding dynamics of B.1.1.7 and non-B.1.1.7 viruses, revealing prolonged pre-peak shedding in saliva but no significant differences in nasal shedding. Altogether, these results provide the first high-resolution, multi-parameter empirical profile of acute SARS-CoV-2 infection in humans and implicate person-to-person variation in infectious virus shedding in driving patterns of epidemiological spread of the pandemic.

Results:

Description of cohort and study design

During the Fall of 2020 and Spring of 2021, all faculty, staff, and students at the University of Illinois at Urbana-Champaign were required to undergo at least twice weekly RTqPCR testing for SARS-CoV-2. We leveraged this large-scale, high-frequency screening program to enroll symptomatic, pre-symptomatic, and asymptomatic SARS-CoV-2-infected individuals. We enrolled university faculty, staff, and students who had a documented negative RTqPCR test

result in the past seven days and were either (a) within 24 hours of a positive RTqPCR result, or (b) within five days of exposure to someone with a confirmed positive RTqPCR result. These criteria ensured that we only enrolled people within the first days of infection.

We collected both nasal and saliva samples daily for up to 14 days to generate a high-resolution portrait of viral dynamics during the early stages of SARS-CoV-2 infection. Participants also completed a daily online symptom survey. Our study cohort was primarily young (median age: 28; range: 19-73), non-Hispanic white, and skewed slightly towards males (**Table S1**). All infections were either mild or asymptomatic, and none of the participants were ever hospitalized for COVID-19.

Early SARS-CoV-2 infection dynamics vary significantly between individuals

To examine viral dynamics at the individual level, we plotted Ct/Cn values from both saliva and nasal samples (the RTqPCR assay used for nasal samples reports Cn rather than Ct values), Quidel SARS Sofia 2 antigen FIA results, and viral culture data from nasal swabs, as a function of time relative to observed peak in nasal viral genome load (**Fig 1A, S1**). In many cases, we captured both the rise and fall of viral genome shedding in nasal and/or saliva samples. A comparison between individuals revealed substantial heterogeneity in shedding dynamics, with obvious differences in the duration of detectable infectious virus shedding, clearance kinetics, and the temporal relationship between shedding in nasal and saliva compartments. Further, a small subset of individuals (9 out of 60 examined) exhibited only low-level sporadic RNA shedding, with no detectable infectious virus in nasal samples.

Generally, earlier positivity results in the viral culture assay (which suggests higher infectious viral loads) were associated with lower Ct values in nasal samples (**Fig 1B**). This is unsurprising, as both nasal viral genome load and viral infectivity were assayed using the same sample. Saliva Ct values tended to be higher than matched nasal samples, likely due in part to the lower molecular sensitivity of the specific saliva RTqPCR assay used which does not include an RNA extraction step (22). For both sample types, the relationship between viral culture results and Ct values was not absolute, as several nasal swab samples with Ct values greater than 30 also tested positive for infectious virus. These data indicate that caution must be exercised when using a simple Ct value cutoff as a surrogate for infectious status.

We also assessed the relationship between antigen FIA and viral culture results and found that participants tested positive by antigen FIA on 93% of the days on which they also tested positive by viral culture (**Fig 1C**). This finding is consistent with earlier cross-sectional studies examining the relationship between antigen tests positivity and infectious virus shedding (23, 24).

While the symptom profiles self-reported by study participants varied widely across individuals, all cases were mild and did not require medical treatment (**Fig S2**). To determine whether any specific symptoms correlated with viral culture positivity, we compared the reported frequencies for each symptom on days where individuals tested viral culture positive or negative (**Fig S3**). Muscle aches, runny nose, and scratchy throat were significantly more likely to be reported on days when participants were viral culture positive, suggesting these specific symptoms as potential indicators of infectious status. No other symptoms examined exhibited a clear association with viral culture status. Self-reported symptom data from this study may be partially skewed by having been collected after participants were notified of their initial positive test result or potential exposure.

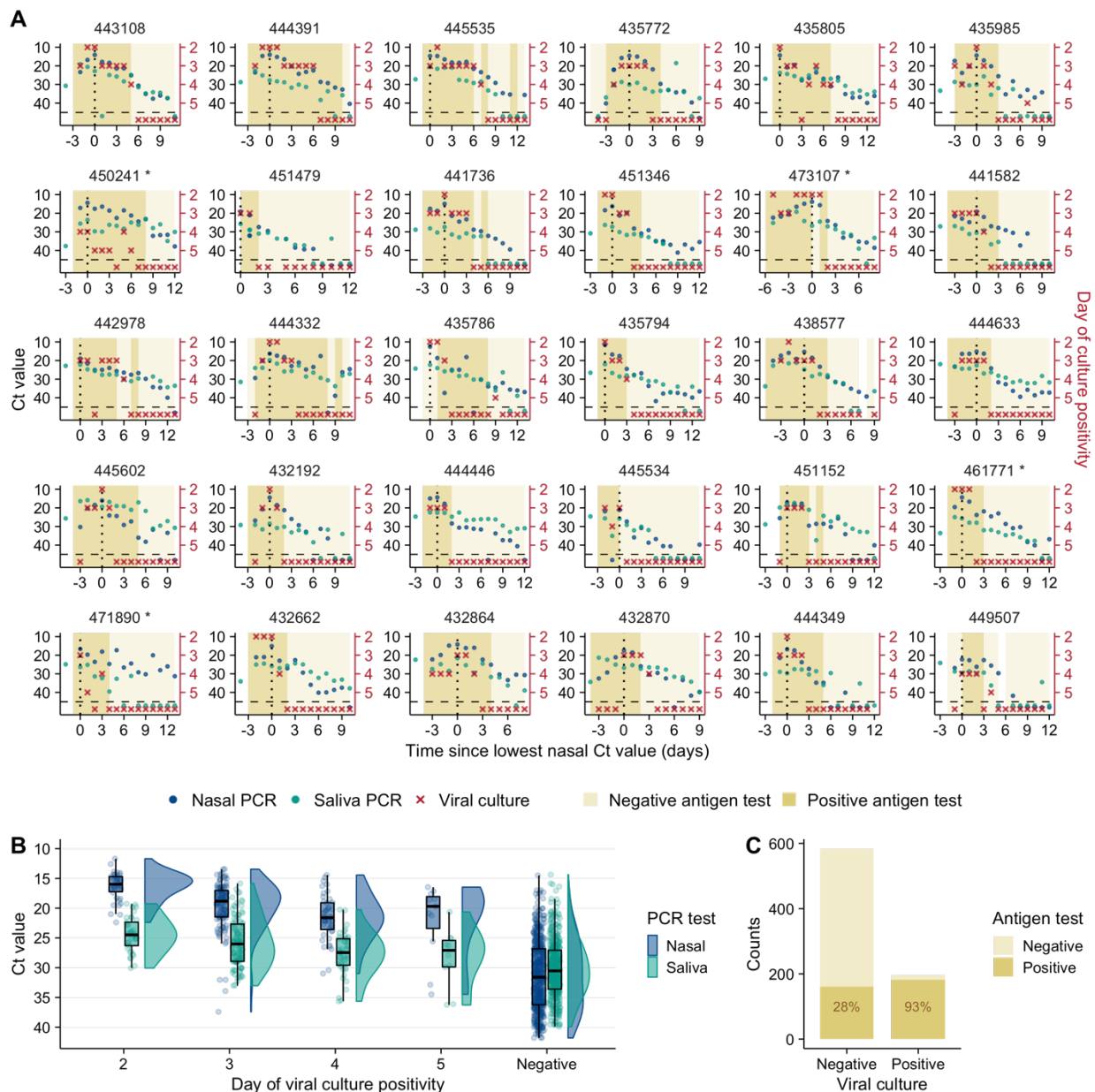


Figure 1: SARS-CoV-2 viral dynamics captured through daily sampling. (A) Temporal trends for the saliva RTqPCR (teal dots), nasal swab RTqPCR (navy blue dots), nasal swab viral culture (red crosses), and positive nasal swab antigen test results (dark mustard shaded area). The X axis shows days since the measured peak in nasal swab viral genome load, which is indicated by a vertical dotted line. The left Y axis indicates Ct values for saliva RTqPCR assay (covidSHIELD) and Cn values for nasal swab RTqPCR assay (Abbott Alinity). The right Y axis indicates results of viral culture assays, where day of culture positivity indicates the day of incubation at which > 50% of Vero-TMPRSS2 cells infected with the sample were positive for cytopathic effect. Horizontal dashed line indicates limit of detection of RTqPCR and viral culture assays. Title of each plot indicates the participant ID for the top 30 individuals with the most data points (the remaining 30 participants are shown in figure S1). Single asterisk next to the participant ID indicates B.1.1.7 variant infection. **(B)** Individual Ct (for saliva) and Cn (for nasal swabs) values from samples plotted based on concurrent results from the viral culture assay.

Negative indicates samples for which viral culture assay showed no viral growth after 5 days. Note the Y axes are reversed in panel (A) and (B). (C) Plot showing antigen FIA results from days where participants tested either positive or negative by viral culture. The text inside the bars indicates the percentage of antigen FIA results that were positive when concurrent viral culture sample was positive or negative.

Within-host mechanistic models capture viral dynamics in nasal and saliva samples

To better quantify the specific features of viral dynamics within individuals, we implemented five within-host models based on models developed previously for SARS-CoV-2 and influenza infection (see Supporting Text for details) (**Figs 2A, S4**) (25, 26). We fit these models to viral genome loads derived from RTqPCR Ct/Cn values using a population mixed effect modeling approach (see Supporting Text). The viral dynamics in nasal and saliva samples were distinct from each other in most individuals, indicating strong compartmentalization of the oral and nasal cavities. We thus fit the models to data from nasal and saliva samples separately. For either nasal or saliva samples, we selected data from 56 out of 60 individuals for model fitting. For each sample type, data from 4 individuals were excluded because viral genome loads were very low throughout the sampling period.

To identify factors that might partially explain the observed variation in individual-level dynamics, for each model we tested whether the age of the participants or the infecting viral genotype (*i.e.* non-B.1.1.7 vs. B.1.1.7) co-vary with any of the estimated model parameters in the model fitting. A total of 107 model variations were tested (see Supporting text for the testing strategy). We compared the relative abilities of these model variations to capture the RTqPCR data using the Akaike Information Criterion (AIC) and found that in general, the refractory model and the effector cell model best describe data from nasal and saliva samples, respectively (**Table S2, S3**). In the refractory model (**Fig 2A**), we assumed that target cells can be rendered refractory to infection through the activity of soluble immune mediators released by infected cells, such as interferon (27). In the best-fit immune effector cell model (**Fig 2A**), we assumed that innate and adaptive immune cells are activated and recruited to eliminate infected cells, leading to increased viral clearance (26). We estimated that all fitted parameters with a random effect differ between non-B.1.1.7 and B.1.1.7 infections. See **Tables S4, S5 and S6** for estimated values of the population and individual parameters and the fixed parameter values, respectively. Overall, these models described the viral genome load data in both nasal and saliva samples very well (**Fig 2B**).

The frequent longitudinal sampling of participants during early infection provided a unique opportunity to precisely quantify viral load kinetics during the viral expansion phase, prior to the peak in genome shedding. We estimated the mean exponential expansion rate r (growth rate, for short) to be 5.4/day (SD: ± 0.98 /day) in nasal swabs and 4.8/day (SD: ± 1.6 /day) in saliva (**Fig 2C-D**). Note that as we will show below, the growth rate of the B.1.1.7 variant is lower than that of non-B.1.1.7 viruses in saliva, leading to a bimodal distribution (**Fig 2D**).

Viral clearance kinetics clearly differed between nasal and saliva samples (**Fig 2B-D**). For nasal samples, viral genome loads decreased rapidly after peak, with an estimated death rate of productively infected cells at 4.0/day (SD: ± 0.6 /day); however, the rate of viral decline decreased over time. In saliva, post-peak viral genome loads declined initially at a slower rate on average (0.9/day; SD: ± 0.3 /day) than the decline in nasal samples. However, our model suggested the existence of a second clearance phase with a more rapid decline, potentially due to the onset of effector cell and/or neutralizing antibody responses.

Interestingly, the model predicts a significant correlation ($p < 0.01$) in nasal samples between age and the Φ parameter which describes the effectiveness of the anti-viral immune response in rendering target cells refractory to infection (**Fig 2E**). This suggests that innate immune responses are less effective at limiting SARS-CoV-2 in the nasal compartment of older individuals within our cohort, consistent with prior studies describing dysregulation of innate immunity to viral infection in aged individuals (28–30). There was no significant correlation between age and either growth rate or clearance rate in nasal samples. In saliva, none of the individual-level model parameter estimates correlated with participant age (**Fig S5**).

Overall, we noted a surprising degree of discordance in viral dynamics between nasal and saliva samples for many participants. In most individuals (41 out of 54 analyzed), viral genome shedding peaked at least 1 day earlier in saliva than in nasal samples (**Fig 2F**). In contrast, the peak in nasal shedding only preceded the saliva peak by at least 1 day in 3 individuals.

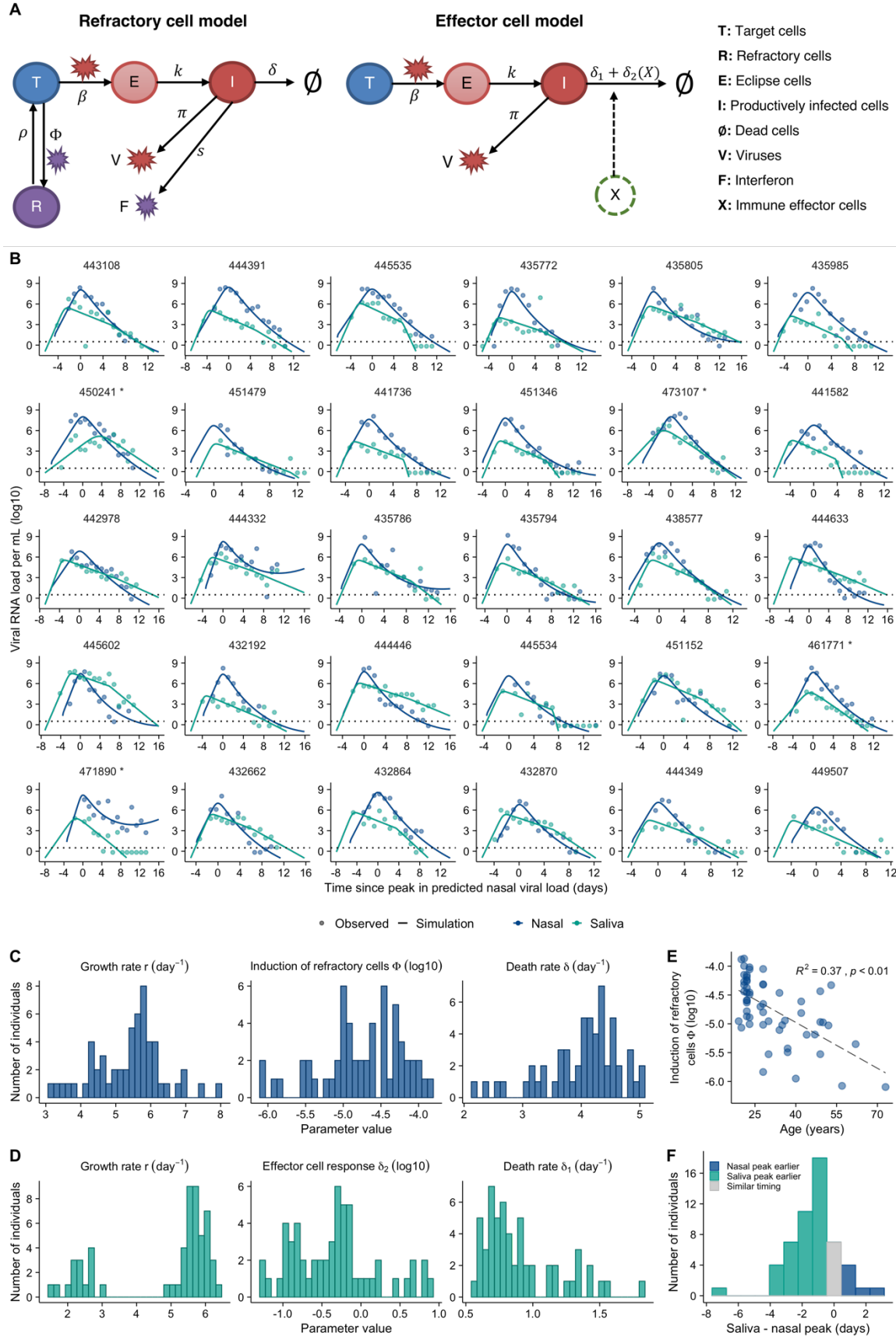


Figure 2: Model fits quantify heterogeneity in viral dynamics and discordance in genome shedding between nasal and saliva samples. (A) Diagrams outlining structures of the refractory cell model and the immune effector cell model that best fit the nasal swab and saliva RTqPCR data, respectively. In the refractory cell model, target cells (T) are infected by viruses (V) at rate β . Infected cells first become eclipse cells (E) and do not produce viruses; at rate k , eclipse cells become productively infected cells (I) producing both viruses and interferon (F) at rate π and s , respectively. They die at rate δ . Binding of interferons with target cells induces an antiviral response that converts target cells into cells refractory to infection (R). The rate of induction of the antiviral response is Φ . Refractory cells can revert to target cells at rate ρ . In the effector cell model, instead of modeling the interferon response explicitly, we assume that over the course of infection, immune effector cells (X) that clear infected cells are activated and recruited. This leads to an increase in infected cell death rate from δ_1 to $\delta_1 + \delta_2$. **(B)** Model fits to nasal sample (navy blue) and saliva (teal) RTqPCR results from the same subset of individuals shown in figure 1A. Dotted lines represent the limit of detection for the RTqPCR assays. The Y axis is shown on a log 10 scale. **(C)** Distributions of key individual parameter estimates from the refractory cell model (nasal data) across individuals in this cohort. **(D)** Distributions of key individual parameter estimates from the immune effector cell model (saliva data). **(E)** Association between age and the estimated strength of the antiviral immune response (Φ) based on nasal sample data. The Y axis is shown on a log 10 scale. **(F)** Distribution of the differences in the estimated times of peak viral genome loads between the saliva and nasal samples. Nasal sample peak is set to time=0 for all participants. Bars colored teal and navy blue represent estimated saliva peaks that occur at least 0.5 day earlier and later than nasal samples, respectively. The gray bar indicates the number of individuals that have similar timing in the peaks.

Significant heterogeneity in the infectious potential of individuals

We next examined the duration of infectious virus shedding in nasal samples, as a surrogate for the infectious potential of an individual; 51 out of 60 individuals tested positive by viral culture on nasal swabs at least once during the sampling period (**Fig 3A**). We found a weak positive correlation between the duration of viral culture positivity and participant age (**Fig 3B**). Of note, many study participants were viral culture positive on the first day of sample collection, suggesting that we failed to capture the onset of viral culture positivity for these individuals and thus may be underestimating the duration of infectious virus shedding for a subset of study participants.

To better quantify the infectious potential of each individual, we first used the viral culture data as a measure for intrinsic infectiousness (infectiousness for short below) to characterize how infectiousness depends on the viral genome load. We fitted three alternative models as proposed before (31) to the paired nasal RTqPCR and viral culture data collected from each individual using a non-linear mixed effect modeling approach (see **Fig. S6** for a workflow and Supporting Text for more details). Comparing models using the AIC scores, we found that the relationship is best described by a saturation model where the infectious virus load is a Hill-type function of the viral genome load (**Fig 3C, S7 and Table S7**). See **Table S8** for the best-fit parameter values.

Using the best-fit model, we estimated the infectiousness of each individual over the course of infection (**Fig. S8**). Note that the dataset only allows us to estimate a quantity that is a constant proportion of the infectious virus load (rather than its absolute value) across time and between individuals, thus we report the predicted the values in arbitrary units (a.u.) as a relative measure

of infectiousness. Our model predicts that infectious virus shedding increases sharply when viral genome loads rise above approximately 10^5 to 10^6 copies/mL (**Fig. 3D**). Importantly, there exists a high level of heterogeneity in infectiousness across different individuals that is not fully explained by differences in viral genome load (**Fig. 3D**). For example, at a viral genome load of 10^8 copies/mL, infectious virus shedding is higher than 10 a.u. in some individuals; while in other individuals it is below 1 a.u. This suggests that viral genome loads are not precisely predictive of infectiousness.

We next estimated the total infectiousness of each individual by integrating the area under the infectious virus load curve over the course of infection. This approach again revealed a large degree of heterogeneity in individual-level infectiousness, with a more than 30-fold difference between the highest and lowest estimated infectiousness (65.9 and 2.1 a.u., respectively) (**Fig 3E**). These data suggest that the previously reported heterogeneity in secondary transmission rates (6, 7) can arise from heterogeneity in intrinsic infectiousness in addition to heterogeneity in contact structure. We then fit a gamma distribution to the individual-level infectiousness data and found that a distribution with a shape parameter of 2 describes the majority of the individuals well, with the exception of one individual who had the highest level of infectiousness (**Fig 3E**). This again emphasizes the potential for a small subset of individuals that exhibit high intrinsic infectiousness to function as ‘superspreaders’ if they have frequent and/or high-risk contacts during the infectious period. Finally, we observed a significant correlation between age and total infectiousness ($p < 0.01$; $R^2 = 0.26$, **Fig 3F**).

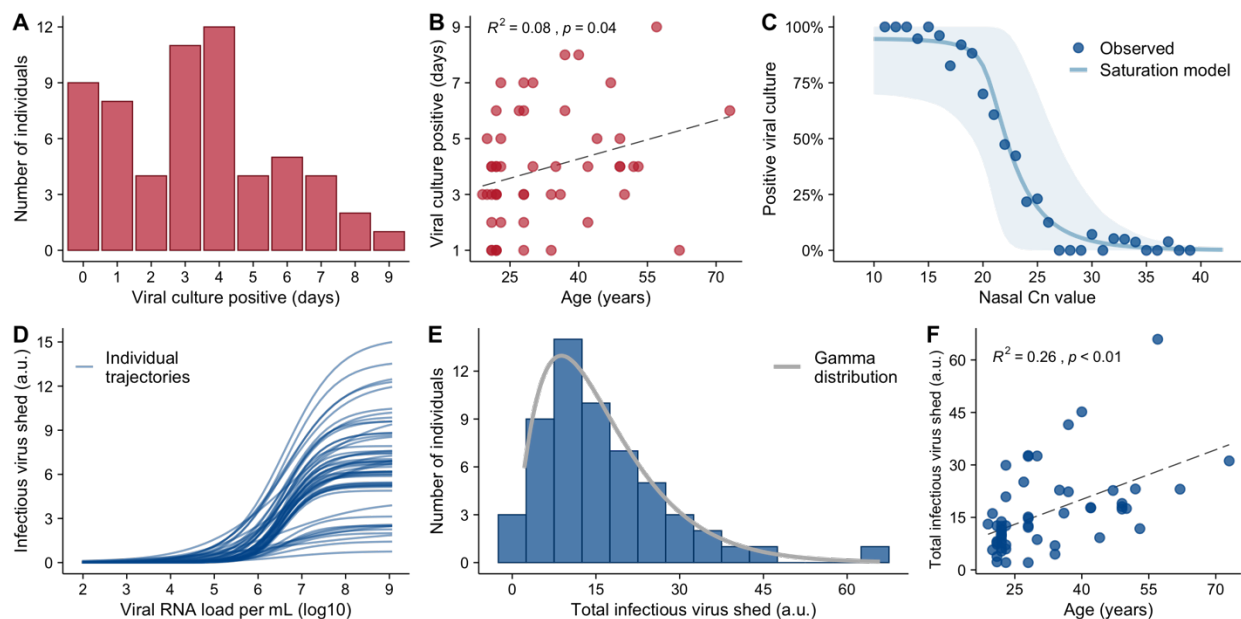


Figure 3: Substantial heterogeneity in infectious virus shedding between individuals. (A) The distribution in the numbers of days in which participants tested positive by viral culture on nasal swabs following study enrollment. (B) Association between age of study participants and the number of days of positive viral culture. (C) Plot shows the relationship between the Cn value in nasal samples and the probability of the sample being viral culture positive summarized across all individuals. Individual samples were binned based on their Cn values. Dots indicate the observed percentage of positive samples within a bin that were viral culture positive. The solid line and the shaded area are the mean and the 90% confidence intervals, respectively, of the trajectories generated using the best-fit parameters of the saturation model (see Fig S7 for

individual fits). **(D)** The relationship between infectious virus shed (in arbitrary units) and viral genome load predicted by the saturation model for 56 individuals in our analysis. **(E)** Distribution of the estimated total cumulative level of infectious virus (in arbitrary units) shed from the nasal passage by each participant over the course of infection. Solid line shows the best-fit gamma distribution with a shape parameter of 2. **(F)** Association between age and the estimated total infectious virus (in arbitrary units) shed by each individual. The R^2 value and the p value from a linear regression (dashed line) are shown in panels (B) and (F).

Analysis of B.1.1.7 (Alpha) viral dynamics

Finally, we asked whether infection with the B.1.1.7 (Alpha) variant of concern (VOC) was associated with any significant differences in viral dynamics that could potentially explain the enhanced transmissibility of this genotype (32–34). Previous studies have suggested that B.1.1.7 infection may result in higher peak viral loads or prolonged shedding compared with other circulating genotypes (35–37). Within our cohort, 15 out of 60 individuals were infected with B.1.1.7.

Both the empirical data and our model analysis suggest that viral genome shedding dynamics in nasal samples are indistinguishable between B.1.1.7 and non-B.1.1.7 infections (none of the latter were VOC genotypes with the exception of a single P.1 (Gamma) infection; see **Table S9**), and that there is no significant difference between B.1.1.7 and non-B.1.1.7 viruses in total infectiousness predicted by our model (**Fig 4A-B**). Comparison of genome shedding in saliva between B.1.1.7 and non-B.1.1.7 infected individuals revealed a slower rise in pre-peak viral loads for B.1.1.7 versus non-B.1.1.7 (**Fig 4C**). This is corroborated by the mathematical modeling results, where viral growth rates were estimated to be significantly slower for B.1.1.7 compared with non-B.1.1.7 viruses (p -value < 0.01) (**Fig 4D**). No clear differences in peak viral genome loads or overall clearance kinetics in saliva were apparent, although the peak viral genome loads for B.1.1.7 seemed to be slightly, but not significantly higher ($p=0.2$). With a slower rise to peak viral load for B.1.1.7 and similar peak viral loads between B.1.1.7 and non-B.1.1.7 viruses, our model predicts a longer pre-peak shedding period for B.1.1.7. These data implicate either the delayed rise in pre-peak shedding in saliva or other mechanisms not reflected in viral dynamics in the increased transmissibility of this variant.

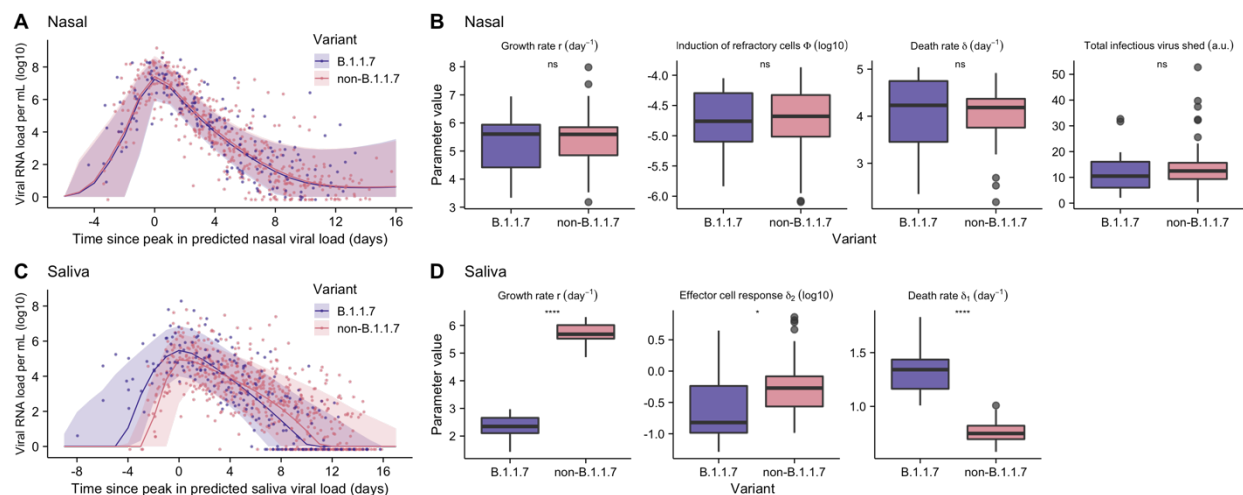


Figure 4: Comparison of viral dynamics between B.1.1.7 and non-B.1.1.7 viruses. (A, C) Viral genome load of B.1.1.7 infections (purple) and non-B.1.1.7 infections (pink) over time in (A) nasal and (C) saliva samples, as measured by RTqPCR. Dots indicate individual viral

genome load measurements. Ribbons indicate 90% confidence intervals of viral genome load trajectories generated using population parameters estimated from modeling analysis (Table S4). (B, D) Estimated values for the indicated model parameters from individual B.1.1.7 and non-B.1.1.7 infections (see Table S5 for parameter values). Total infectious virus shed after adjusting for age is shown in panel (B) (see Methods).

Discussion:

This study describes the results of daily multi-compartment sampling of viral dynamics within dozens of individuals newly infected with SARS-CoV-2, and provides the most comprehensive, high-resolution description of SARS-CoV-2 (or any acute RNA virus) shedding and clearance dynamics in humans to date.

Superspreading, in which a small subset of infected individuals are responsible for a disproportionately large share of transmission events, has been identified as a major driver of community spread of SARS-CoV-2, SARS-CoV, and many other acute viral pathogens (6, 7, 9). Superspreading is believed to arise from heterogeneity in both (1) contact structure between individuals arising from behavioral and environmental factors, and (2) the intrinsic infectiousness of individuals (9, 38). While heterogeneity in contact structure has been studied extensively (39–42), the extent of heterogeneity in infectiousness arising from individual-level infection processes has not been empirically measured during natural human infection. As a result, the extent to which individual-level heterogeneity in viral dynamics contributes to superspreading remains unknown.

To address this question, we empirically quantified infectious virus shedding through daily longitudinal sampling of individuals infected with SARS-CoV-2. The substantial heterogeneity in infectious virus shedding that we observed among individuals indicates that superspreading is likely driven by individual-level variation in specific features of the infection process, in addition to behavioral and environmental factors. We also found that heterogeneity in infectious virus shedding was only partly explained by individual-level heterogeneity in viral genome load dynamics, suggesting that additional factors such as variation in the timing and magnitude of the neutralizing antibody response might contribute (43). Our results here suggest caution in assessing the infectiousness of an individual using viral genome load data alone. Further, the absence of clear viral genetic correlates of infectiousness within this dataset suggests the existence of specific host determinants of superspreading potential. While we identified age as a significant correlate of infectiousness, additional determinants likely exist. Identification of these correlates could aid future efforts to mitigate community spread of the virus.

Our finding that viral shedding often peaks earlier in saliva versus the nasal compartment, sometimes by several days, corroborates a recent study of four individuals (44) and has several important implications. First, saliva screening may be a more effective sample type than nasal swabs for detecting infected individuals prior to or early in the infectious period (45). Early detection and isolation of infected individuals is absolutely critical for breaking transmission chains (15). Moreover, early viral shedding from the oral cavity may contribute to the high prevalence of pre-symptomatic SARS-CoV-2 transmission. We were unable to directly assess viral infectivity in saliva so it remains unclear whether the earlier peaks in viral RNA shedding that we observed in saliva reflect earlier shedding of transmission-competent virus. The earlier detection of virus in saliva also raises questions about the initial site of SARS-CoV-2 infection. A recent study demonstrated that both salivary glands and oral mucosal epithelium can support SARS-CoV-2 replication, suggesting that infection could be initiated within the oral cavity (46). Alternatively, if infection is initiated in the nasopharynx or soft palette, viral RNA might be

detectable in saliva prior to detection in the mid-turbinate swabs used in this study. The discordance in shedding dynamics between oral and nasal samples that we observed in many participants is consistent with a significant degree of compartmentalization between these adjacent but distinct tissue sites, as has been observed in animal models of influenza virus infection (47).

The specific mechanisms driving the enhanced transmissibility of the B.1.1.7 variant remain poorly understood. Recent studies have identified alterations in the structural conformation of the spike protein and enhanced antagonism of innate immunity by B.1.1.7 as potential contributors (48, 49). Contrary to previous clinical studies, we observed no significant differences in either peak viral loads or clearance kinetics between B.1.1.7 and non-B.1.1.7 viruses as measured in either nasal swabs or saliva. Previous reports of differences in Ct values or duration of shedding were based on either small sample sizes (36) or cross-sectional analyses of individuals who sought testing and were thus likely symptomatic at the time of sampling (21). Our results are consistent with studies demonstrating the absence of a growth advantage for B.1.1.7 in primary human respiratory epithelial cells (50, 51). If the timing of symptom onset differs between B.1.1.7 and non-B.1.1.7 infections, it could potentially explain why cross-sectional analyses of viral loads might register lower Ct values for B.1.1.7 samples. Interestingly, we did observe a significantly slower pre-peak growth rate of B.1.1.7 than non-B.1.1.7 viruses in saliva but not nasal swabs, resulting in more prolonged shedding in saliva prior to peak viral load. Because symptom onset seems to occur around the time of peak viral load (1, 25, 52) the prolonged pre-peak shedding in saliva by B.1.1.7 could potentially enhance transmissibility by extending the pre-symptomatic shedding period of the virus, however more data are needed to determine whether there is a mechanistic link between pre-peak shedding kinetics and transmission potential. Alternatively, the enhanced transmissibility of the B.1.1.7 variant may also be driven by features not reflected in shedding dynamics, e.g. enhanced environmental stability or a lower infectious dose threshold.

Altogether, our data provide an unprecedented view of the longitudinal viral dynamics of SARS-CoV-2 infection in humans and implicate both individual-level heterogeneity in viral shedding and the oral compartment and as playing critical roles in transmission.

Methods:

This study was approved by the Western Institutional Review Board, and all participants provided informed consent.

Participants

All on-campus students and employees of the University of Illinois at Urbana-Champaign are required to submit saliva for RTqPCR testing every 2-4 days as part of the SHIELD campus surveillance testing program. Individuals testing positive were instructed to isolate and were eligible to enroll in this study for a period of 24 hours following receipt of their positive test result. Close contacts of individuals who test positive (particularly those co-housed with them) are instructed to quarantine and were eligible to enroll for up to 5 days after their last known exposure to an infected individual. All participants were also required to have received a negative saliva RTqPCR result 7 days prior to enrollment.

Individuals were recruited via either a link shared in an automated text message providing isolation information sent within 30 minutes of a positive test result, a call from a study recruiter, or a link shared by an enrolled study participant or included in information provided to all

quarantining close contacts. In addition, signs were used at each testing location and a website was available to inform the community about the study.

Participants were required to be at least 18 years of age, have a valid university ID, speak English, have internet access, and live within 8 miles of the university campus. After enrollment and consent, participants completed an initial survey to collect information on demographics and health history and were provided with sample collection supplies. Participants who tested positive prior to enrollment or during quarantine were followed for up to 14 days. Quarantining participants who continued to test negative by saliva RTqPCR were followed for up to 7 days after their last exposure. All participants' data and survey responses were collected in the Eureka digital study platform.

Sample collection

Each day, participants were remotely observed by trained study staff collecting:

1. 2 mL of saliva into a 50mL conical tube.
2. 1 nasal swab from a single nostril using a foam-tipped swab that was placed within a dry collection tube.
3. 1 nasal swab from the other nostril using a flocked swab that was subsequently placed in a collection vial containing viral transport media (VTM).

The order of nostrils (left vs. right) used for the two different swabs was randomized. For nasal swabs, participants were instructed to insert the soft tip of the swab at least 1 cm into the indicated nostril until they encountered mild resistance, rotate the swab around the nostril 5 times, leaving it in place for 10-15 seconds. After daily sample collection, participants completed a symptom survey. A courier collected all participant samples within 1 hour of collection using a no-contact pickup protocol designed to minimize courier exposure to infected participants.

Saliva RTqPCR

After collection, saliva samples were stored at room temperature and RTqPCR was run within 12 hours of initial collection. The protocol for the covidSHIELD direct saliva-to-RTqPCR assay used has been detailed previously (22). In brief, saliva samples were heated at 95°C for 30 minutes, followed by the addition of 2X Tris/Borate/EDTA buffer (TBE) at a 1:1 ratio (final concentration 1X TBE) and Tween-20 to a final concentration of 0.5%. Samples were assayed using the Thermo Taqpath COVID-19 assay.

Antigen testing

Foam-tipped nasal swabs were placed in collection tubes, transported with cold packs, and stored at 4°C overnight based on guidance from the manufacturer. The morning after collection, swabs were run through the Sofia SARS antigen FIA on Sofia 2 devices according to the manufacturer's protocol.

Nasal swab RTqPCR

Collection tubes containing VTM and flocked nasal swabs were stored at -80°C after collection and were subsequently shipped to Johns Hopkins University for RTqPCR and virus culture testing. After thawing, VTM was aliquoted for RTqPCR and infectivity assays. One ml of VTM from the nasal swab was assayed on the Abbott Alinity per manufacturer's instructions in a College of American Pathologist and CLIA-certified laboratory. Calibration curve for Alinity assay was determined using digital droplet PCR (ddPCR) as previously described (53).

Virus culture from nasal swabs

Vero-TMPRSS2 cells were grown in complete medium (CM) consisting of DMEM with 10% fetal bovine serum (Gibco), 1 mM glutamine (Invitrogen), 1 mM sodium pyruvate (Invitrogen), 100 U/ml of penicillin (Invitrogen), and 100 µg/ml of streptomycin (Invitrogen) (54). Viral infectivity was assessed on Vero-TMPRSS2 cells as previously described using infection media (IM; identical to CM except the FBS is reduced to 2.5%) (24). When a cytopathic effect was visible in >50% of cells in a given well, the supernatant was harvested. The presence of SARS-CoV-2 was confirmed through RTqPCR as described previously by extracting RNA from the cell culture supernatant using the Qiagen viral RNA isolation kit and performing RTqPCR using the N1 and N2 SARS-CoV-2-specific primers and probes in addition to primers and probes for human RNaseP gene using synthetic RNA target sequences to establish a standard curve (55).

Viral genome sequencing and analysis

Viral RNA was extracted from 140 µL of heat inactivated (30 minutes at 95°C, as part of protocol detailed in (22)) saliva samples using the QIAamp viral RNA mini kit (QIAGEN). 100ng of viral RNA was used to generate cDNA using the SuperScript IV first strand synthesis kit (Invitrogen). Viral cDNA was then used to generate sequencing libraries using the Swift SNAP Amplicon SARS CoV2 kit with additional coverage panel and unique dual indexing (Swift Biosciences) which were sequenced on an Illumina Novaseq SP lane. Data were run through the nf-core/viralrecon workflow (<https://nf-co.re/viralrecon/1.1.0>), using the Wuhan-Hu-1 reference genome (NCBI accession NC_045512.2). Swift v2 primer sequences were trimmed prior to variant analysis from iVar version 1.3.1 (<https://genomebiology.biomedcentral.com/articles/10.1186/s13059-018-1618-7>) retaining all calls with a minimum allele frequency of 0.01 and higher. Viral lineages were called using the Pangolin tool (<https://github.com/cov-lineages/pangolin>) version 2.4.2, pango version 1.2.6, and the 5/19/21 version of the pangoLEARN model; based on the nomenclature system described in (56).

Mathematical modeling and statistical analyses

See **Fig S6** for an overview of the modeling and analysis workflow of various datasets. See supplementary material for descriptions of the viral dynamic models, models describing the relationship between the infectious viral load and the viral genome load, model fitting and model comparison procedures.

Associations between the age of individuals and each parameter of interest (on a linear or a log scale) are examined using a linear regression analysis. The R-squared value and the p-value are reported. The difference in the distribution of a parameter of interest between the non-B.1.1.7 infection group and the B.1.1.7 infection group is assessed using a univariate analysis and the p-value is calculated using the Mann-Whitney test. The comparison of infectious virus shedding between the non-B.1.1.7 infection group and the B.1.1.7 infection group is performed using a multivariate analysis with age as an additional variate. The predicted levels of infectious viral shedding after adjusting for age by assuming age is at 28, i.e. the median age of the cohort, are shown in Fig. 4C.

Acknowledgments. We wish to thank Shumon Ahmed, Carly Bell, Nate Bouton, Callie Brennen, Justin Brown, Coleco Buie, Emmaline Cler, Gary Cole, Trey Coleman, Alastair Dunnett, Lauren Engels, Savannah Feher, Kelsey Fox, Lexi Freeman, Yesenia Gonzalez, Montez Harris, Darcy Henness, Dan Hiser, Ayesha Hussain, Daryl Jackson, Junko Jarrett, Michael Jenkins, Kalombo Kalonji, Syntyche Kanku, Steven Krauklis, Mary Krouse, Elmore Leshoure, Joe Lewis, Maggie Li, Angel Lopez, Guadalupe Lopez, Emily Luna, Chun Huai Luo, Colby Mackey, Skyler McLain, Yared Berhanu Melesse, Madison O'Donnell, Savanna Pflugmacher, Denver Piatt, Skyler Pierce, Gina Quitanilla, Ameera Samad, MacKenzie

Scroggins, Monique Settles, Macie Sinn, Pete Varney, Evette Vlach, Raeshun Williams-Chatman, and Todd Young for their efforts supporting recruitment, enrollment, logistics, and/or sample collection and processing. We also thank Jeffrey Olgin, Noah Peyser, and Xochitl Butcher for assistance with the Eureka platform, Michelle Lore for assistance with REDcap, Melanie Loots for assistance with administration, Gillian Snyder for assistance in development of study protocols and logistics, and Erin Iturriaga and Jue Chen for study protocol development. Finally, we are grateful to Dr. Alvaro Hernandez and Ms. Chris Wright of the DNA Services Lab within the Roy J. Carver Biotechnology Center for expert assistance in establishment of a SARS-CoV-2 genomic sequencing protocol. Vero-TMPRSS2 cells were kindly provided by the National Institute of Infectious Diseases, Japan. Sofia 2 devices and associated supplies were provided to Carle Foundation Hospital by Quidel, however Quidel played no role in the design of the study or the interpretation or presentation of the data.

Funding. This work was supported by the National Heart, Lung, and Blood Institute at the National Institutes of Health [3U54HL143541-02S2] through the RADx-Tech program. The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Institute of Biomedical Imaging and Bioengineering; the National Heart, Lung, and Blood Institute; the National Institutes of Health, or the U.S. Department of Health and Human Services.

Author contributions.

Conceptualization: RK, RLS, WJH, YCM, AP, LLG, CBB

Data curation: RLS, BB, PL

Formal Analysis: RK, PPM, RLS

Funding acquisition: DDM, LLG

Investigation: RK, PPM, RLS, AM, MC, NG, CHL, JJ, AC, TL, MF, KW, CJF, LW, RF, MEB, KKC, HC, KRS, ANO, JB, MLR

Methodology: RK, RLS, CBB

Project administration: DCE, KRS, SB, SLG, CR, JY, JQ

Software: RK, PPM, RLS

Supervision: HHM, YCM, AP, LLG, CBB

Visualization: RK, PPM, RLS, CBB

Writing – original draft: RK, CBB

Writing – review & editing: RK, PPM, RLS, YCM, AP, LLG, CBB

Competing interests. CBB and LW are listed as inventors on a pending patent application for the saliva RTqPCR test used in this study.

Data and materials availability: Due to institutional (LANL) policies, the codes for the viral dynamic models cannot be made publicly available. The simulation results of the models can be fully reproduced from the equations in the Supporting Text and the parameter values listed in Supplementary Tables. Raw data will be made available at the time of publication.

REFERENCES

1. X. He, E. H. Y. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan, X. Mo, Y. Chen, B. Liao, W. Chen, F. Hu, Q. Zhang, M. Zhong, Y. Wu, L. Zhao, F. Zhang, B. J. Cowling, F. Li, G. M. Leung, Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med.* **26**, 672–675 (2020).

2. L. Ferretti, A. Ledda, C. Wymant, L. Zhao, V. Ledda, L. Abeler-Dörner, M. Kendall, A. Nurtay, H.-Y. Cheng, T.-C. Ng, H.-H. Lin, R. Hinch, J. Masel, A. M. Kilpatrick, C. Fraser, “The timing of COVID-19 transmission” (preprint, *Epidemiology*, 2020), , doi:10.1101/2020.09.04.20188516.
3. C. M. Szablewski, K. T. Chang, M. M. Brown, V. T. Chu, A. R. Yousaf, N. Anyalechi, P. A. Aryee, H. L. Kirking, M. Lumsden, E. Mayweather, C. J. McDaniel, R. Montierth, A. Mohammed, N. G. Schwartz, J. A. Shah, J. E. Tate, E. Dirlikov, C. Drenzek, T. M. Lanzieri, R. J. Stewart, SARS-CoV-2 Transmission and Infection Among Attendees of an Overnight Camp — Georgia, June 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 1023–1025 (2020).
4. Q.-X. Long, X.-J. Tang, Q.-L. Shi, Q. Li, H.-J. Deng, J. Yuan, J.-L. Hu, W. Xu, Y. Zhang, F.-J. Lv, K. Su, F. Zhang, J. Gong, B. Wu, X.-M. Liu, J.-J. Li, J.-F. Qiu, J. Chen, A.-L. Huang, Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat Med.* **26**, 1200–1204 (2020).
5. R. Li, S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, J. Shaman, Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science.* **368**, 489–493 (2020).
6. K. Sun, W. Wang, L. Gao, Y. Wang, K. Luo, L. Ren, Z. Zhan, X. Chen, S. Zhao, Y. Huang, Q. Sun, Z. Liu, M. Litvinova, A. Vespignani, M. Ajelli, C. Viboud, H. Yu, Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. *Science.* **371**, eabe2424 (2021).
7. D. C. Adam, P. Wu, J. Y. Wong, E. H. Y. Lau, T. K. Tsang, S. Cauchemez, G. M. Leung, B. J. Cowling, Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. *Nat Med.* **26**, 1714–1719 (2020).
8. Q. Bi, Y. Wu, S. Mei, C. Ye, X. Zou, Z. Zhang, X. Liu, L. Wei, S. A. Truelove, T. Zhang, W. Gao, C. Cheng, X. Tang, X. Wu, Y. Wu, B. Sun, S. Huang, Y. Sun, J. Zhang, T. Ma, J. Lessler, T. Feng, Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *The Lancet Infectious Diseases.* **20**, 911–919 (2020).
9. J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, W. M. Getz, Superspreading and the effect of individual variation on disease emergence. *Nature.* **438**, 355–359 (2005).
10. N. Néant, G. Lingas, Q. Le Hingrat, J. Ghosn, I. Engelmann, Q. Lepiller, A. Gaymard, V. Ferré, C. Hartard, J.-C. Plantier, V. Thibault, J. Marlet, B. Montes, K. Bouiller, F.-X. Lescure, J.-F. Timsit, E. Faure, J. Poissy, C. Chidiac, F. Raffi, A. Kimmoun, M. Etienne, J.-C. Richard, P. Tattevin, D. Garot, V. Le Moing, D. Bachelet, C. Tardivon, X. Duval, Y. Yazdanpanah, F. Mentré, C. Laouénan, B. Visseaux, J. Guedj, for the French COVID Cohort Investigators and French Cohort Study groups, Modeling SARS-CoV-2 viral kinetics and association with mortality in hospitalized patients from the French COVID cohort. *Proc Natl Acad Sci USA.* **118**, e2017962118 (2021).
11. J. Silva, C. Lucas, M. Sundaram, B. Israelow, P. Wong, J. Klein, M. Tokuyama, P. Lu, A. Venkataraman, F. Liu, T. Mao, J. E. Oh, A. Park, A. Casanovas-Massana, C. B. F. Vogels, M. C. Muenker, J. Zell, J. B. Fournier, M. Campbell, M. Chiorazzi, E. R. Fuentes, M. E. Petrone, C. C. Kalinich, I. M. Ott, A. Watkins, A. J. Moore, M. Nakahata, Yale IMPACT

- Team, N. D. Grubaugh, S. Farhadian, C. D. Cruz, A. I. Ko, W. L. Schulz, A. Ring, S. Ma, S. Omer, A. L. Wyllie, A. Iwasaki, “Saliva viral load is a dynamic unifying correlate of COVID-19 severity and mortality” (preprint, *Infectious Diseases (except HIV/AIDS)*, 2021), , doi:10.1101/2021.01.04.21249236.
12. The Massachusetts Consortium for Pathogen Readiness, J. Fajnzylber, J. Regan, K. Coxen, H. Corry, C. Wong, A. Rosenthal, D. Worrall, F. Giguel, A. Piechocka-Trocha, C. Atyeo, S. Fischinger, A. Chan, K. T. Flaherty, K. Hall, M. Dougan, E. T. Ryan, E. Gillespie, R. Chishti, Y. Li, N. Jilg, D. Hanidziar, R. M. Baron, L. Baden, A. M. Tsibris, K. A. Armstrong, D. R. Kuritzkes, G. Alter, B. D. Walker, X. Yu, J. Z. Li, SARS-CoV-2 viral load is associated with increased disease severity and mortality. *Nat Commun.* **11**, 5493 (2020).
 13. S. Zheng, J. Fan, F. Yu, B. Feng, B. Lou, Q. Zou, G. Xie, S. Lin, R. Wang, X. Yang, W. Chen, Q. Wang, D. Zhang, Y. Liu, R. Gong, Z. Ma, S. Lu, Y. Xiao, Y. Gu, J. Zhang, H. Yao, K. Xu, X. Lu, G. Wei, J. Zhou, Q. Fang, H. Cai, Y. Qiu, J. Sheng, Y. Chen, T. Liang, Viral load dynamics and disease severity in patients infected with SARS-CoV-2 in Zhejiang province, China, January-March 2020: retrospective cohort study. *BMJ*, m1443 (2020).
 14. M. Marks, P. Millat-Martinez, D. Ouchi, C. h Roberts, A. Alemany, M. Corbacho-Monné, M. Ubals, A. Tobias, C. Tebé, E. Ballana, Q. Bassat, B. Baro, M. Vall-Mayans, C. G-Beiras, N. Prat, J. Ara, B. Clotet, O. Mitjà, Transmission of COVID-19 in 282 clusters in Catalonia, Spain: a cohort study. *The Lancet Infectious Diseases*, S1473309920309853 (2021).
 15. D. B. Larremore, B. Wilder, E. Lester, S. Shehata, J. M. Burke, J. A. Hay, M. Tambe, M. J. Mina, R. Parker, Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening. *Sci. Adv.* **7**, eabd5393 (2021).
 16. S. M. Kissler, J. R. Fauver, C. Mack, S. W. Olesen, C. Tai, K. Y. Shiue, C. C. Kalinich, S. Jednak, I. M. Ott, C. B. F. Vogels, J. Wohlgemuth, J. Weisberger, J. DiFiori, D. J. Anderson, J. Mancell, D. D. Ho, N. D. Grubaugh, Y. H. Grad, “SARS-CoV-2 viral dynamics in acute infections” (preprint, *Epidemiology*, 2020), , doi:10.1101/2020.10.21.20217042.
 17. J. Y. Kim, J.-H. Ko, Y. Kim, Y.-J. Kim, J.-M. Kim, Y.-S. Chung, H. M. Kim, M.-G. Han, S. Y. Kim, B. S. Chin, Viral Load Kinetics of SARS-CoV-2 Infection in First Two Patients in Korea. *J Korean Med Sci.* **35**, e86 (2020).
 18. R. Wölfel, V. M. Corman, W. Guggemos, M. Seilmaier, S. Zange, M. A. Müller, D. Niemeyer, T. C. Jones, P. Vollmar, C. Rothe, M. Hoelscher, T. Bleicker, S. Brünink, J. Schneider, R. Ehmann, K. Zwirgmaier, C. Drosten, C. Wendtner, Virological assessment of hospitalized patients with COVID-2019. *Nature.* **581**, 465–469 (2020).
 19. F.-X. Lescure, L. Bouadma, D. Nguyen, M. Parisey, P.-H. Wicky, S. Behillil, A. Gaymard, M. Bouscambert-Duchamp, F. Donati, Q. Le Hingrat, V. Enouf, N. Houhou-Fidouh, M. Valette, A. Mailles, J.-C. Lucet, F. Mentre, X. Duval, D. Descamps, D. Malvy, J.-F. Timsit, B. Lina, S. van-der-Werf, Y. Yazdanpanah, Clinical and virological data of the first cases of COVID-19 in Europe: a case series. *The Lancet Infectious Diseases.* **20**, 697–706 (2020).
 20. B. E. Young, S. W. X. Ong, S. Kalimuddin, J. G. Low, S. Y. Tan, J. Loh, O.-T. Ng, K. Marimuthu, L. W. Ang, T. M. Mak, S. K. Lau, D. E. Anderson, K. S. Chan, T. Y. Tan, T. Y. Ng, L. Cui, Z. Said, L. Kurupatham, M. I.-C. Chen, M. Chan, S. Vasoo, L.-F. Wang, B. H. Tan, R. T. P. Lin, V. J. M. Lee, Y.-S. Leo, D. C. Lye, for the Singapore 2019 Novel

Coronavirus Outbreak Research Team, Epidemiologic Features and Clinical Course of Patients Infected With SARS-CoV-2 in Singapore. *JAMA*. **323**, 1488 (2020).

21. T. C. Jones, G. Biele, B. Mühlemann, T. Veith, J. Schneider, J. Beheim-Schwarzbach, T. Bleicker, J. Tesch, M. L. Schmidt, L. E. Sander, F. Kurth, P. Menzel, R. Schwarzer, M. Zuchowski, J. Hofmann, A. Krumbholz, A. Stein, A. Edelmann, V. M. Corman, C. Drosten, Estimating infectiousness throughout SARS-CoV-2 infection course. *Science*, eabi5273 (2021).
22. D. R. E. Ranoa, R. L. Holland, F. G. Alnaji, K. J. Green, L. Wang, C. B. Brooke, M. D. Burke, T. M. Fan, P. J. Hergenrother, “Saliva-Based Molecular Testing for SARS-CoV-2 that Bypasses RNA Extraction” (preprint, *Microbiology*, 2020), , doi:10.1101/2020.06.18.159434.
23. B. Monel, D. Planas, L. Grzelak, N. Smith, N. Robillard, I. Staropoli, P. Goncalves, F. Porrot, F. Guivel-Benhassine, N. D. Guinet, J. Rodary, J. Puech, V. Euzen, L. Bélec, G. Orvoen, L. Nunes, V. Moulin, J. Fourgeaud, M. Wack, S. Imbeaud, P. Campagne, D. Duffy, J. P. Di Santo, T. Bruel, H. Péré, D. Veyer, O. Schwartz, “Release of infectious virus and cytokines in nasopharyngeal swabs from individuals infected with non-B.1.1.7 or B.1.1.7 SARS-CoV-2 variants” (preprint, *Infectious Diseases (except HIV/AIDS)*, 2021), , doi:10.1101/2021.05.20.21257393.
24. A. Pekosz, V. Parvu, M. Li, J. C. Andrews, Y. C. Manabe, S. Kodsi, D. S. Gary, C. Roger-Dalbert, J. Leitch, C. K. Cooper, Antigen-Based Testing but Not Real-Time Polymerase Chain Reaction Correlates With Severe Acute Respiratory Syndrome Coronavirus 2 Viral Culture. *Clin Infect Dis* (2021), doi:10.1093/cid/ciaa1706.
25. R. Ke, C. Zitzmann, R. M. Ribeiro, A. S. Perelson, “Kinetics of SARS-CoV-2 infection in the human upper and lower respiratory tracts and their relationship with infectiousness” (preprint, *Infectious Diseases (except HIV/AIDS)*, 2020), , doi:10.1101/2020.09.25.20201772.
26. K. A. Pawelek, G. T. Huynh, M. Quinlivan, A. Cullinane, L. Rong, A. S. Perelson, Modeling Within-Host Dynamics of Influenza Virus Infection Including Immune Responses. *PLoS Comput Biol*. **8**, e1002588 (2012).
27. D. Blanco-Melo, B. E. Nilsson-Payant, W.-C. Liu, S. Uhl, D. Hoagland, R. Møller, T. X. Jordan, K. Oishi, M. Panis, D. Sachs, T. T. Wang, R. E. Schwartz, J. K. Lim, R. A. Albrecht, B. R. tenOever, Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell*. **181**, 1036-1045.e9 (2020).
28. A. C. Shaw, D. R. Goldstein, R. R. Montgomery, Age-dependent dysregulation of innate immunity. *Nat Rev Immunol*. **13**, 875–887 (2013).
29. R. D. Molony, J. T. Nguyen, Y. Kong, R. R. Montgomery, A. C. Shaw, A. Iwasaki, Aging impairs both primary and secondary RIG-I signaling for interferon induction in human monocytes. *Sci. Signal*. **10**, eaan2392 (2017).
30. R. Angioni, R. Sánchez-Rodríguez, F. Munari, N. Bertoldi, D. Arcidiacono, S. Cavinato, D. Marturano, A. Zaramella, S. Realdon, A. Cattelan, A. Viola, B. Molon, Age-severity

- matched cytokine profiling reveals specific signatures in Covid-19 patients. *Cell Death Dis.* **11**, 957 (2020).
31. R. Ke, C. Zitzmann, D. D. Ho, R. M. Ribeiro, A. S. Perelson, “In vivo kinetics of SARS-CoV-2 infection and its relationship with a person’s infectiousness” (preprint, *Infectious Diseases (except HIV/AIDS)*, 2021), , doi:10.1101/2021.06.26.21259581.
 32. N. L. Washington, K. Gangavarapu, M. Zeller, A. Bolze, E. T. Cirulli, K. M. Schiabor Barrett, B. B. Larsen, C. Anderson, S. White, T. Cassens, S. Jacobs, G. Levan, J. Nguyen, J. M. Ramirez, C. Rivera-Garcia, E. Sandoval, X. Wang, D. Wong, E. Spencer, R. Robles-Sikisaka, E. Kurzban, L. D. Hughes, X. Deng, C. Wang, V. Servellita, H. Valentine, P. De Hoff, P. Seaver, S. Sathe, K. Gietzen, B. Sickler, J. Antico, K. Hoon, J. Liu, A. Harding, O. Bakhtar, T. Basler, B. Austin, D. MacCannell, M. Isaksson, P. G. Febbo, D. Becker, M. Laurent, E. McDonald, G. W. Yeo, R. Knight, L. C. Laurent, E. de Feo, M. Worobey, C. Y. Chiu, M. A. Suchard, J. T. Lu, W. Lee, K. G. Andersen, Emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. *Cell*, S0092867421003834 (2021).
 33. N. G. Davies, S. Abbott, R. C. Barnard, C. I. Jarvis, A. J. Kucharski, J. D. Munday, C. A. B. Pearson, T. W. Russell, D. C. Tully, A. D. Washburne, T. Wenseleers, A. Gimma, W. Waites, K. L. M. Wong, K. van Zandvoort, J. D. Silverman, CMMID COVID-19 Working Group¹‡, COVID-19 Genomics UK (COG-UK) Consortium[‡], K. Diaz-Ordaz, R. Keogh, R. M. Eggo, S. Funk, M. Jit, K. E. Atkins, W. J. Edmunds, Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science*. **372**, eabg3055 (2021).
 34. The COVID-19 Genomics UK (COG-UK) consortium, E. Volz, S. Mishra, M. Chand, J. C. Barrett, R. Johnson, L. Geidelberg, W. R. Hinsley, D. J. Laydon, G. Dabrera, Á. O’Toole, R. Amato, M. Ragonnet-Cronin, I. Harrison, B. Jackson, C. V. Ariani, O. Boyd, N. J. Loman, J. T. McCrone, S. Gonçalves, D. Jorgensen, R. Myers, V. Hill, D. K. Jackson, K. Gaythorpe, N. Groves, J. Sillitoe, D. P. Kwiatkowski, S. Flaxman, O. Ratmann, S. Bhatt, S. Hopkins, A. Gandy, A. Rambaut, N. M. Ferguson, Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature* (2021), doi:10.1038/s41586-021-03470-x.
 35. P. Calistri, L. Amato, I. Puglia, F. Cito, A. Di Giuseppe, M. L. Danzetta, D. Morelli, M. Di Domenico, M. Caporale, S. Scialabba, O. Portanti, V. Curini, F. Perletta, C. Cammà, M. Ancora, G. Savini, G. Migliorati, N. D’Alterio, A. Lorusso, Infection sustained by lineage B.1.1.7 of SARS-CoV-2 is characterised by longer persistence and higher viral RNA loads in nasopharyngeal swabs. *International Journal of Infectious Diseases*. **105**, 753–755 (2021).
 36. S. M. Kissler, J. R. Fauver, C. Mack, C. G. Tai, M. I. Breban, A. E. Watkins, R. M. Samant, D. J. Anderson, D. D. Ho, N. D. Grubaugh, Y. H. Grad, “Densely sampled viral trajectories suggest longer duration of acute infection with B.1.1.7 variant relative to non-B.1.1.7 SARS-CoV-2” (preprint, *Epidemiology*, 2021), , doi:10.1101/2021.02.16.21251535.
 37. C. L. Althaus, S. Baggio, M. L. Reichmuth, E. B. Hodcroft, J. Riou, R. A. Neher, F. Jacquerioz, H. Spechbach, J. Salamun, P. Vetter, C. Williamson, N. Hsiao, W. Preisler, M.-A. Davies, R. J. Lessells, T. de Olivera, L. Kaiser, I. Eckerle, “A tale of two variants: Spread of SARS-CoV-2 variants Alpha in Geneva, Switzerland, and Beta in South Africa” (preprint, *Epidemiology*, 2021), , doi:10.1101/2021.06.10.21258468.

38. S. S. Lakdawala, V. D. Menachery, Catch me if you can: superspreading of COVID-19. *Trends Microbiol*, S0966-842X(21)00124–4 (2021).
39. J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, W. J. Edmunds, Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. *PLoS Med*. **5**, e74 (2008).
40. J. Zhang, M. Litvinova, Y. Liang, Y. Wang, W. Wang, S. Zhao, Q. Wu, S. Merler, C. Viboud, A. Vespignani, M. Ajelli, H. Yu, Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science*. **368**, 1481–1486 (2020).
41. D. Mistry, M. Litvinova, A. Pastore y Piontti, M. Chinazzi, L. Fumanelli, M. F. C. Gomes, S. A. Haque, Q.-H. Liu, K. Mu, X. Xiong, M. E. Halloran, I. M. Longini, S. Merler, M. Ajelli, A. Vespignani, Inferring high-resolution human mixing patterns for disease modeling. *Nat Commun*. **12**, 323 (2021).
42. J. Wallinga, P. Teunis, M. Kretzschmar, Using Data on Social Contacts to Estimate Age-specific Transmission Parameters for Respiratory-spread Infectious Agents. *American Journal of Epidemiology*. **164**, 936–944 (2006).
43. J. J. A. van Kampen, D. A. M. C. van de Vijver, P. L. A. Fraaij, B. L. Haagmans, M. M. Lamers, N. Okba, J. P. C. van den Akker, H. Endeman, D. A. M. P. J. Gommers, J. J. Cornelissen, R. A. S. Hoek, M. M. van der Eerden, D. A. Hesselink, H. J. Metselaar, A. Verbon, J. E. M. de Steenwinkel, G. I. Aron, E. C. M. van Gorp, S. van Boheemen, J. C. Voermans, C. A. B. Boucher, R. Molenkamp, M. P. G. Koopmans, C. Geurtsvankessel, A. A. van der Eijk, Duration and key determinants of infectious virus shedding in hospitalized patients with coronavirus disease-2019 (COVID-19). *Nat Commun*. **12**, 267 (2021).
44. E. S. Savelle, A. Winnett, A. E. Romano, M. K. Porter, N. Shelby, R. Akana, J. Ji, M. M. Cooper, N. W. Schlenker, J. A. Reyes, A. M. Carter, J. T. Barlow, C. Tognazzini, M. Feaster, Y.-Y. Goh, R. F. Ismagilov, “SARS-CoV-2 is detectable using sensitive RNA saliva testing days before viral load reaches detection range of low-sensitivity nasal swab tests” (preprint, Infectious Diseases (except HIV/AIDS), 2021), , doi:10.1101/2021.04.02.21254771.
45. R. L. Smith, L. L. Gibson, P. P. Martinez, R. Ke, A. Mirza, M. Conte, N. Gallagher, A. Conte, L. Wang, R. Fredrickson, D. C. Edmonson, M. E. Baughman, K. K. Chiu, H. Choi, T. W. Jensen, K. R. Scardina, S. Bradley, S. L. Gloss, C. Reinhart, J. Yedetore, A. N. Owens, J. Broach, B. Barton, P. Lazar, D. Hennessy, T. Young, A. Dunnett, M. L. Robinson, H. H. Mostafa, A. Pekosz, Y. C. Manabe, W. J. Heetderks, D. D. McManus, C. B. Brooke, Longitudinal assessment of diagnostic test performance over the course of acute SARS-CoV-2 infection. *The Journal of Infectious Diseases*, jia337 (2021).
46. NIH COVID-19 Autopsy Consortium, HCA Oral and Craniofacial Biological Network, N. Huang, P. Pérez, T. Kato, Y. Mikami, K. Okuda, R. C. Gilmore, C. D. Conde, B. Gasmi, S. Stein, M. Beach, E. Pelayo, J. O. Maldonado, B. A. Lafont, S.-I. Jang, N. Nasir, R. J. Padilla, V. A. Murrah, R. Maile, W. Lovell, S. M. Wallet, N. M. Bowman, S. L. Meinig, M. C. Wolfgang, S. N. Choudhury, M. Novotny, B. D. Aevermann, R. H. Scheuermann, G. Cannon, C. W. Anderson, R. E. Lee, J. T. Marchesan, M. Bush, M. Freire, A. J. Kimple, D. L. Herr, J. Rabin, A. Grazioli, S. Das, B. N. French, T. Pranzatelli, J. A. Chiorini, D. E.

- Kleiner, S. Pittaluga, S. M. Hewitt, P. D. Burbelo, D. Chertow, K. Frank, J. Lee, R. C. Boucher, S. A. Teichmann, B. M. Warner, K. M. Byrd, SARS-CoV-2 infection of the oral cavity and saliva. *Nat Med* (2021), doi:10.1038/s41591-021-01296-8.
47. S. S. Lakdawala, A. Jayaraman, R. A. Halpin, E. W. Lamirande, A. R. Shih, T. B. Stockwell, X. Lin, A. Simenauer, C. T. Hanson, L. Vogel, M. Paskel, M. Minai, I. Moore, M. Orandle, S. R. Das, D. E. Wentworth, R. Sasisekharan, K. Subbarao, The soft palate is an important site of adaptation for transmissible influenza viruses. *Nature*. **526**, 122–125 (2015).
 48. Y. Cai, J. Zhang, T. Xiao, C. L. Lavine, S. Rawson, H. Peng, H. Zhu, K. Anand, P. Tong, A. Gautam, S. Lu, S. M. Sterling, R. M. Walsh, S. Rits-Volloch, J. Lu, D. R. Wesemann, W. Yang, M. S. Seaman, B. Chen, Structural basis for enhanced infectivity and immune evasion of SARS-CoV-2 variants. *Science*, eabi9745 (2021).
 49. L. G. Thorne, M. Bouhaddou, A.-K. Reuschl, L. Zuliani-Alvarez, B. Polacco, A. Pelin, J. Batra, M. V. X. Whelan, M. Ummadi, A. Rojc, J. Turner, K. Obernier, H. Braberg, M. Soucheray, A. Richards, K.-H. Chen, B. Harjai, D. Memon, M. Hosmillo, J. Hiatt, A. Jahun, I. G. Goodfellow, J. M. Fabius, K. Shokat, N. Jura, K. Verba, M. Noursadeghi, P. Beltrao, D. L. Swaney, A. Garcia-Sastre, C. Jolly, G. J. Towers, N. J. Krogan, “Evolution of enhanced innate immune evasion by the SARS-CoV-2 B.1.1.7 UK variant” (preprint, *Microbiology*, 2021), , doi:10.1101/2021.06.06.446826.
 50. J. C. Brown, D. H. Goldhill, J. Zhou, T. P. Peacock, R. Frise, N. Goonawardane, L. Baillon, R. Kugathasan, A. L. Pinto, P. F. McKay, J. Hassard, M. Moshe, A. Singanayagam, T. Burgoyne, the ATACCC Investigators, PHE Virology Consortium, W. S. Barclay, “Increased transmission of SARS-CoV-2 lineage B.1.1.7 (VOC 202012/01) is not accounted for by a replicative advantage in primary airway cells or antibody escape” (preprint, *Microbiology*, 2021), , doi:10.1101/2021.02.24.432576.
 51. L. Ulrich, N. J. Halwe, A. Taddeo, N. Ebert, J. Schön, C. Devisme, B. S. Trüeb, B. Hoffmann, M. Wider, M. Bekliz, M. Essaidi-Laziosi, M. L. Schmidt, D. Niemeyer, V. M. Corman, A. Kraft, A. Godel, L. Laloli, J. N. Kelly, A. Breithaupt, C. Wylezich, I. B. Veiga, M. Gultom, K. Adea, B. Meyer, C. Eberhardt, L. Thomann, M. Gsell-Albert, F. Labroussaa, J. Jores, A. Summerfield, C. Drosten, I. A. Eckerle, R. Dijkman, D. Hoffmann, V. Thiel, M. Beer, C. Benarafa, “Enhanced fitness of SARS-CoV-2 variant of concern B.1.1.7, but not B.1.351, in animal models” (preprint, *Microbiology*, 2021), , doi:10.1101/2021.06.28.450190.
 52. M. Cevik, M. Tate, O. Lloyd, A. E. Maraolo, J. Schafers, A. Ho, SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. *The Lancet Microbe*. **2**, e13–e22 (2021).
 53. V. Gniazdowski, C. P. Morris, S. Wohl, T. Mehoke, S. Ramakrishnan, P. Thielen, H. Powell, B. Smith, D. T. Armstrong, M. Herrera, C. Reifsnyder, M. Sevdali, K. C. Carroll, A. Pekosz, H. H. Mostafa, Repeat COVID-19 Molecular Testing: Correlation of SARS-CoV-2 Culture with Molecular Assays and Cycle Thresholds. *Clin Infect Dis* (2020), doi:10.1093/cid/ciaa1616.
 54. S. Matsuyama, N. Nao, K. Shirato, M. Kawase, S. Saito, I. Takayama, N. Nagata, T. Sekizuka, H. Kato, F. Kato, M. Sakata, M. Tahara, S. Kutsuna, N. Ohmagari, M. Kuroda,

T. Suzuki, T. Kageyama, M. Takeda, Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. *Proc Natl Acad Sci U S A*. **117**, 7001–7003 (2020).

55. J. J. Waggoner, V. Stittleburg, R. Pond, Y. Saklawi, M. K. Sahoo, A. Babiker, L. Hussaini, C. S. Kraft, B. A. Pinsky, E. J. Anderson, N. Rouphael, Triplex Real-Time RT-PCR for Severe Acute Respiratory Syndrome Coronavirus 2. *Emerg. Infect. Dis.* **26**, 1633–1635 (2020).
56. A. Rambaut, E. C. Holmes, Á. O’Toole, V. Hill, J. T. McCrone, C. Ruis, L. du Plessis, O. G. Pybus, A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol.* **5**, 1403–1407 (2020).

Supplemental figure legends:

Fig S1: Remainder of individual plots. Plots of longitudinal assay results from study participants not shown in Figure 1A. Single asterisk next to the participant ID indicates B.1.1.7 variant infection, while double asterisks indicate P1 variant infection.

Fig S2: Individual-level symptom data. Self-reported symptom data from study participants, overlaid with viral culture status. Participants were asked to complete a survey through the Eureka digital study platform inquiring about the presence or absence of the indicated set of symptoms each day after sample collection.

Fig S3: Comparison of symptoms and viral culture status. Plots show the proportions of either viral culture negative or viral culture positive days for which participants reported the indicated symptoms. Data are only shown for individuals who reported the indicated symptom at least once.

Fig S4: Model structures. Diagrams showing the structures of the additional three models (not shown in Figure 2A) considered for describing viral load data. See Supporting Text for descriptions of the models.

Fig S5: Model parameter estimates as a function of age. Plots showing the relationship between age and the indicated model parameter estimates for **(A)** the refractory cell model (nasal data) and the **(B)** the immune effector cell model (saliva data). Linear regressions were performed on the data. R^2 values and p-values are shown.

Fig S6: Analysis workflow. Diagram indicating how empirical RTqPCR and viral culture data were used to generate estimations of individual level viral dynamics and infectiousness.

Fig S7: The saturation model accurately predicts the cell culture positivity data. Lines denote the predicted probability of cell culture being positive. Dots denotes cell culture positivity data, where a dot is at 1 or 0 when the cell culture is positive or negative, respectively.

Fig S8: Individual infectiousness plots. Estimated infectiousness over time plotted for individual study participants. Dashed lines indicate inferred peak in infectiousness.

Supplemental tables:

Table S1: Study cohort demographics data

Table S2: Comparing model fits to longitudinal viral load data from nasal swab samples using AIC scores

Table S3: Comparing model fits to longitudinal viral load data from saliva samples using AIC scores

Table S4: Estimated population parameter values from the best models

Table S5: Individual parameter values estimated from the best fit models

Table S6: The fixed parameters and their values in all viral dynamic models

Table S7. Model comparison using AIC scores of three models describing the relationship between the viral genome load and cell culture positivity

Table S8. The estimated parameters in the best-fit model, i.e. the saturation model, to the paired RTqPCR and cell culture positivity data

Table S9: Viral genotype data