

1 **SARS-CoV-2 viral dynamics in acute infections**

2
3 Stephen M. Kissler^{*1}, Joseph R. Fauver^{*2}, Christina Mack^{*3,4}, Scott W. Olesen¹, Caroline Tai³, Kristin Y.
4 Shiue^{3,4}, Chaney C. Kalinich², Sarah Jednak⁵, Isabel M. Ott², Chantal B.F. Vogels², Jay Wohlgemuth⁶,
5 James Weisberger⁷, John DiFiori⁸, Deverick J. Anderson⁹, Jimmie Mancell¹⁰, David D. Ho¹¹, Nathan D.
6 Grubaugh^{†2}, Yonatan H. Grad^{†1}

7
8 ¹Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health,
9 Boston, MA

10 ²Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT

11 ³IQVIA, Real World Solutions, Durham, NC

12 ⁴Department of Epidemiology, University of North Carolina-Chapel Hill, Chapel Hill, NC

13 ⁵Department of Health Management and Policy, University of Michigan School of Public Health, Ann
14 Arbor, MI

15 ⁶Quest Diagnostics, San Juan Capistrano, CA

16 ⁷Bioreference Laboratories, Elmwood Park, NJ

17 ⁸Hospital for Special Surgery, and the National Basketball Association, New York, NY

18 ⁹Duke Center for Antimicrobial Stewardship and Infection Prevention, Durham, NC

19 ¹⁰Department of Medicine, University of Tennessee Health Science Center, Memphis, TN

20 ¹¹Aaron Diamond AIDS Research Center, Columbia University Vagelos College of Physicians and
21 Surgeons, New York, NY

22
23 * denotes equal contribution

24 † denotes co-senior authorship

25 Correspondence and requests for materials should be addressed to:

26 Email: ygrad@hsph.harvard.edu

27 Telephone: 617.432.2275

31 **Research in Context.**

32

33 **Evidence before this study.**

34 SARS-CoV-2 viral dynamics affect clinical and public health measures, informing patient care, testing
35 algorithms, contact tracing protocols, and clinical trial design. We searched Web of Science using the search
36 terms “ALL = ((SARS-CoV-2 OR COVID-19) AND (viral OR RNA) AND (load OR concentration OR
37 shedding) AND (dynamic* OR kinetic* OR trajector*))” which returned 83 references. Of these, 22 were
38 not pertinent to within-host SARS-CoV-2 viral dynamics. The remaining 61 studies tracked SARS-CoV-2
39 viral trajectories in a variety of geographic locations and patient populations. Together, these studies report
40 that viral titers normally peak at or before the onset of symptoms and that a long tail of intermittent positive
41 tests can follow a period of acute infection. Plasma but not nasopharyngeal viral concentration is associated
42 with increased disease severity. Most studies tracked hospitalized patients after the onset of symptoms. Two
43 of the studies tracked pre-symptomatic and/or asymptomatic patients, but these were too sparsely sampled
44 to clearly discern viral dynamics during the earliest stage of infection.

45

46 **Added value of this study.**

47 We implemented prospective longitudinal real time quantitative reverse transcriptase polymerase chain
48 reaction (RT-qPCR) testing for SARS-CoV-2 in a cohort of individuals during the resumption of the 2019-
49 20 National Basketball Association season. This allowed us to explicitly measure viral titers during the full
50 course of 46 acute infections. Consistent with other studies, we find that peak viral concentrations do not
51 differ substantially between symptomatic and asymptomatic individuals but that symptomatic individuals
52 take longer to clear the virus than asymptomatic individuals. For both symptomatic and asymptomatic
53 individuals, viral titers normally peak within 3 days of the first positive test. This study is the first to describe
54 the time course of viral concentrations during the earliest stage of infection when individuals are most likely
55 to be infectious.

56

57 **Implications of all the available evidence.**

58 Symptomatic and asymptomatic individuals follow similar SARS-CoV-2 viral trajectories. Due to the rapid
59 progression from first possible detection to peak viral concentration, frequent rapid-turnaround testing is
60 needed to screen individuals prior to them becoming infectious.

61 **Abstract (250 words)**

62

63 **Background.** SARS-CoV-2 infections are characterized by viral proliferation and clearance phases and
64 can be followed by low-level viral RNA shedding. The dynamics of viral RNA concentration, particularly
65 in the early stages of infection, can inform clinical measures and interventions such as test-based screening.

66

67 **Methods.** We used prospective longitudinal RT-qPCR testing to measure the viral RNA trajectories for 68
68 individuals during the resumption of the 2019-20 National Basketball Association season. For 46
69 individuals with acute infections, we inferred the peak viral concentration and the duration of the viral
70 proliferation and clearance phases.

71

72 **Findings.** On average, viral RNA concentrations peaked 2.7 days (95% credible interval [1.2, 3.8]) after
73 first detection at a cycle threshold value of 22.4 [20.6, 24.1]. The viral clearance phase lasted longer for
74 symptomatic individuals (10.5 days [6.5, 14.0]) than for asymptomatic individuals (6.7 days [3.2, 9.2]). A
75 second test within 2 days after an initial positive PCR substantially improves certainty about a patient's
76 infection phase. The effective sensitivity of a test intended to identify infectious individuals declines
77 substantially with test turnaround time.

78

79 **Interpretation.** SARS-CoV-2 viral concentrations peak rapidly regardless of symptoms. Sequential tests
80 can help reveal a patient's progress through infection stages. Frequent rapid-turnaround testing is needed
81 to effectively screen individuals before they become infectious.

82

83 **Funding.** NWO Rubicon 019.181EN.004 (CBFV); clinical research agreement with the NBA and NBPA
84 (NDG); Huffman Family Donor Advised Fund (NDG); Fast Grant funding from the Emergent Ventures at
85 the Mercatus Center; George Mason University (NDG); the Morris-Singer Fund for the Center for
86 Communicable Disease Dynamics at the Harvard T.H. Chan School of Public Health (YHG).

87 **Introduction.**

88

89 As mortality from the COVID-19 pandemic surpasses one million, SARS-CoV-2 continues to cause
90 hundreds of thousands of daily new infections¹. A critical strategy to curb the spread of the virus without
91 imposing widespread lockdowns is to rapidly identify and isolate infectious individuals. Because symptoms
92 are an unreliable indicator of infectiousness and infections are frequently asymptomatic², testing is key to
93 determining whether a person is infected and may be contagious.

94

95 Real time quantitative reverse transcriptase polymerase chain reaction (RT-qPCR) tests are the gold
96 standard for detecting SARS-CoV-2 infection. Normally, these tests yield a binary positive/negative
97 diagnosis based on detection of viral RNA. However, they can also quantify the viral titer via the cycle
98 threshold (Ct). The Ct is the number of thermal cycles needed to amplify sampled viral RNA to a detectable
99 level: the higher the sampled viral RNA concentration, the lower the Ct. This inverse correlation between
100 Ct and viral concentration makes RT-qPCR tests far more valuable than a binary diagnostic, as they can be
101 used to reveal a person's progress through key stages of infection³, with the potential to assist clinical and
102 public health decision-making. However, the dynamics of the Ct during the earliest stages of infection,
103 when contagiousness is rapidly increasing, have been unclear, because diagnostic testing is usually
104 performed after the onset of symptoms, when viral RNA concentration has peaked and already begun to
105 decline, and performed only once^{4,5}. Without a clear picture of the course of SARS-CoV-2 viral
106 concentrations across the full duration of acute infection, it has been impossible to specify key elements of
107 testing algorithms such as the frequency rapid at-home testing⁶ that will be needed to reliably screen
108 infectious individuals before they transmit infection. Here, we fill this gap by analyzing the prospective
109 longitudinal SARS-CoV-2 RT-qPCR testing performed for players, staff, and vendors during the
110 resumption of the 2019-20 National Basketball Association (NBA) season.

111

112 **Methods.**

113 Data collection.

114 The study period began in teams' local cities from June 23rd through July 9th, 2020, and testing continued
115 for all teams as they transitioned to Orlando, Florida through September 7th, 2020. A total of 68 individuals
116 (90% male) were tested at least five times during the study period and recorded at least one positive test
117 with Ct value <40. Many individuals were being tested daily as part of Orlando campus monitoring. Due
118 to a lack of new infections among players and team staff after clearing quarantine in Orlando, all players
119 and team staff included in the results pre-date the Orlando phase of the restart. A diagnosis of "acute" or
120 "persistent" infection was abstracted from physician records. "Acute" denoted a likely new infection.
121 "Persistent" indicated the presence of virus in a clinically recovered individual, likely due to infection that
122 developed prior to the onset of the study. There were 46 acute infections; the remaining 22 individuals were
123 assumed to be persistently shedding SARS-CoV-2 RNA due to a known infection that occurred prior to the
124 study period. This persistent RNA shedding can last for weeks after an acute infection and likely represented
125 non-infectious viral RNA⁷. Of the individuals included in the study, 27 of the 46 with acute infections and
126 40 of the 68 overall were from staff and vendors. The Ct values for all tests for the 68 individuals included
127 in the analysis with their designations of acute or persistent infection are depicted in **Supplemental Figures**
128 **1–4**. A schematic diagram of the data collection and analysis pipeline is given in **Figure 1**.

129

130 Statistical analysis.

131 We used a Bayesian statistical model to infer the peak Ct value and the durations of the proliferation and
132 clearance stages for the 46 acute infections with at least one Ct value below 35 (**Figure 1**; **Supplemental**
133 **Methods**). We assumed that the viral concentration trajectories consisted of a proliferation phase, with
134 exponential growth in viral RNA concentration, followed by a clearance phase characterized by exponential
135 decay in viral RNA concentration⁸. Since Ct values are roughly proportional to the negative logarithm of

136 viral concentration³, this corresponds to a linear decrease in Ct followed by a linear increase. We therefore
137 constructed a piecewise-linear regression model to estimate the peak Ct value, the time from infection onset
138 to peak (*i.e.*, the duration of the proliferation stage), and the time from peak to infection resolution (*i.e.*, the
139 duration of the clearance stage). We estimated the parameters of the regression model by fitting to the
140 available data using a Hamiltonian Monte Carlo algorithm⁹ yielding simulated draws from the Bayesian
141 posterior distribution for each parameter. Full details on the fitting procedure are given in the
142 **Supplemental Methods**. Code is available at <https://github.com/gradlab/CtTrajectories>.

143

144 Inferring stage of infection.

145 Next, we determined whether individual or paired Ct values can reveal whether an individual is in the
146 proliferation, clearance, or persistent stage of infection. To do so, we extracted all observed Ct values within
147 a 5-unit window (*e.g.*, between 30 and 35 Ct) and measured how frequently these values sat within the
148 proliferation stage, the clearance stage, or the persistent stage. We measured these frequencies across
149 10,000 posterior parameter draws to account for that fact that Ct values near stage transitions (*e.g.*, near the
150 end of the clearance stage) could be assigned to different infection stages depending on the parameter values
151 (see **Figure 1**, bottom-right). We did this for 23 windows with midpoint spanning from Ct = 37.5 to Ct =
152 15.5 in increments of 1 Ct.

153

154 To calculate the probability that a Ct value falling within the 5-unit window corresponded to an acute
155 infection (*i.e.*, either the proliferation or the clearance stage), we summed the proliferation and clearance
156 frequencies for all samples within that window and divided by the total number of samples in the window.
157 We similarly calculated the probability that a Ct falling within the 5-unit window corresponded to just the
158 proliferation phase.

159

160 To assess the information gained by conducting a second test within two days of an initial positive, we
161 restricted our attention to all samples within the 5-unit window that had a subsequent sample taken within
162 two days. We repeated the above calculations, stratifying by whether the second test had a higher or lower
163 Ct than the first.

164

165 Measuring the effective sensitivity of screening tests.

166 The sensitivity of a test is defined as the probability that the test correctly identifies an individual who is
167 positive for some criterion of interest. For clinical diagnostic SARS-CoV-2 tests, the criterion of interest is
168 current infection with SARS-CoV-2. However, a common goal is to predict infectiousness at some point in
169 the future, as in the context of test-based screening prior to a social gathering. The ‘effective sensitivity’ of
170 a test in this context (*i.e.*, its ability to predict future infectiousness) may differ substantially from its clinical
171 sensitivity (*i.e.*, its ability to detect current infection). A test’s effective sensitivity depends on its inherent
172 characteristics, such as its limit of detection and sampling error rate, as well as the viral dynamics of infected
173 individuals.

174

175 To illustrate this, we estimated the effective sensitivity of (a) a test with limit of detection of 40 Ct and a
176 1% sampling error probability (akin to RT-qPCR), and (b) a test with limit of detection of 35 Ct a 5%
177 sampling error probability (akin to some rapid antigen tests). We measured the frequency with which these
178 tests would successfully screen an individual who would become infectious (assumed to correspond to Ct
179 ≤ 30)¹⁰ at the time of a gathering when the test was administered between 0 and 3 days prior to the gathering.
180 We measured this effective sensitivity with respect to 2,500 simulated viral concentration trajectories
181 sampled from the posterior distributions from the best-fit model, restricting to trajectories with peak Ct
182 below 30 (any samples with peak Ct > 30 would, by these criteria, never be infectious and so would not
183 factor into the sensitivity calculation). Full details are given in the **Supplemental Methods**.

184

185 We also estimated the number of individuals who would be expected to arrive at a 1,000-person gathering
186 while infectious assuming a 2% prevalence of infectiousness in the population under each testing strategy.
187 To do so, we simulated 1,000 events for which the number of infectious individuals at the event was
188 simulated using a Binomial distribution with 1,000 draws and ‘success’ probability of 0.02. Then, we
189 sampled the number of these infectious individuals who would be screened using a second Binomial draw
190 with ‘success’ probability dictated by the previously calculated effective sensitivity. Further details are
191 given in the **Supplemental Methods**. To facilitate the exploration of different scenarios, we have generated
192 an online tool ([https://stephenkissler.shinyapps.io /shiny/](https://stephenkissler.shinyapps.io/shiny/)) where users can input test and population
193 characteristics and calculate the effective sensitivity and expected number of infectious individuals at a
194 gathering.
195

196 Results.

197 Of the 46 individuals with acute infections, 13 reported symptoms at the time of diagnosis; the timing of
198 the onset of symptoms was not recorded. The mean peak Ct value for symptomatic individuals was 22.2
199 (95% credible interval [19.1, 25.1]), the mean duration of the proliferation phase was 2.9 days [0.7, 4.7],
200 and the mean duration of clearance was 10.5 days [6.5, 14.0] (**Figure 2**). This compares with 22.4 Ct [20.2,
201 24.5], 3.0 days [1.3, 4.3], and 6.7 days [3.2, 9.2], respectively, for individuals who did not report symptoms
202 at the time of diagnosis (**Figure 2**). This yielded a slightly longer overall duration of acute infection for
203 individuals who reported symptoms (13.4 days [9.3, 17.1]) vs. those who did not (9.7 days [6.0, 12.5]). For
204 all individuals regardless of symptoms, the mean peak Ct value, proliferation duration, clearance duration,
205 and duration of acute shedding were 22.4 Ct [20.6, 24.1], 2.7 days [1.2, 3.8], 7.4 days [3.9, 9.6], and 10.1
206 days [6.5, 12.6] (**Supplemental Figure 5**). There was a substantial amount of individual-level variation in
207 the peak Ct value and the proliferation and clearance stage durations (**Supplemental Figures 6–11**).
208

209 Using the full dataset of 68 individuals, we estimated the frequency with which a given Ct value was
210 associated with an acute infection (*i.e.*, the proliferation or clearance phase, but not the persistence phase),
211 and if so, the probability that it was associated with the proliferation stage alone. The probability of an acute
212 infection increased rapidly with decreasing Ct (increasing viral load), with Ct < 30 virtually guaranteeing
213 an acute infection in this dataset (**Figure 3A**). However, a single Ct value provided little information about
214 whether an acute infection was in the proliferation or the clearance stage (**Figure 3B**).
215

216 We assessed whether a second test within two days of the first could improve these predictions. A positive
217 test followed by a second test with lower Ct (higher viral RNA concentration) was slightly more likely to
218 be associated with an active infection than a positive test alone (**Figure 3C**). Similarly, a positive test
219 followed by a second test with lower Ct (higher viral RNA concentration) was much more likely to be
220 associated with the proliferation phase than with the clearance phase (**Figure 3D**).
221

222 We next estimated how the effective sensitivity of a pre-event screening test declines with increasing time
223 to the event. For a test with limit of detection of 40 Ct and a 1% chance of sampling error, the effective
224 sensitivity declines from 99% when the test coincides with the start of the event to 76% when the test is
225 administered two days prior to the event (**Figure 4A**), assuming a threshold of infectiousness at 30 Ct¹⁰.
226 This two-day-ahead sensitivity is slightly lower than the effective sensitivity of a test with a limit of
227 detection at 35 Ct and a 5% sampling error administered one day before the event (82%), demonstrating
228 that limitations in testing technology can be compensated for by reducing turnaround time. Using these
229 effective sensitivities, we estimated the number of infectious individuals expected to arrive at an event with
230 1,000 people when 2% of the population is infectious. Just as the effective sensitivity declines with time to
231 the event, the predicted number of infectious individuals rises (**Figure 4B**).
232

233

234 **Discussion.**

235

236 This report provides first comprehensive data on the early-infection RT-qPCR Ct dynamics associated with
237 SARS-CoV-2 infection. Viral titers increase quickly, normally within 3 days of the first possible RT-qPCR
238 detection, regardless of symptoms. Our findings highlight that repeated PCR tests can be used to infer the
239 stage of a patient's infection. While a single test can inform on whether a patient is in the acute or persistent
240 viral RNA shedding stages, a subsequent test can help identify whether viral RNA concentrations are
241 increasing or decreasing, thus informing clinical care. We also show that the effective sensitivity of pre-
242 event screening tests declines rapidly with test turnaround time due to the rapid progression from
243 detectability to peak viral titers. Due to the transmission risk posed by large gatherings¹¹, the trade-off
244 between test speed and sensitivity must be weighed carefully. These data offer the first direct measurements
245 capable of informing such decisions.

246

247 Our findings on the duration of SARS-CoV-2 viral RNA shedding expand on and agree with previous
248 studies^{12,13} and with observations that peak Ct does not differ substantially between symptomatic and
249 asymptomatic individuals⁴. While previous studies have largely relied on serial sampling of admitted
250 hospital patients, our study used prospective sampling of ambulatory infected individuals to characterize
251 complete viral dynamics for the presymptomatic stage and for individuals who did not report symptoms.
252 This allowed us to assess differences between the viral RNA proliferation and clearance stages for
253 individuals with and without reported symptoms. The similarity in the early-infection viral RNA dynamics
254 for both symptomatic and asymptomatic individuals underscores the need for SARS-CoV-2 screening
255 regardless of symptoms. The rapid progression from a negative test to a peak Ct value 2-4 days later
256 provides empirical support for screening and surveillance strategies that employ frequent rapid testing to
257 identify potentially infectious individuals^{14,15}. Taken together, the dynamics of viral RNA shedding
258 substantiate the need for frequent population-level SARS-CoV-2 screening and a greater availability of
259 diagnostic tests.

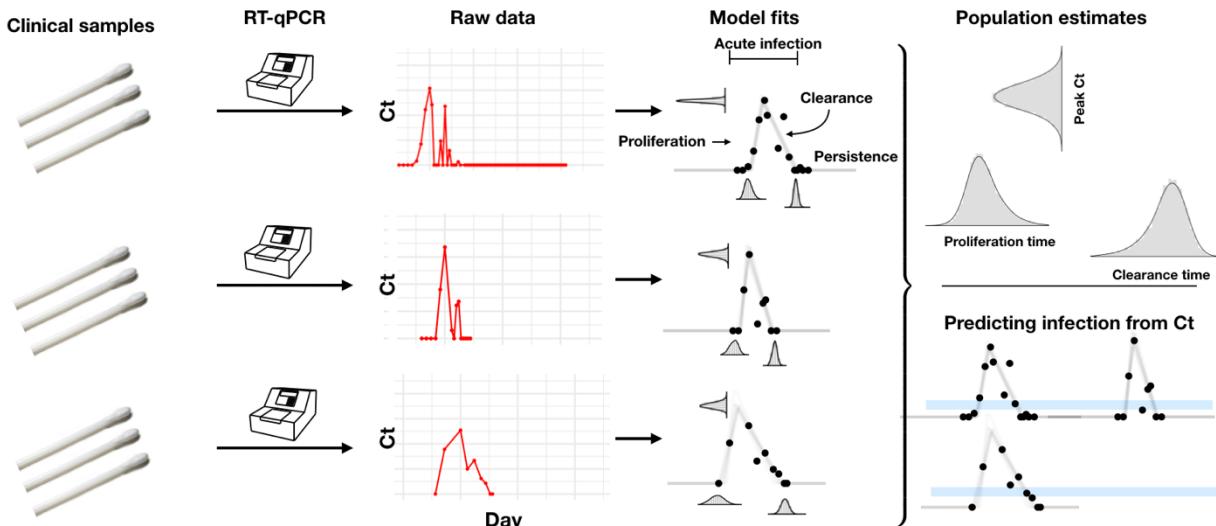
260

261 Our findings are limited for several reasons. The cohort does not constitute a representative sample from
262 the population, as it was a predominantly male, healthy, young population inclusive of professional athletes.
263 Some of the trajectories were sparsely sampled, limiting the precision of our posterior estimates. Symptom
264 reporting was imperfect, particularly after initial evaluation as follow-up during course of disease was not
265 systematic for all individuals. As with all predictive tests, the probabilities that link Ct values with infection
266 stages (**Figure 3**) pertain to the population from which they were calibrated and do not necessarily
267 generalize to other populations for which the prevalence of infection and testing protocols may differ. Still,
268 we anticipate that the central patterns will hold across populations: first, that low Cts (<30) strongly predict
269 acute infection, and second, that a follow-up test collected within two days of an initial positive test can
270 substantially help to discern whether a person is closer to the beginning or the end of their infection. Our
271 study did not test for the presence of infectious virus, though previous studies have documented a close
272 inverse correlation between Ct values and culturable virus¹⁰. Our assessment of pre-event testing assumed
273 that individuals become infectious immediately upon passing a threshold and that this threshold is the same
274 for the proliferation and for the clearance phase. In reality, the threshold for infectiousness is unlikely to be
275 at a fixed viral concentration for all individuals and may be at a higher Ct/lower viral concentration during
276 the proliferation stage than during the clearance stage. Further studies that measure culturable virus during
277 the various stages of infection and that infer infectiousness based on contact tracing combined with
278 prospective longitudinal testing will help to clarify the relationship between viral concentration and
279 infectiousness.

280

281 To manage the spread of SARS-CoV-2, we must develop novel technologies and find new ways to extract
282 more value from the tools that are already available. Our results suggest that integrating the quantitative

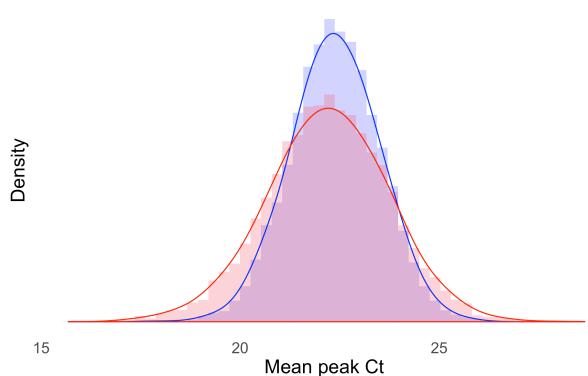
283 viral RNA trajectory into algorithms for clinical management could offer benefits. The ability to chart a
284 patient's progress through their infection underpins our ability to provide appropriate clinical care and to
285 institute effective measures to reduce the risk of onward transmission. Marginally more sophisticated
286 diagnostic and screening algorithms may greatly enhance our ability to manage the spread of SARS-CoV-
287 2 using tests that are already available.



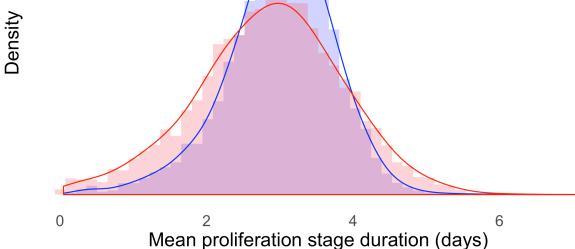
288
289
290

291 **Figure 1. Illustration of the analysis pipeline.** Combined anterior nares and oropharyngeal swabs were tested using a RT-qPCR
292 assay to generate longitudinal Ct values ('Raw data', red points) for each person. Using a statistical model, we estimated Ct
293 trajectories consistent with the data, represented by the thin lines under the 'Model fits' heading. These produced posterior
294 probability distributions for the peak Ct value, the duration of the proliferation phase (infection onset to peak Ct), and the duration
295 of the clearance phase (peak Ct to resolution of acute infection) for each person. We estimated population means for these quantities.
296 The model fits also allowed us to determine how frequently a given Ct value or pair of Ct values within a five-unit window (blue
297 bars, bottom-right pane) was associated with the proliferation phase, the clearance phase, or a persistent infection.

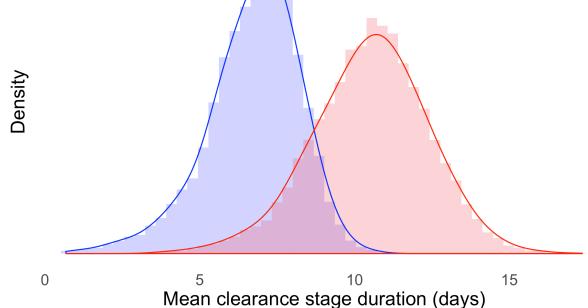
298 A)



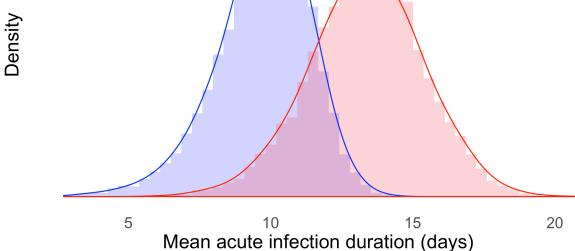
B)



299 C)

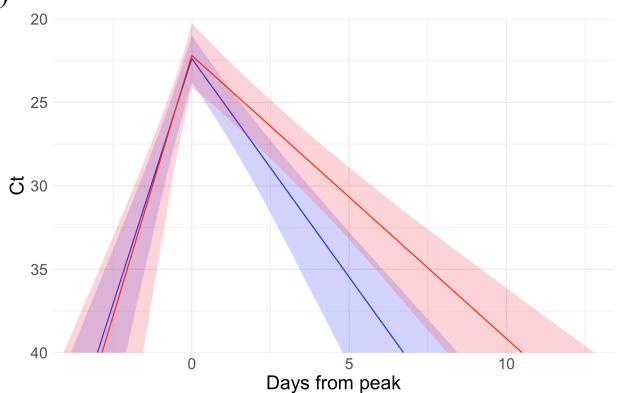


D)



301
302
303

E)

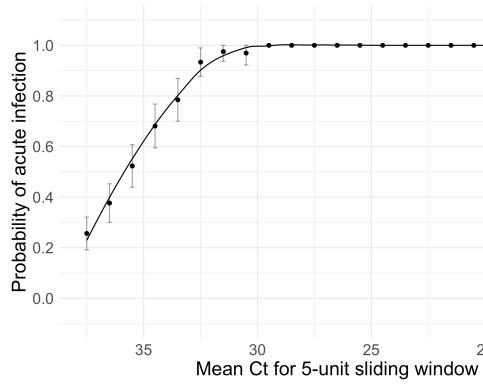


304
305
306
307

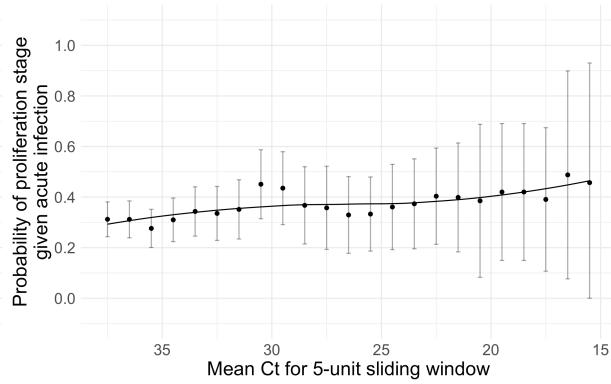
308 **Figure 2. Peak Ct value and infection stage duration distributions according to symptoms reported at time of diagnosis.**
309 Histograms (colored bars) of 10,000 simulated draws from the posterior distributions for mean peak Ct value (A),
310 mean duration of the proliferation stage (infection detection to peak Ct, B), mean duration of the clearance stage
311 (peak Ct to resolution of acute RNA shedding, C), and total duration of acute shedding (D) across the 46 individuals with an acute infection.
312 The histograms are separated according to whether the person reported symptoms (red, 13 individuals)
313 or did not report symptoms (blue, 33 individuals). The red and blue curves are kernel density estimators for the histograms
314 to assist with visualizing the shapes of the histograms. The mean Ct trajectory corresponding to the mean values for peak Ct,
315 proliferation duration, and clearance duration for symptomatic vs. asymptomatic individuals is depicted in (E) (solid lines), where shading depicts the 90% credible intervals.

316

A)



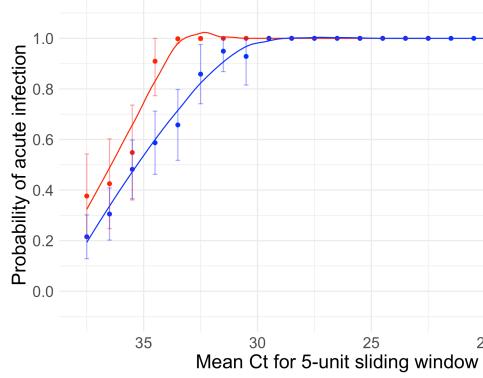
B)



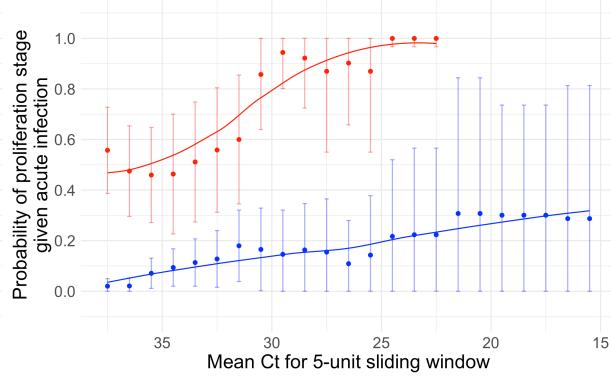
317

318

C)



D)



319

— Ct decrease/viral load increase — Ct increase/viral load decrease

320

321

322

323

324

325

326

327

Figure 3. Relationship between single/paired Ct values and infection stage. Probability that a given Ct value lying within a 5-unit window (horizontal axis) corresponds to an acute infection (A, C) or to the proliferation phase of infection assuming an acute infection (B, D). Sub-figures A and B depict the predictive probabilities for a single Ct value, while sub-figures C and D depict the predictive probabilities for a positive test paired with a subsequent test with either lower (red) or higher (blue) Ct. The curves are LOESS smoothing curves to better visualize the trends. Error bars represent the 90% Wald confidence interval.

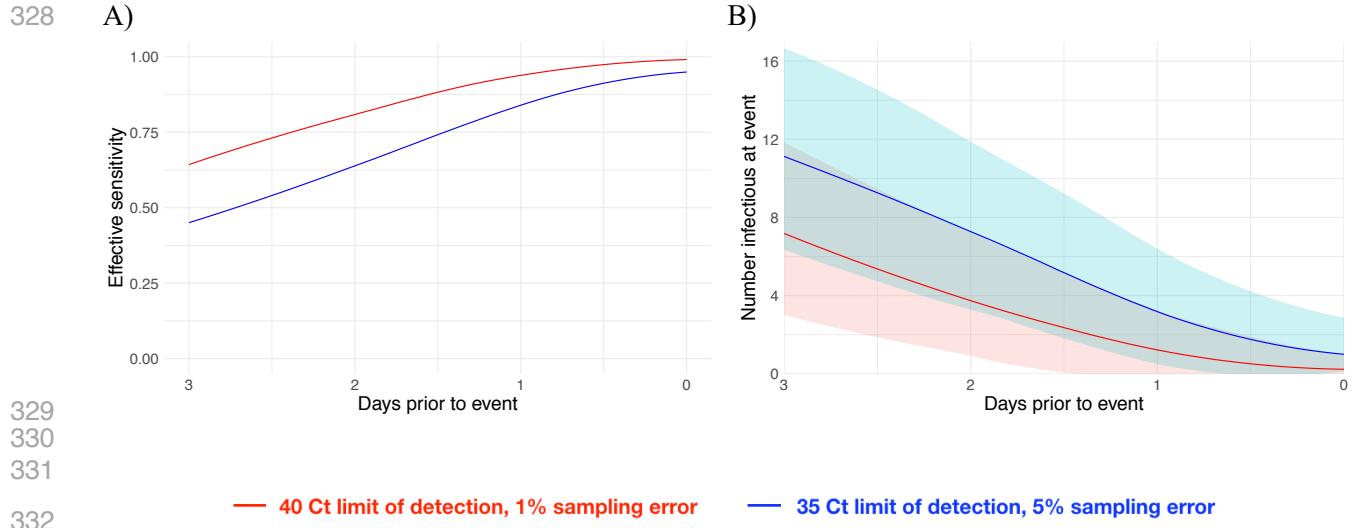


Figure 4. Effective sensitivity and expected number of infectious event attendees for tests with varying sensitivity. Effective sensitivity for a test with limit of detection of 40 Ct and 1% sampling error probability (red) and 35 Ct and 5% sampling error probability (blue) (A). Number of infectious individuals expected to attend an event of size 1000 assuming a population prevalence of 2% infectious individuals for a test with limit of detection of 40 Ct and 1% sampling error probability (red) and 35 Ct and 5% sampling error probability (blue). Shaded bands represent 90% prediction intervals generated from the quantiles of 1,000 simulated events and capture uncertainty both in the number of infectious individuals who would arrive at the event in the absence of testing and in the probability that the test successfully identifies infectious individuals.

341 **References**

- 342
- 343 1. World Health Organization. Coronavirus disease (COVID-19) Situation Report.
- 344 2. Furukawa NW, Brooks JT, Sobel J. Evidence Supporting Transmission of Severe Acute
- 345 Respiratory Syndrome Coronavirus 2 While Presymptomatic or Asymptomatic. *Emerg Infect Dis*.
- 346 2020;26(7). doi:10.3201/eid2607.201595
- 347 3. Tom MR, Mina MJ. To Interpret the SARS-CoV-2 Test, Consider the Cycle Threshold Value.
- 348 *Clin Infect Dis*. 2020;02115(Xx):1-3. doi:10.1093/cid/ciaa619
- 349 4. Walsh KA, Jordan K, Clyne B, et al. SARS-CoV-2 detection, viral load and infectivity over the
- 350 course of an infection. *J Infect*. 2020;81(3):357-371. doi:10.1016/j.jinf.2020.06.067
- 351 5. Wyllie AL, Fournier J, Casanovas-Massana A, et al. Saliva or Nasopharyngeal Swab Specimens
- 352 for Detection of SARS-CoV-2. *N Engl J Med*. 2020;February(Coorespondance):NEJMc2016359.
- 353 doi:10.1056/NEJMc2016359
- 354 6. Larremore DB, Wilder B, Lester E, et al. Test sensitivity is secondary to frequency and turnaround
- 355 time for COVID-19 surveillance. *Sci Adv*. 2020. doi:10.1126/sciadv.abd5393
- 356 7. Xiao AT, Tong YX, Zhang S. Profile of RT-PCR for SARS-CoV-2: A Preliminary Study From 56
- 357 COVID-19 Patients. *Clin Infect Dis*. April 2020. doi:10.1093/cid/ciaa460
- 358 8. Cleary B, Hay JA, Blumenstiel B, Gabriel S, Regev A, Mina MJ. Efficient prevalence estimation
- 359 and infected sample identification with group testing for SARS-CoV-2. *medRxiv*. 2020.
- 360 9. Carpenter B, Gelman A, Hoffman MD, et al. Stan : A Probabilistic Programming Language. *J Stat*
- 361 *Softw*. 2017;76(1). doi:10.18637/jss.v076.i01
- 362 10. Singanayagam A, Patel M, Charlett A, et al. Duration of infectiousness and correlation with RT-
- 363 PCR cycle threshold values in cases of COVID-19, England, January to May 2020. *Euro Surveill*.
- 364 2020;25(32):1-5. doi:10.2807/1560-7917.ES.2020.25.32.2001483
- 365 11. Cevik M, Marcus J, Buckee C, Smith T. SARS-CoV-2 Transmission Dynamics Should Inform
- 366 Policy. *SSRN Electron J*. 2020. doi:10.2139/ssrn.3692807
- 367 12. Cevik M, Tate M, Lloyd O, Maraolo AE, Schafers J, Ho A. SARS-CoV-2, SARS-CoV-1 and
- 368 MERS-CoV viral load dynamics, duration of viral shedding and infectiousness: a living systematic
- 369 review and meta-analysis. *medRxiv*. 2020.
- 370 13. Houlihan C, Vora N, Byrne T, et al. SARS-CoV-2 virus and antibodies in front-line Health Care
- 371 Workers in an acute hospital in London: preliminary results from a longitudinal study. *medRxiv*.
- 372 2020.
- 373 14. Larremore DB, Wilder B, Lester E, et al. Test sensitivity is secondary to frequency and turnaround
- 374 time for COVID-19 surveillance. *medRxiv*. 2020:2020.06.22.20136309.
- 375 doi:10.1101/2020.06.22.20136309
- 376 15. Paltiel AD, Zheng A, Walensky RP. Assessment of SARS-CoV-2 Screening Strategies to Permit
- 377 the Safe Reopening of College Campuses in the United States. *JAMA Netw Open*.
- 378 2020;3(7):e2016818. doi:10.1001/jamanetworkopen.2020.16818
- 379 16. U.S. Food and Drug Administration. *Quest Diagnostics Infectious Disease, Inc. ("Quest*
- 380 *Diagnostics") SARS-CoV-2 RNA Qualitative Real-Time RT-PCR Emergency Use Authorization*.
- 381 Washington, DC; 2020.
- 382 17. U.S. Food and Drug Administration. *Roche Molecular Systems, Inc. Cobas SARS-CoV-2*
- 383 *Emergency Use Authorization*. Washington, DC; 2020.
- 384 18. Kudo E, Israelow B, Vogels CBF, et al. Detection of SARS-CoV-2 RNA by multiplex RT-qPCR.
- 385 Sugden B, ed. *PLOS Biol*. 2020;18(10):e3000867. doi:10.1371/journal.pbio.3000867
- 386 19. Ott IM, Vogels C, Grubaugh N, Wyllie AL. *Saliva Collection and RNA Extraction for SARS-CoV-*
- 387 *2 Detection V.2*. New Haven, CT; 2020.
- 388 20. R Development Core Team R. R: A Language and Environment for Statistical Computing. Team
- 389 RDC, ed. *R Found Stat Comput*. 2011;1(2.11.1):409. doi:10.1007/978-3-540-74686-7
- 390 21. Vogels C, Fauver J, Ott IM, Grubaugh N. *Generation of SARS-COV-2 RNA Transcript Standards*
- 391 *for QRT-PCR Detection Assays*; 2020. doi:10.17504/protocols.io.bdv6i69e
- 392
- 393

394 **Supplemental Methods.**

395 Ethics.

396 Residual de-identified viral transport media from anterior nares and oropharyngeal swabs collected NBA
397 players, staff, and vendors were obtained from Quest Diagnostics or BioReference Laboratories. In
398 accordance with the guidelines of the Yale Human Investigations Committee, this work with de-identified
399 samples was approved for research not involving human subjects by the Yale Internal Review Board (HIC
400 protocol # 2000028599). This project was designated exempt by the Harvard IRB (IRB20-1407).

401

402 Additional testing protocol details.

403 Clinical samples were obtained by combined swabs of the anterior nares and oropharynx administered by
404 a trained provider. The samples were initially tested by either Quest Diagnostics (while teams were in local
405 markets using the Quest SARS-CoV-2 RT-qPCR¹⁶) or BioReference Laboratories (while teams were in
406 Orlando using the cobas SARS-CoV-2 test¹⁷). Viral transport media from positive samples were sent to
407 Yale University for subsequent RT-qPCR testing using a multiplexed version of the assay from the US
408 Centers for Disease Control and Prevention¹⁸ to normalize Ct values across testing platforms. A total of
409 234 samples from BioReference and 128 from Quest were tested at Yale; 49 positive samples had Ct values
410 assigned on first testing but did not undergo repeat testing at the Yale laboratory. To account for the
411 different calibration of the testing instruments, we used a linear conversion (**Supplemental Figures 12-14**,
412 **Methods: Converting Ct values**) to adjust these samples to the Yale laboratory scale. Subsequent analysis
413 is based on the N1 Ct value from the Yale multiplex assay and on the adjusted Roche cobas target 1 assay.

414

415 Residual viral transport media (VTM) from Quest Diagnostics or BioReference Laboratories were shipped
416 overnight to Yale on dry ice. VTM was thawed on ice and 300 µL was used for RNA extraction using the
417 MagMAX Viral/Pathogen Nucleic Acid Isolation Kit and the KingFisher Flex robot (Thermo Fisher
418 Scientific, Waltham, MA¹⁹). Total nucleic acid was eluted into 75ul of elution buffer and SARS-CoV-2
419 RNA was quantified from 5 µL of extracted total RNA using a multiplexed version of the CDC RT-qPCR
420 assay that contains the 2019-nCoV_N1 (N1), 2019-nCoV-N2 (N2), and human RNase P (RP) primer-probe
421 sets¹⁸. The RT-qPCR was performed using the Luna Universal Probe One-Step RT-qPCR Kit (New
422 England Biolabs, Ipswich, MA, US) and the following thermocycler conditions: (1) reverse transcription
423 for 10 minutes at 55°C, (2) initial denaturation for 1 min at 95°C, and PCR for 45 cycles of 10 seconds at
424 95°C and 30 seconds at 55°C on the CFX96 qPCR machine (Bio-Rad, Hercules, CA, US).

425

426 Converting Ct values. Most ($n = 226$) of the 312 positive samples in the raw dataset underwent RT-qPCR
427 at the Yale laboratory. We used the Yale Ct value whenever it was available. Still, 86 samples underwent
428 initial diagnostic testing at BioReference Laboratories but not confirmatory testing at the Yale laboratory.
429 Both platforms rely on a multiplex RT-qPCR strategy. The two testing platforms yield slightly different Ct
430 values, as evidenced by the 94 samples the underwent RT-qPCR at both facilities (**Supplemental Figure**
431 **12**). For comparison between platforms, target 1 from the Roche cobas assay, which is specific to SARS-
432 CoV-2, and the N1 target from the Yale multiplex assay were used. For the 86 samples that were not
433 processed at the Yale laboratory, we adjusted the Ct values using the best-fit (minimum sum of squares)
434 linear regression between the initial Ct value and the Yale Ct value for the samples that were processed in
435 both facilities. To do so, we estimated the coefficients β_0 and β_1 in the following regression equation:

436
$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

437

438 Here, y_i denotes the i^{th} Ct value from Yale, x_i denotes the i^{th} Ct value from the initial test, and ϵ_i is an error
439 term with mean 0 and constant variance across all samples. The resulting fit (**Supplemental Figure 12**)
440 was strong ($R^2 = 0.86$) with homoscedastic residuals (**Supplemental Figure 13**) that are approximately
441 normally distributed, as evidenced by a Q-Q plot (**Supplemental Figure 14**).

442

443 Data parsing. The raw dataset included 3,207 test results for 102 individuals. We excluded 21 individuals
444 who had 5 or fewer tests, since the data for these individuals were too sparse to reliably infer a Ct trajectory.
445 We also excluded 13 individuals who did not record any Ct values that surpassed the RT-qPCR limit of
446 detection (40). We removed 146 entries for which the test result was recorded as ‘positive’ but there was
447 no associated Ct value; these tests were initially conducted on an instrument that provided only a binary
448 diagnosis and the samples were not available for confirmatory testing. This left 2,411 total tests for 68
449 individuals for the main analysis. The median number of tests administered to each of the 68 individuals
450 was 41 (IQR [14, 51]; Range [5, 70]). The median number of Ct values with viral concentration above the
451 limit of detection recorded for each person was 3 (IQR [2, 4]; Range [1, 9]). We trivially shifted the date
452 indices so that date 0 corresponded to the time of the minimum Ct. We set the Ct value for negative tests
453 equal to the limit of detection. For the statistical analysis, we removed any sequences of 3 or more
454 consecutive negative tests to avoid overfitting to these trivial values.
455

456 Model fitting.

457 We assumed that the viral concentration trajectories consisted of a proliferation phase, with exponential
458 growth in viral RNA concentration, followed by a clearance phase characterized by exponential decay in
459 viral RNA concentration⁸. Since Ct values are roughly proportional to the negative logarithm of viral
460 concentration³, this corresponds to a linear decrease in Ct followed by a linear increase. We therefore
461 constructed a piecewise-linear regression model to estimate the peak Ct value, the time from infection onset
462 to peak (*i.e.* the duration of the proliferation stage), and the time from peak to infection resolution (*i.e.* the
463 duration of the clearance stage). This idealized trajectory is depicted in **Supplemental Figure 15**. The
464 trajectory may be represented by the equation
465

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

466
467 Here, $E[Ct(t)]$ represents the expected value of the Ct at time t , “l.o.d” represents the RT-qPCR limit of
468 detection, δ is the absolute difference in Ct between the limit of detection and the peak (lowest) Ct, and t_o ,
469 t_p , and t_r are the onset, peak, and recovery times, respectively.
470
471

472 Before fitting, we re-parametrized the model using the following definitions:
473

- 474 • $\Delta Ct(t) = \text{l.o.d.} - Ct(t)$ is the difference between the limit of detection and the observed Ct value at
475 time t .
- 476 • $\omega_p = t_p - t_o$ is the duration of the proliferation stage.
- 477 • $\omega_c = t_r - t_p$ is the duration of the clearance stage.
- 478

479 We constrained $0 \leq \omega_p \leq 14$ days and $0 \leq \omega_p \leq 30$ days to prevent inferring unrealistically large values for
480 these parameters for trajectories that were missing data prior to the peak and after the peak, respectively.
481 We also constrained $0 \leq \delta \leq 40$ as Ct values can only take values between 0 and the limit of detection (40).

482
483 We next assumed that the observed $\Delta Ct(t)$ could be described the following mixture model:
484

$$\Delta Ct(t) \sim \lambda \text{Normal}(E[\Delta Ct(t)], \sigma(t)) + (1 - \lambda) \text{Exponential}(\log(10)) \Big|_0^{\text{l.o.d.}}$$

487 where $E[\Delta Ct(t)] = \text{l.o.d.} - E[Ct(t)]$ and λ is the sensitivity of the q-PCR test, which we fixed at 0.99. The
488 bracket term on the right-hand side of the equation denotes that the distribution was truncated to ensure Ct
489 values between 0 and the limit of detection. This model captures the scenario where most observed Ct
490 values are normally distributed around the expected trajectory with standard deviation $\sigma(t)$, yet there is a
491 small (1%) probability of an exponentially-distributed false negative near the limit of detection. The log(10)
492 rate of the exponential distribution was chosen so that 90% of the mass of the distribution sat below 1 Ct
493 unit and 99% of the distribution sat below 2 Ct units, ensuring that the distribution captures values
494 distributed at or near the limit of detection. We did not estimate values for λ or the exponential rate because
495 they were not of interest in this study; we simply needed to include them to account for some small
496 probability mass that persisted near the limit of detection to allow for the possibility of false negatives.
497

498 For the 86 samples that were not tested in the Yale laboratory, we included additional uncertainty in the
499 observed Ct value by inflating $\sigma(t)$, such that
500

$$\sigma(t) = (\tilde{\sigma}^2 + \epsilon^2 I_{adj})^{1/2}$$

501 Here, $\sigma(\tilde{\sigma})$ is a constant, ϵ is the standard deviation of the residuals from the linear fit between the initial
502 test and the Yale laboratory test, and I_{adj} is an indicator variable that is 1 if the sample at time t was adjusted
503 and 0 otherwise.
504

505 We used a hierarchical structure to describe the distributions of ω_p , ω_r , and δ for each individual based on
506 their respective population means $\mu_{\omega p}$, $\mu_{\omega r}$, and μ_δ and population standard deviations $\sigma_{\omega p}$, $\sigma_{\omega r}$, and σ_δ such
507 that
508

$$\begin{aligned} 511 \quad \omega_p &\sim \text{Normal}(\mu_{\omega p}, \sigma_{\omega p}) \\ 512 \quad \omega_r &\sim \text{Normal}(\mu_{\omega r}, \sigma_{\omega r}) \\ 513 \quad \delta &\sim \text{Normal}(\mu_\delta, \sigma_\delta) \end{aligned}$$

514
515 We inferred separate population means (μ) for symptomatic and asymptomatic individuals. We used a
516 Hamiltonian Monte Carlo fitting procedure implemented in Stan (version 2.24)⁹ and R (version 3.6.2)²⁰ to
517 estimate the individual-level parameters ω_p , ω_r , δ , and t_p as well as the population-level parameters $\sigma(\tilde{\sigma})$,
518 $\mu_{\omega p}$, $\mu_{\omega r}$, μ_δ , $\sigma_{\omega p}$, $\sigma_{\omega r}$, and σ_δ . We used the following priors:
519

520 *Hyperparameters:*

$$\begin{aligned} 522 \quad \sigma(\tilde{\sigma}) &\sim \text{Cauchy}(0, 5) [0, \infty] \\ 523 \\ 524 \quad \mu_{\omega p} &\sim \text{Normal}(14/2, 14/6) [0, 14] \\ 525 \quad \mu_{\omega r} &\sim \text{Normal}(30/2, 30/6) [0, 30] \\ 526 \quad \mu_\delta &\sim \text{Normal}(40/2, 40/6) [0, 40] \\ 527 \\ 528 \quad \sigma_{\omega p} &\sim \text{Cauchy}(0, 14/\tan(\pi(0.95-0.5))) [0, \infty] \\ 529 \quad \sigma_{\omega r} &\sim \text{Cauchy}(0, 30/\tan(\pi(0.95-0.5))) [0, \infty] \\ 530 \quad \sigma_\delta &\sim \text{Cauchy}(0, 40/\tan(\pi(0.95-0.5))) [0, \infty] \end{aligned}$$

531
532 *Individual-level parameters:*

$$\begin{aligned} 533 \quad \omega_p &\sim \text{Normal}(\mu_{\omega p}, \sigma_{\omega p}) [0, 14] \\ 534 \quad \omega_r &\sim \text{Normal}(\mu_{\omega r}, \sigma_{\omega r}) [0, 30] \\ 535 \quad \delta &\sim \text{Normal}(\mu_\delta, \sigma_\delta) [0, 40] \end{aligned}$$

536 $t_p \sim \text{Normal}(0, 2)$

537

538 The values in square brackets denote truncation bounds for the distributions. We chose a vague half-Cauchy
539 prior with scale 5 for the observation variance $\sigma(\tilde{\sigma})$. The priors for the population mean values (μ_\cdot) are
540 normally-distributed priors spanning the range of allowable values for that parameter; this prior is vague
541 but expresses a mild preference for values near the center of the allowable range. The priors for the
542 population standard deviations (σ_\cdot) are half Cauchy-distributed with scale chosen so that 90% of the
543 distribution sits below the maximum value for that parameter; this prior is vague but expresses a mild
544 preference for standard deviations close to 0.

545

546 We ran four MCMC chains for 5,000 iterations each with a target average proposal acceptance probability
547 of 0.99. The first half of each chain was discarded as the warm-up. The Gelman R-hat statistic was less than
548 1.1 for all parameters except for the t_p and ω_r associated with person ID 1370, as the posterior distributions
549 for those parameters were multi-modal (see **Supplemental Figures 8-9**). This indicates good overall
550 mixing of the chains. There were fewer than 10 divergent iterations (<0.1% of the transitions after warm-
551 up), indicating good exploration of the parameter space.

552

553 The posterior distributions for μ_δ , μ_{ω_p} , and μ_{ω_r} , estimated separately for symptomatic and asymptomatic
554 individuals, are reported in **Figure 2** (main text). We fit a second model that did not distinguish between
555 symptomatic and asymptomatic individuals. The posterior distributions for these same parameters under
556 this model are depicted in **Supplemental Figure 5**. The posterior distributions for the individual-level
557 parameters ω_p , ω , and δ are depicted in **Supplemental Figures 6-8**, with 500 sampled trajectories from
558 these posterior distributions for each individual depicted in **Supplemental Figure 9**. The overall combined
559 posterior distributions for the individual-level parameters ω_p , ω_r , and δ are depicted in **Supplemental**
560 **Figure 10**. We estimated the best-fit normal (for δ) and gamma (for ω_p and ω_r) distributions using the
561 ‘fitdistrplus’ package implemented in R (version 3.6.2)²⁰.

562

563 Converting Ct values to viral genome equivalents.

564 CT values were fitted to a standard curve in order to convert Ct value data to RNA copies or genome
565 equivalents (GE). Synthetic T7 RNA transcripts corresponding to a 1,363 b.p. segment of the SARS-CoV-
566 2 nucleocapsid gene were serially diluted from 10^6 - 10^0 GE/ μl in duplicate to generate a standard curve²¹
(Supplemental Table 1). The average Ct value for each dilution was used to calculate the slope (-3.60971)
567 and intercept (40.93733) of the linear regression of Ct on log-10 transformed standard RNA concentration,
568 and Ct values from subsequent RT-qPCR runs were converted to GE using the following equation:
569

570

$$571 \log_{10}([\text{RNA}]) = (Ct - 40.93733)/(-3.60971) + \log_{10}(250)$$

572

573 Here, [RNA] represents the RNA concentration in GE/ml. The $\log_{10}(250)$ term accounts for the extraction
574 (300 μl) and elution (75 μl) volumes associated with processing the clinical samples as well as the 1,000
575 $\mu\text{l}/\text{ml}$ unit conversion.

576

577 Inferring the stage of infection using single and paired Ct values

578

579 To determine whether individual or paired Ct values can reveal a patient’s stage of infection, we measured
580 the frequency with which a Ct value falling within a 5-unit band, possibly followed by a second Ct value
581 of higher or lower magnitude, was associated with the proliferation stage, the clearance stage, or the
582 persistent stage. First, we assigned to each positive test the probability that it was collected during each of
583 the three stages of infection. To do so, we began with the positive samples from the 46 individuals with
584 acute infections and calculated the frequency with which each sample sat within the proliferation stage, the

585 clearance stage, or the persistent stage (*i.e.*, neither of the previous two stages) across 10,000 posterior
586 parameter draws for that person. For the remaining 22 individuals, all positive samples were assigned to
587 the persistent stage. Next, we calculated the probability that a Ct value falling within a 5-unit window
588 corresponded to an active infection (*i.e.*, either the proliferation or the clearance stage) by summing the
589 proliferation and clearance probabilities for all positive samples with that window and dividing by the total
590 number of positive samples in the window. We considered windows with midpoints spanning from Ct =
591 37.5 to Ct = 15.5 (**Figure 3A**). We performed a similar calculation to determine the probability that a Ct
592 falling within a given 5-unit window corresponded to just the proliferation phase, assuming it had already
593 been determined that the sample fell within an active infection (**Figure 3B**). Finally, to assess the
594 information gained by conducting a second test within two days of an initial positive, we collected all
595 positive samples with a subsequent sample (positive or negative) that was taken within two days and
596 repeated the above calculations, separating by whether the second test had a higher or lower Ct than the
597 first.

598

599 Calculating effective sensitivity.

600 The sensitivity of a test is defined as the probability that the test correctly identifies an individual who is
601 positive for some criterion of interest. For clinical SARS-CoV-2 tests, the criterion of interest is current
602 infection with SARS-CoV-2. However, for the time-to-event analysis presented in the main text, the
603 criterion of interest is infectiousness at some point in the future. The effective sensitivity of a test (with
604 respect to future infectiousness) may differ substantially from its clinical sensitivity (with respect to current
605 infection).

606

607 The effective sensitivity of a test intended to detect future infectiousness depends on the test's
608 characteristics (its limit of detection and sampling error rate) as well as the viral dynamics of infected
609 individuals. To determine the effective sensitivity of a test n hours before an event, we first sampled 2,500
610 posterior draws for the proliferation time, clearance time, and peak Ct value from the MCMC fits to the
611 viral trajectory data. We included only draws with a peak Ct ≤ 30 , as we assumed Ct = 30 to be the threshold
612 of infectiousness (individuals with a peak Ct > 30 would never be infectious according to this threshold and
613 therefore would never satisfy the criterion of interest, *i.e.*, infectiousness at the event). We trivially defined
614 the event's start time to be $t = 0$ and assumed that the event lasted for 3 hours. For each of the 2,500
615 individuals, we identified the range of possible onset times for the proliferation stage that would ensure the
616 person would be infectious (Ct < 30) at some point during the event. We uniform-randomly drew a
617 proliferation onset time from this range for each individual. We then simulated a test at time $-n$. Any
618 individuals with Ct greater than the test's limit of detection at time $-n$ went undetected. Any individuals
619 with Ct less than the test's limit of detection at time $-n$ were detected with probability (1-sampling error).
620 The effective sensitivity was calculated as the number of individuals successfully detected divided by the
621 total number of individuals who could have been detected (2,500). We repeated this calculation for 1-hour
622 increments through 3 days prior to the event. We considered two tests, one with a limit of detection at 40
623 Ct and a sampling error of 1% (analogous to RT-qPCR) and one with a limit of detection at 35 Ct and a
624 sampling error of 5% (analogous to a rapid antigen test).

625

626 To calculate the expected number of individuals who might attend an event while infectious given a test n
627 hours before the event, we simulated 1,000 events. For each event, we drew the number of individuals who
628 would have attended the event while infectious in the absence of testing, η , from a Binomial(N, p)
629 distribution where N is the number of event attendees (1,000 for the examples presented in the main text)
630 and p is the prevalence of infectious individuals in the population (2% for the examples in the main text).
631 We then calculated the number of individuals who were successfully identified using a Binomial(η, eff_se)
632 distribution where η is the number of individuals who would have been infectious at the event in the absence
633 of testing and eff_se is the effective sensitivity of the test when administered at time $-n$. The remaining

634 individuals were missed by the test. This value (its mean and 90% quantiles across the 1,000 simulated
635 events) is reported in **Figure 4b** (main text).

636

637 The online calculator differs slightly from the above procedure. Rather than drawing directly from the
638 posterior distributions for the MCMC fits, the calculator allows the user to input different values specifying
639 the population distribution of proliferation times, clearance times, and peak Ct values. The proliferation
640 times and clearance times are described by independent Gamma distributions with user-input mean and
641 standard deviation. The peak Ct values are defined by independent normal distributions with user-input
642 mean and standard deviation, truncated to ensure that the values lie between 0 and 40 Ct. The default values
643 align with the best-fit values for the respective Gamma and normal distributions reported in the caption of
644 **Supplemental Figure 10**. The effective sensitivity values and expected number of infectious attendees
645 therefore differ slightly from the values reported in the main text due to these distributional approximations
646 and the fact that the values are drawn independently (in the posterior draws, there is some correlation
647 between the parameters). Still, the calculations align closely and allow for greater flexibility in allowing
648 the user to update the viral trajectory parameters to reflect different populations.

649 **Acknowledgements.** We thank the NBA, National Basketball Players Association (NBPA), and all of the
650 study participants who are committed to applying what they learned from sports towards enhancing public
651 health. In particular, we thank D. Weiss of the NBA for his continuous support and leadership. We are
652 appreciative of the discussions from the COVID-19 Sports and Society Working Group. We also thank D.
653 Larremore for comments on the manuscript, J. Hay and R. Niehus for suggestions on the statistical approach
654 and P. Jack and S. Taylor for laboratory support.

655

656 **Funding.** This study was funded by the NWO Rubicon 019.181EN.004 (CBFV), a clinical research
657 agreement with the NBA and NBPA (NDG), the Huffman Family Donor Advised Fund (NDG), Fast Grant
658 funding support from the Emergent Ventures at the Mercatus Center, George Mason University (NDG),
659 and the Morris-Singer Fund for the Center for Communicable Disease Dynamics at the Harvard T.H. Chan
660 School of Public Health (YHG).

661

662 **Role of funding source.** The funding sources did not play a role in the data collection, analysis, or
663 interpretation of this study.

664

665 **Author contributions.** SMK conceived of the study, conducted the statistical analysis, and wrote the
666 manuscript. JRF conceived of the study, conducted the laboratory analysis, and wrote the manuscript. CM
667 conceived of the study, collected the data, and wrote the manuscript. SWO conducted the statistical analysis.
668 CT analyzed the data and edited the manuscript. KYS analyzed the data and edited the manuscript. CCK
669 conducted the laboratory analysis and edited the manuscript. SJ conducted the laboratory analysis and
670 edited the manuscript. IMO conducted the laboratory analysis. CBFV conducted the laboratory analysis.
671 JW conducted laboratory analysis and edited the manuscript. JW conducted laboratory analysis and edited
672 the manuscript. JD conceived of the study and edited the manuscript. DJA contributed to data analysis and
673 edited the manuscript. JM contributed to data analysis and edited the manuscript. DDH conceived of the
674 study and edited the manuscript. NDG conceived of the study, oversaw the study, and wrote the manuscript.
675 YHG conceived of the study, oversaw the study, and wrote the manuscript.

676

677 **Competing interests.**

678 JW is an employee of Quest Diagnostics. JW is an employee of Bioreference Laboratories. NDG receives
679 financial support from Tempus to develop SARS-CoV-2 diagnostic tests. SMK, SWO, and YHG have a
680 consulting agreement with the NBA.

681

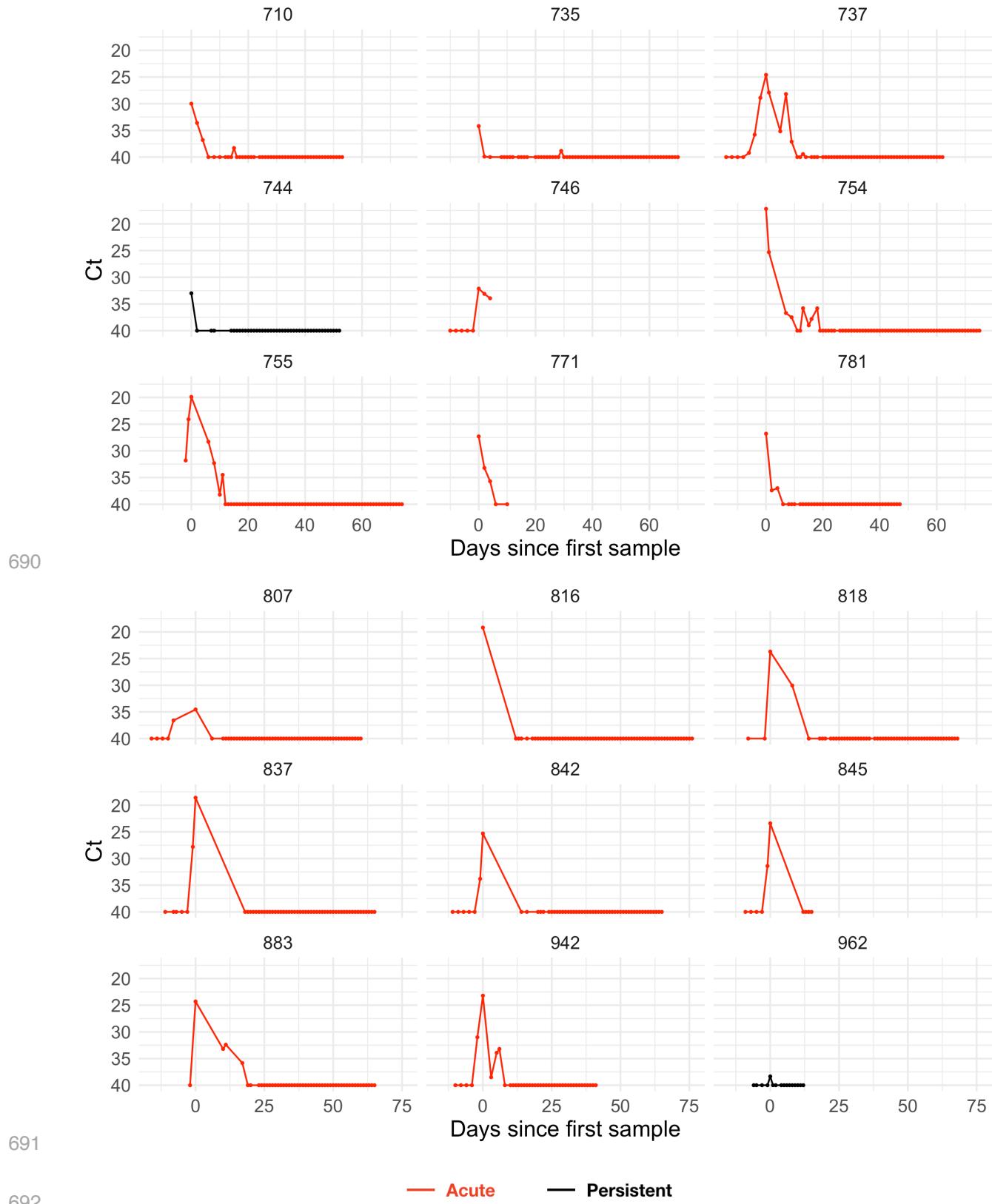
682 **Supplementary Information** is available for this paper.

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

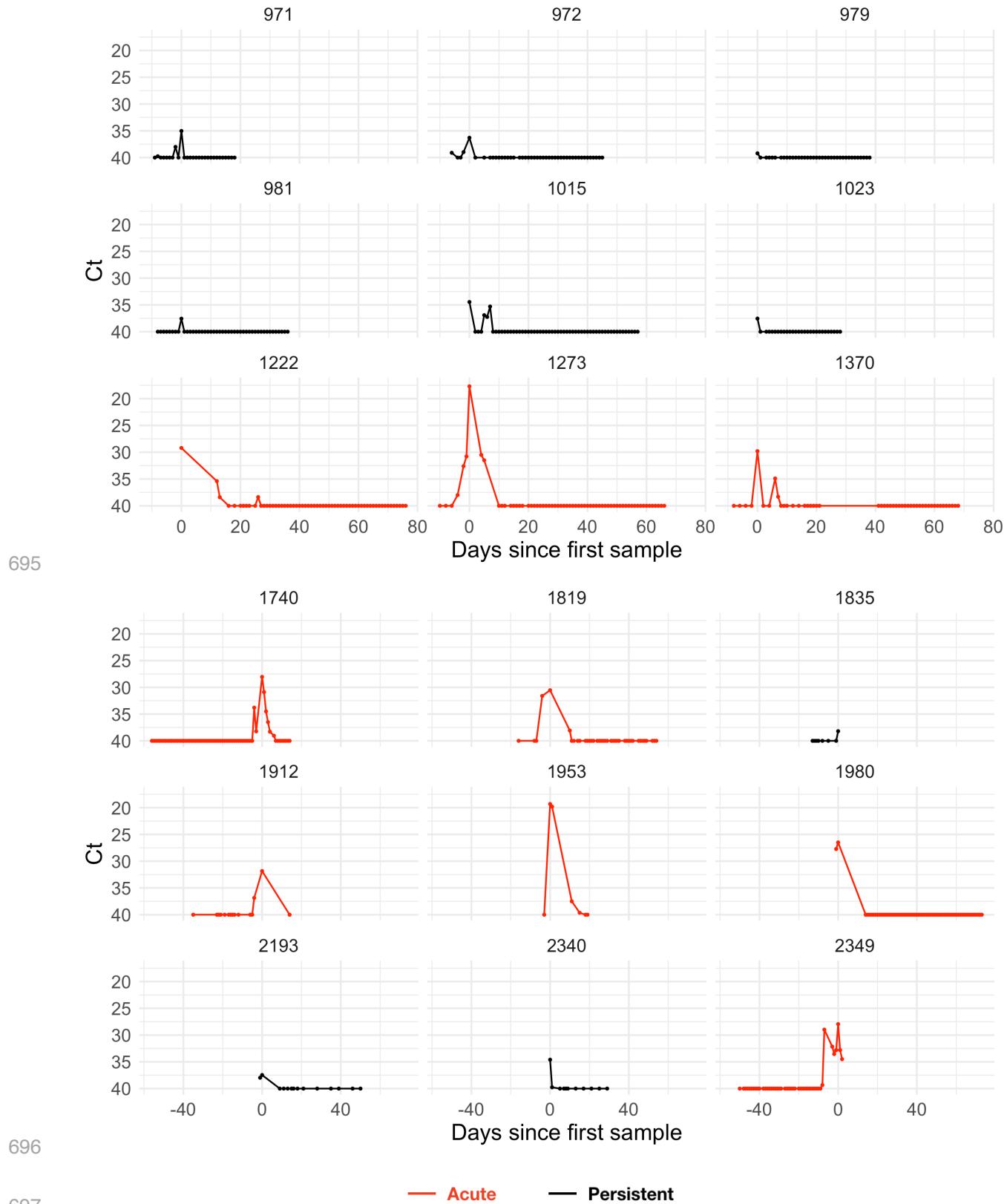
683

684 **Supplemental Table 1. Standard curve relationship between genome equivalents and Ct values.**

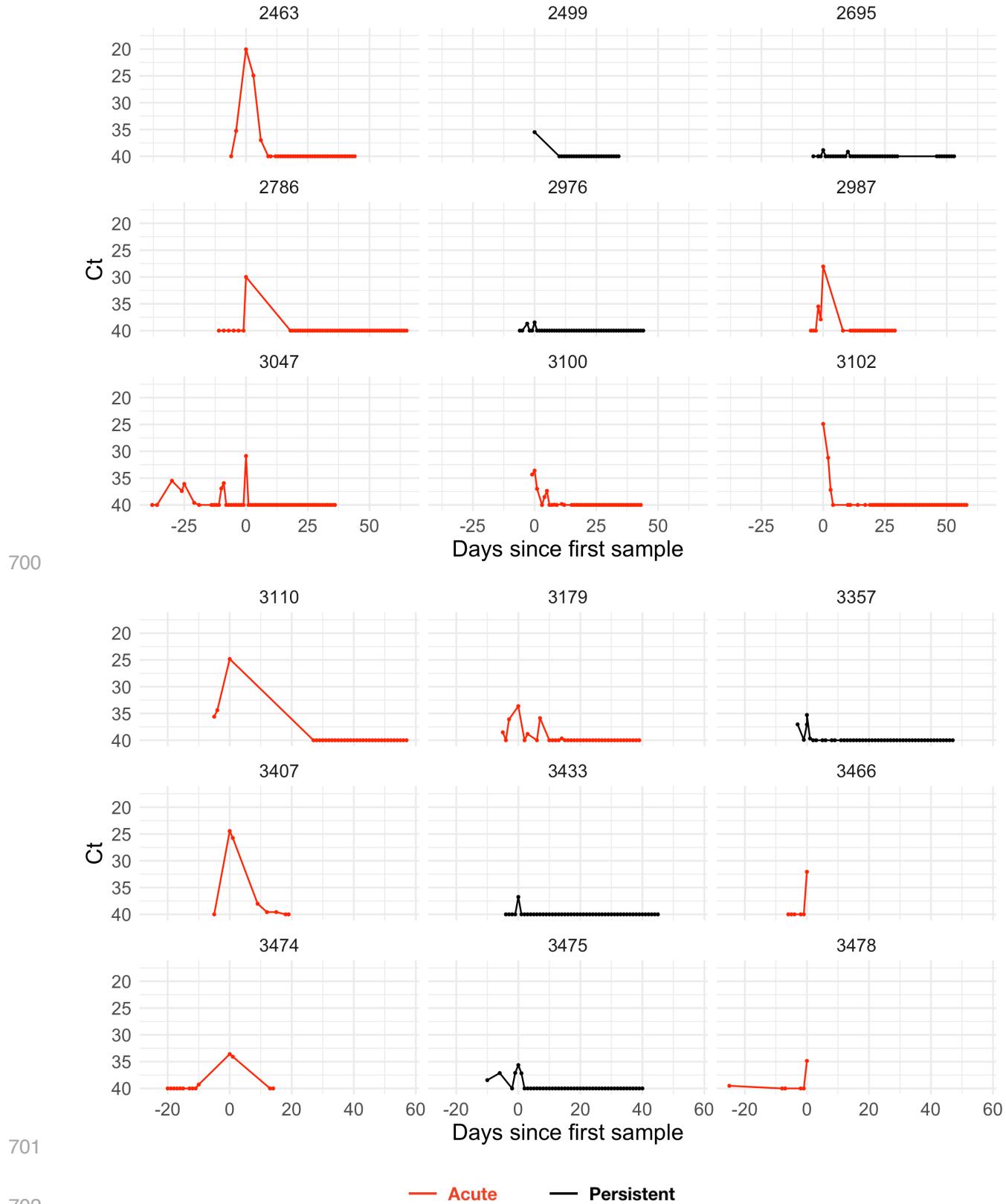
685 Synthetic T7 RNA transcripts corresponding to a 1,363 base pair segment of the SARS-CoV-2 nucleocapsid
686 gene were serially diluted from 10⁶-10⁰ and evaluated in duplicate with RT-qPCR. The best-fit linear
687 regression of the average Ct on the log10-transformed standard values had slope -3.60971 and intercept
688 40.93733 ($R^2 = 0.99$).
689



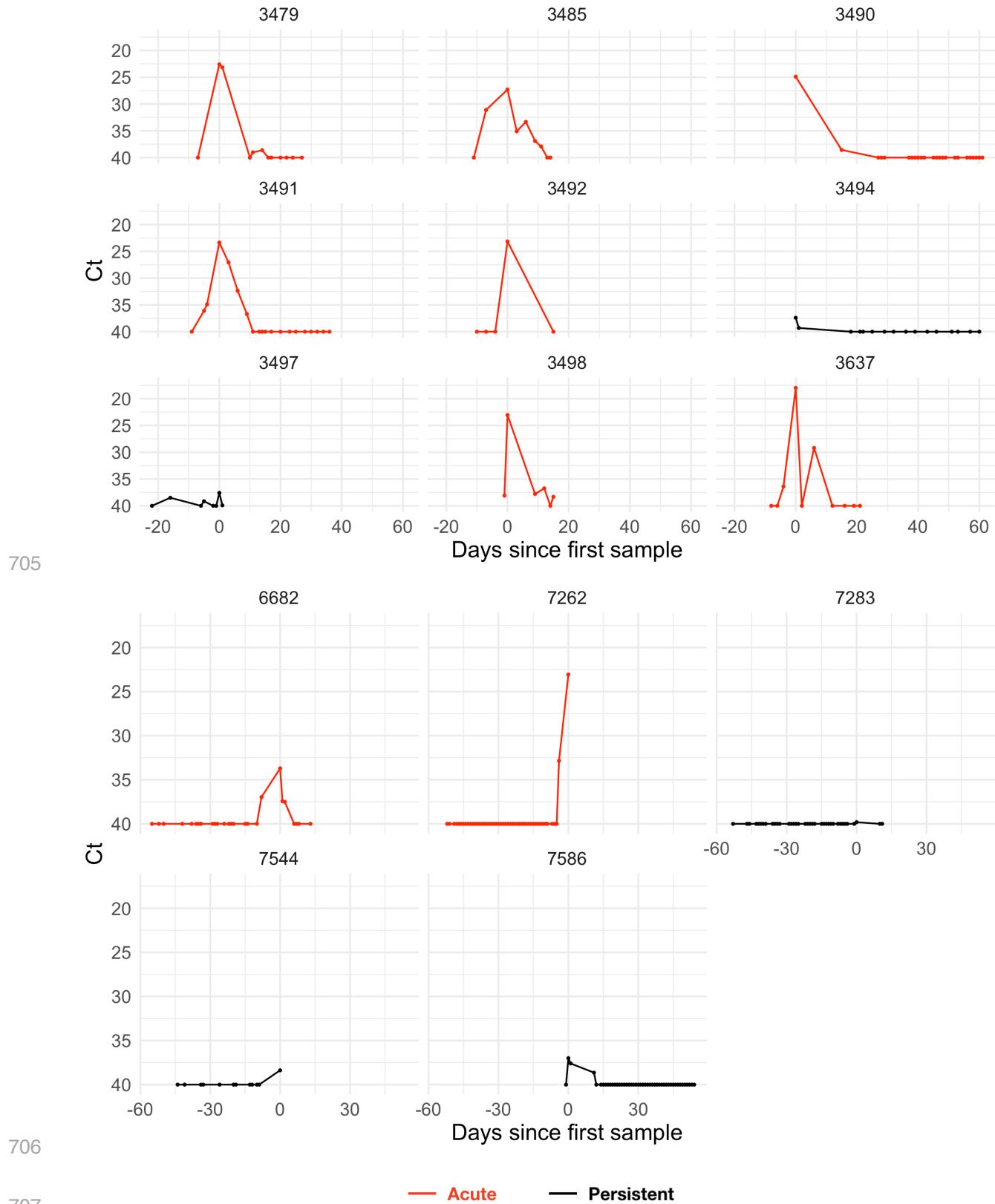
Supplemental Figure 1. Observed Ct values from the study participants (1/4). Points depict observed Ct values, which are connected with lines to better visualize trends. Individuals with presumed acute infections are marked in red. All others are in black.



698 **Supplemental Figure 2. Observed Ct values from the study participants (2/4).** Points depict observed Ct values, which are
699 connected with lines to better visualize trends. Individuals with presumed acute infections are marked in red. All others are in black.



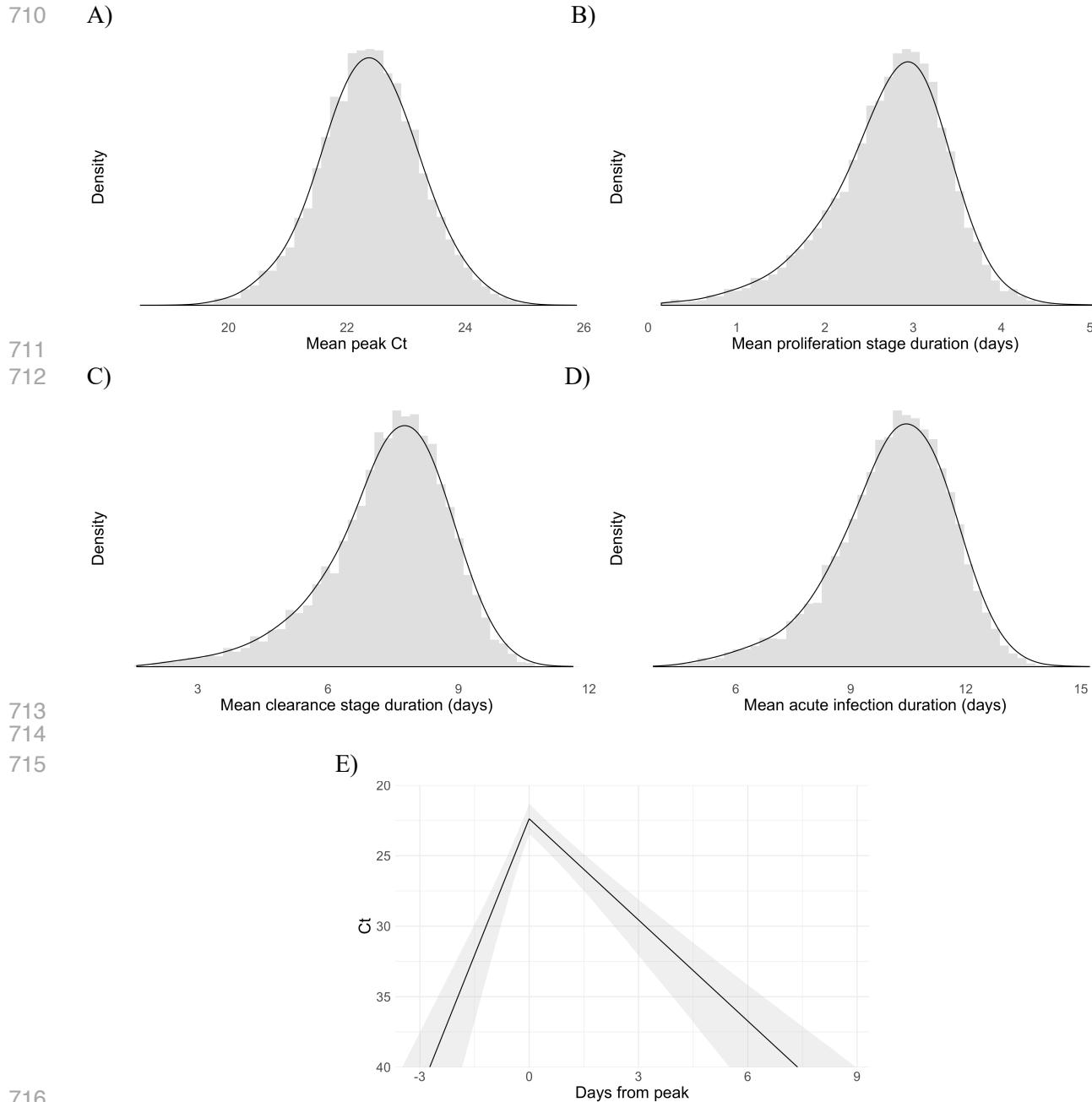
Supplemental Figure 3. Observed Ct values from the study participants (3/4). Points depict observed Ct values, which are connected with lines to better visualize trends. Individuals with presumed acute infections are marked in red. All others are in black.



— Acute — Persistent

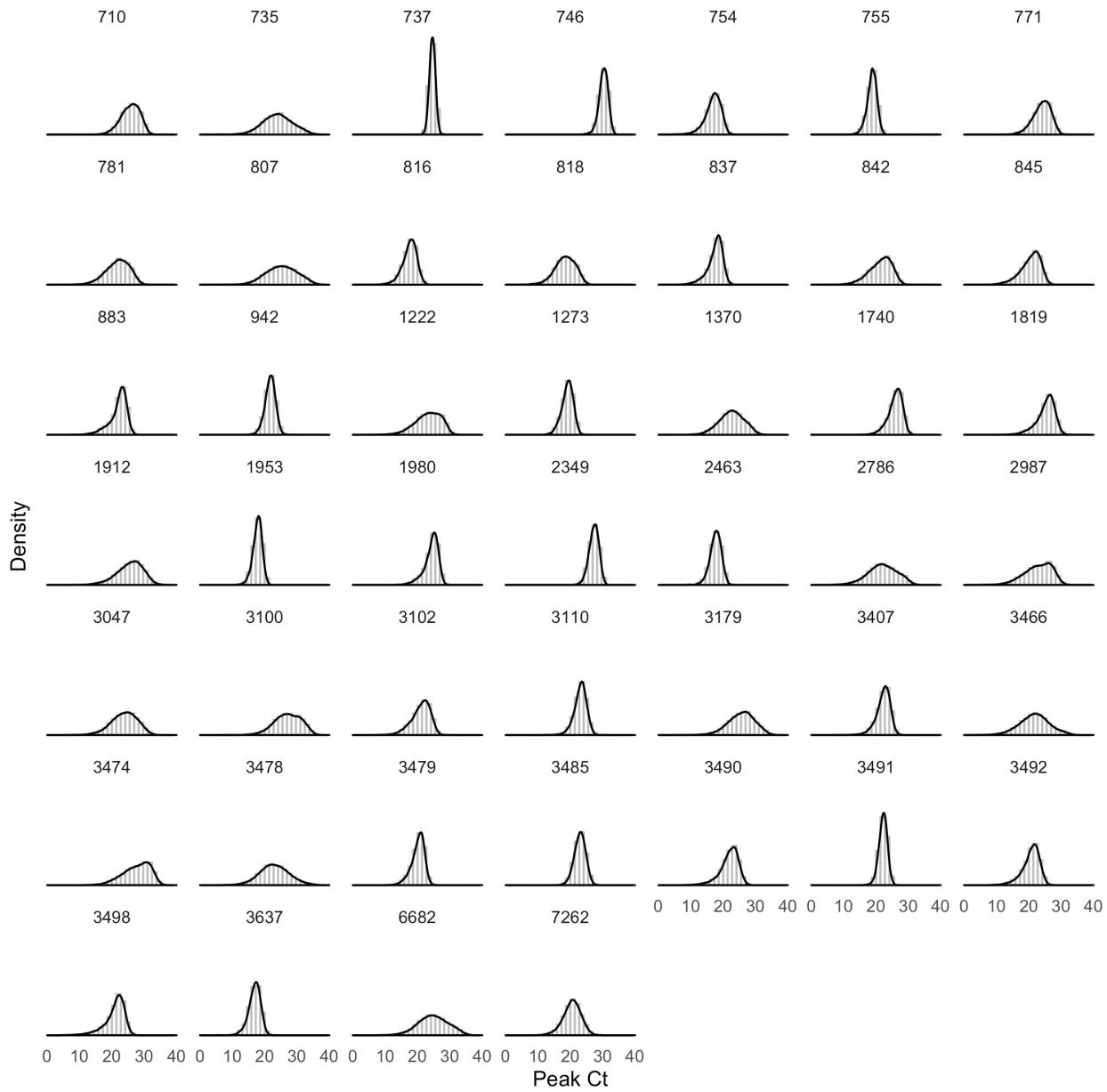
707

708 **Supplemental Figure 4. Observed Ct values from the study participants (4/4).** Points depict observed Ct values, which are
709 connected with lines to better visualize trends. Individuals with presumed acute infections are marked in red. All others are in black.



716
717
718 **Supplemental Figure 5. Mean peak Ct value and waiting time distributions for individuals with acute infections.** Histograms
719 (colored bars) of 10,000 posterior draws from the distributions for peak Ct value (A), duration of the proliferation stage (infection
720 detection to peak Ct, B), duration of the clearance stage (peak Ct to resolution of acute RNA shedding, C), and total duration of
721 acute shedding (D) across the 46 individuals with a verified infection. The curves are kernel density estimators for the histograms
722 to assist with visualizing the shapes of the histograms. The mean Ct trajectory corresponding to the mean values for peak Ct,
723 proliferation duration, and clearance duration is depicted in (E) (solid lines), where shading depicts the 90% credible interval.

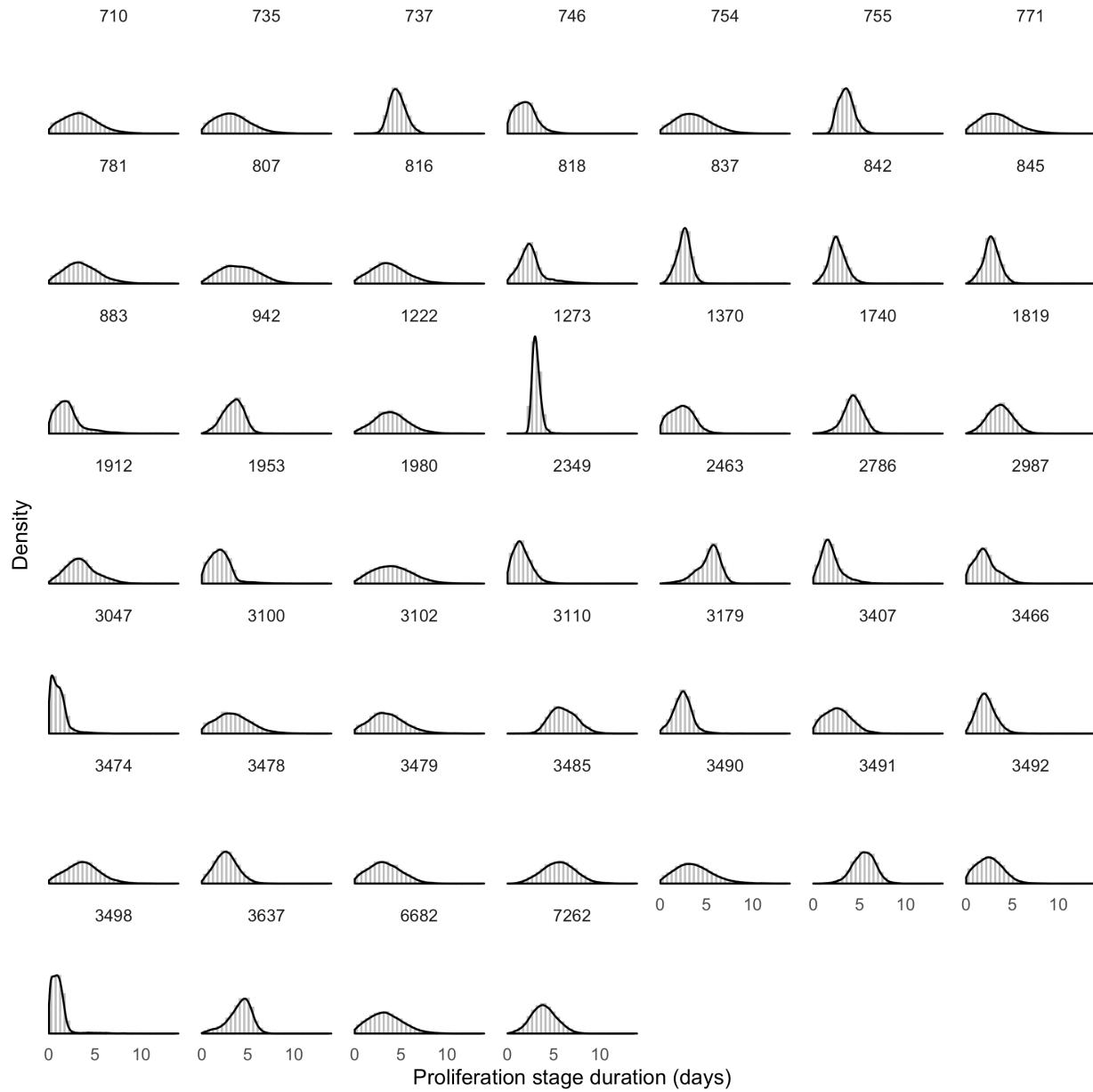
724



725

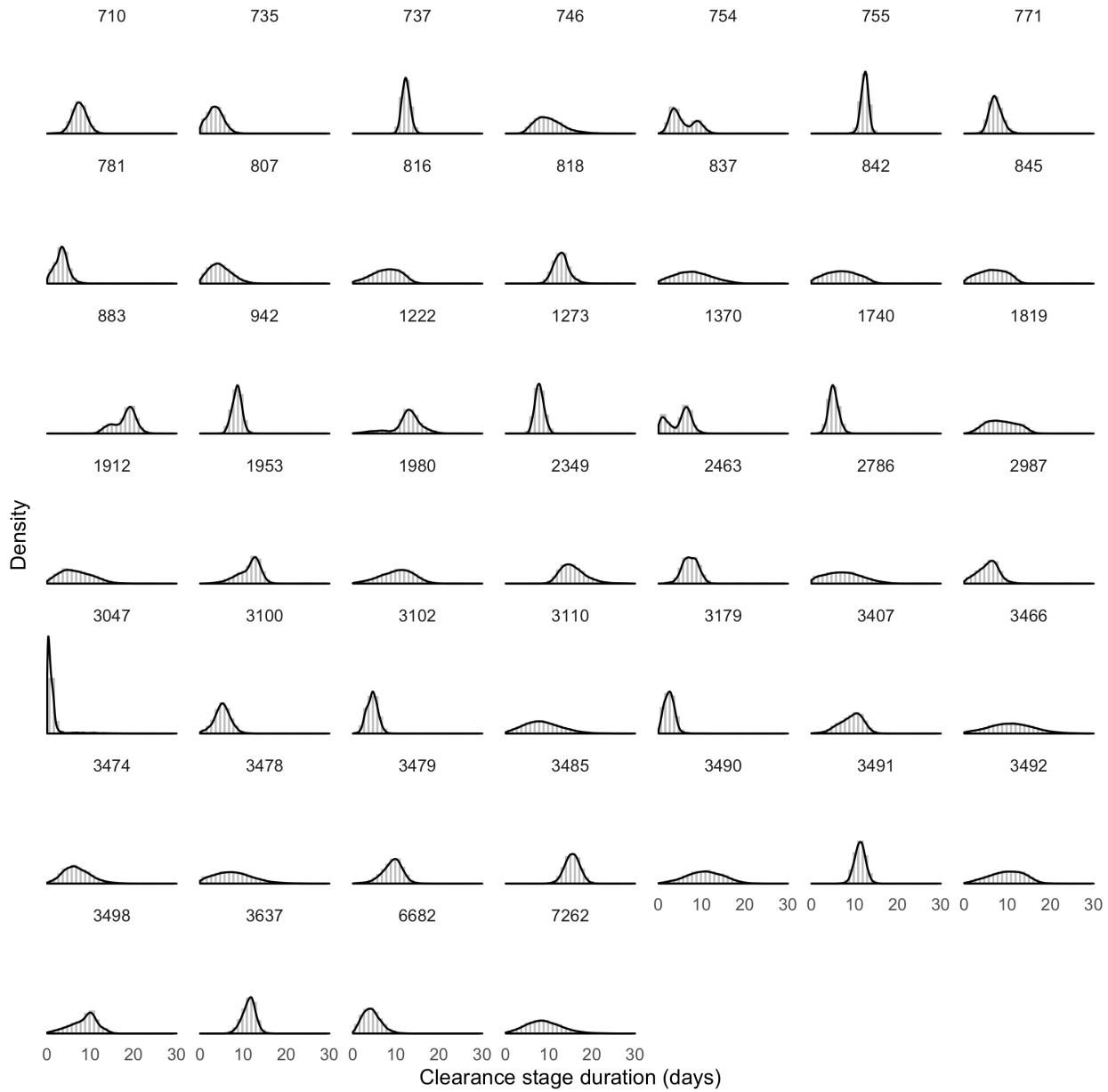
726

Supplemental Figure 6. Posterior peak Ct value distributions for the 46 individuals with acute infections.



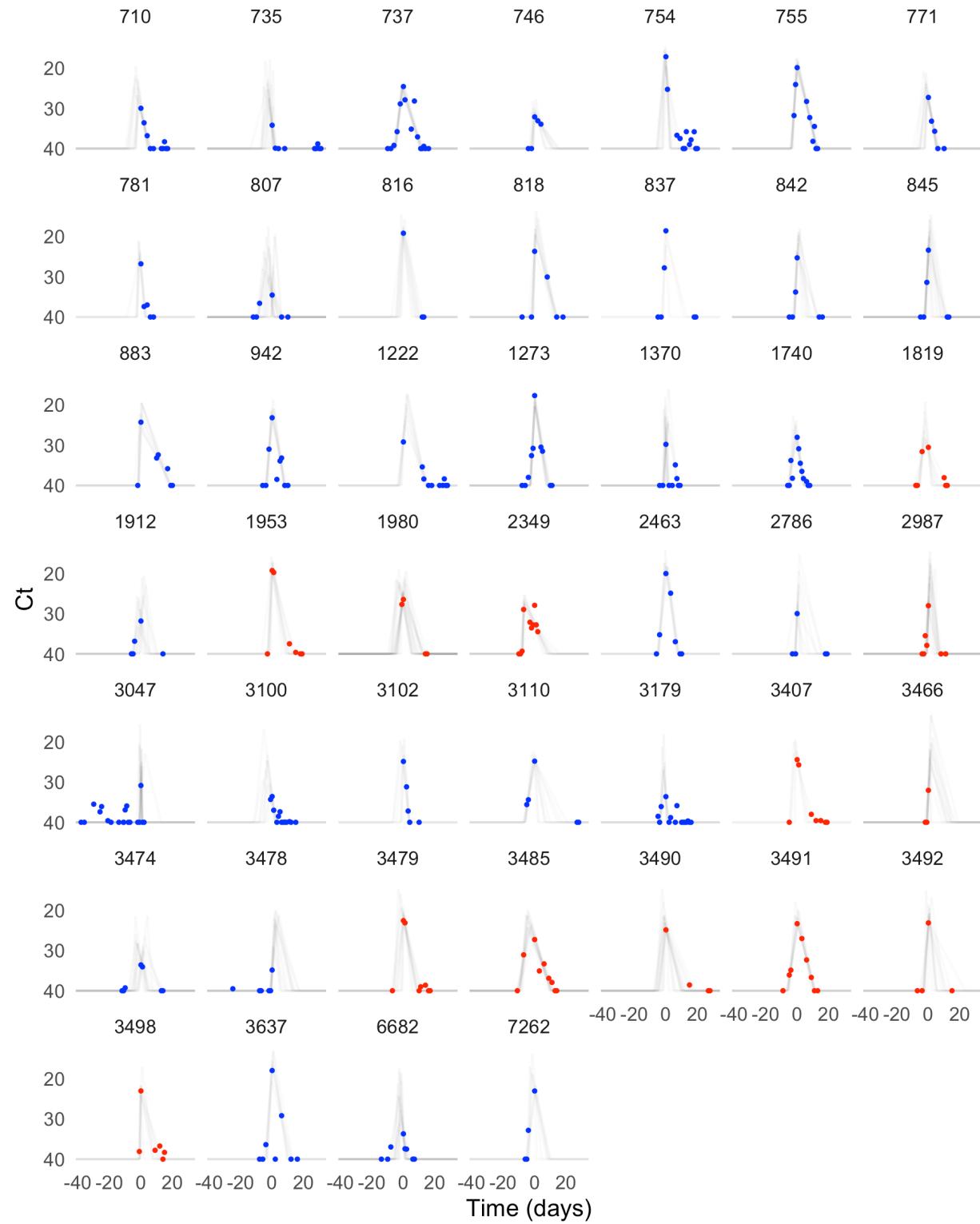
727
728
729

Supplemental Figure 7. Posterior distributions for the duration of the proliferation stage for 46 individuals with acute infections.



730
731

Supplemental Figure 8. Posterior distributions for the clearance stage duration for 46 individuals with acute infections.



732

● Symptoms reported ● Symptoms not reported

733

734

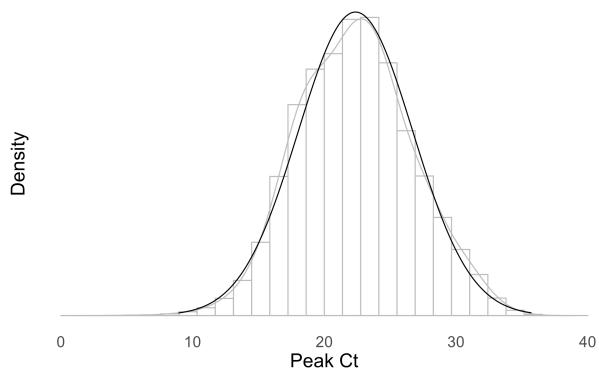
735

736

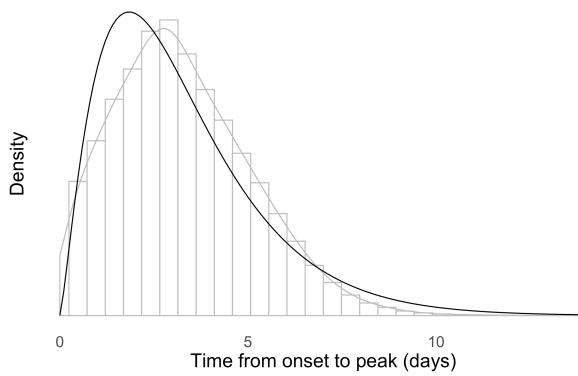
Supplemental Figure 9. Best-fit Ct trajectories for the 46 individuals with acute infections. Thin grey lines depict 500 sampled trajectories. Points represent the observed data, with symptomatic individuals represented in red and asymptomatic individuals in blue.

737

738 A)

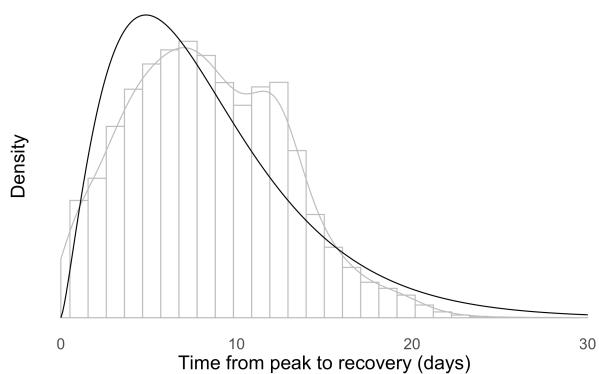


B)

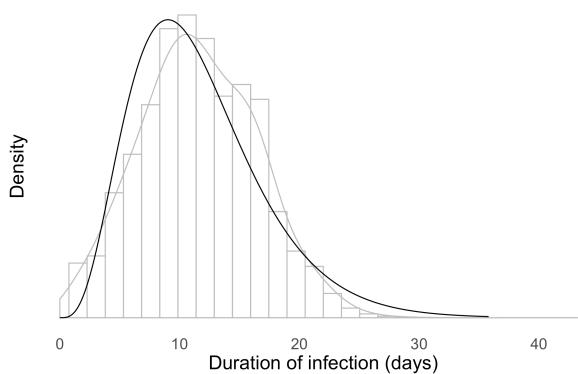


739

740 C)



D)



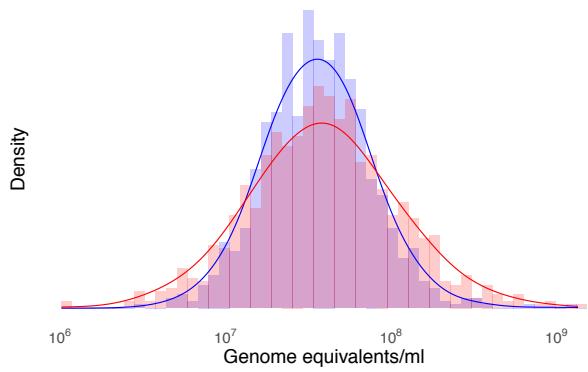
741

742

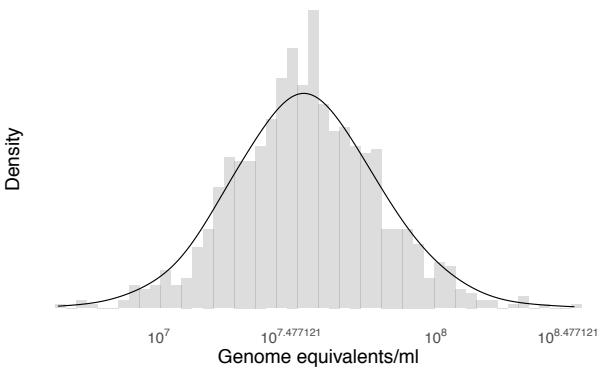
743 **Supplemental Figure 10. Individual-level peak Ct value and waiting time distributions.** Histograms (grey bars) of 10,000
744 posterior draws from the distributions for peak Ct value (A), time from onset to peak (B), time from peak to recovery (C), and total
745 duration of infection (D) across the 46 individuals with an acute infection. Grey curves are kernel density estimators to more clearly
746 exhibit the shape of the histograms. Black curves represent the best-fit normal (A) or gamma (B, C, D) distributions to the
747 histograms. The duration of infection is the sum of the time from onset to peak and the time from peak to recovery. The best-fit
748 normal distribution to the posterior peak Ct value distribution had mean 22.3 and standard deviation 4.2; the best-fit gamma
749 distribution to the proliferation stage duration had shape parameter 2.3 and inverse scale parameter 0.7; the best-fit gamma
750 distribution to the clearance stage duration had shape parameter 2.4 and inverse scale parameter 0.3; and the best-fit gamma
751 distribution to the total duration of infection had shape parameter 4.3 and inverse scale parameter 0.4.
752

753

754 A)



B)

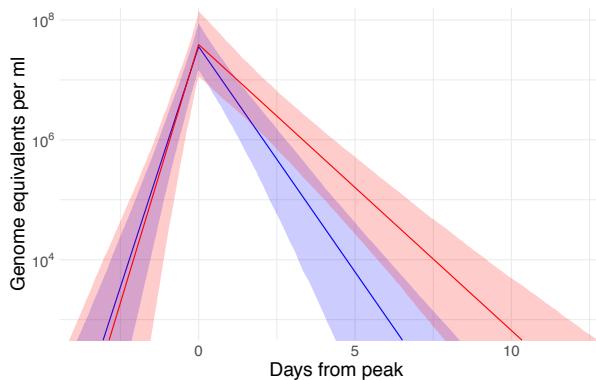


755

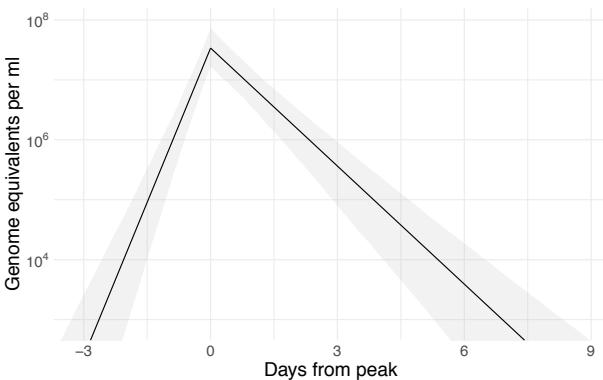
756

757

C)



D)



758

759

— Symptoms reported — Symptoms not reported

760

761

762

763

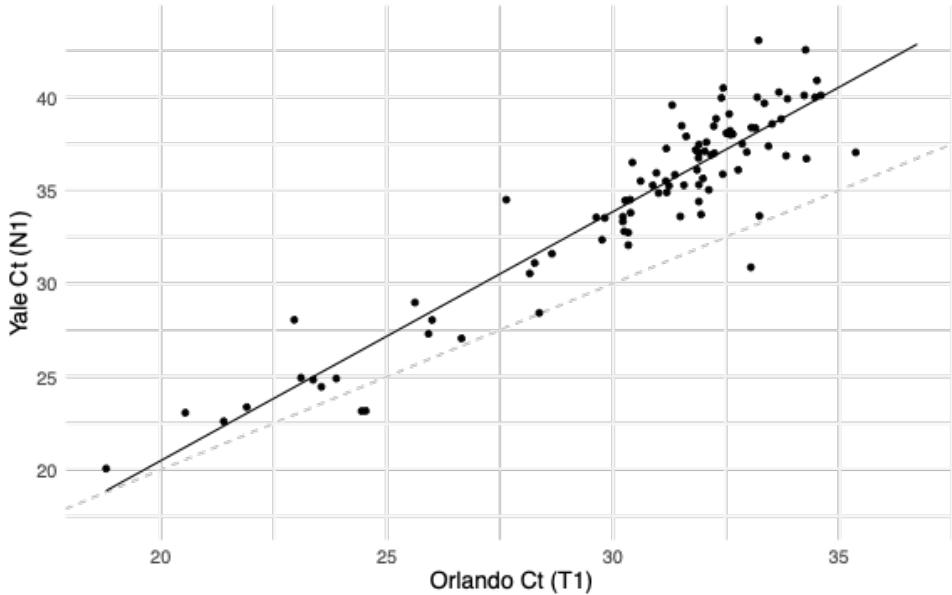
764

765

766

767

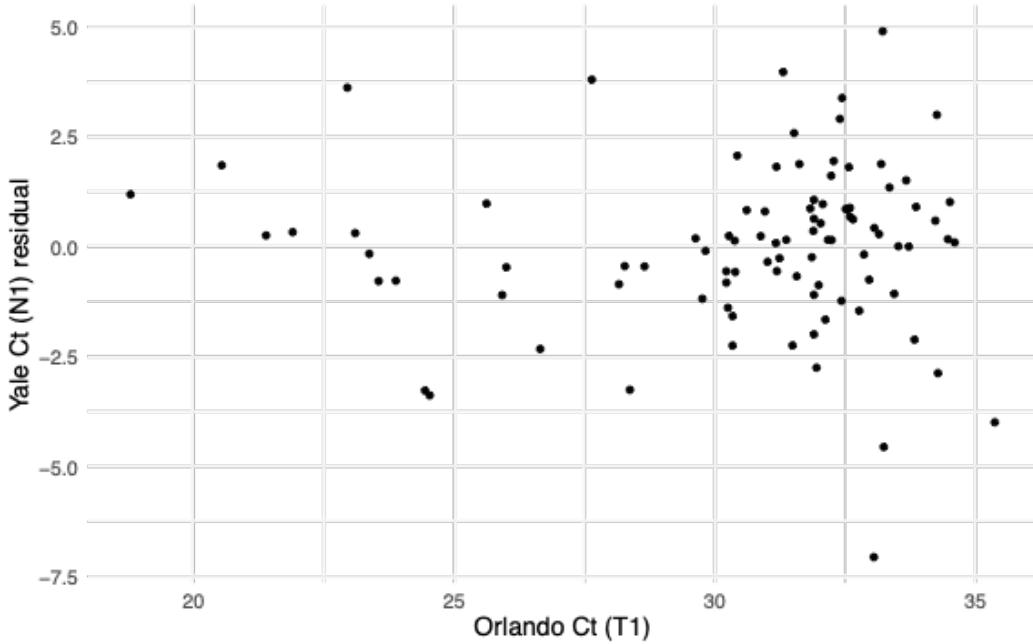
Supplemental Figure 11. Peak viral concentration and overall posterior viral concentration trajectories in terms of genome equivalents per ml. Posterior peak viral concentration distribution for symptomatic (red) and asymptomatic (blue) individuals (A) and for all individuals combined (B). Posterior viral concentration trajectory for symptomatic (red) and asymptomatic (blue) individuals (C) and for all individuals combined (D). The shaded bands in (C) and (D) represent 90% credible intervals for the trajectories.



768

769 **Supplemental Figure 12. Ct values from the Yale and Florida labs.** Points depict the Ct values for SARS-CoV-2 nasal swab
770 samples that were tested in both Florida and Yale labs. Ct values from Florida represent Target 1 (ORF1ab) on the Roche cobas
771 system, and Ct values from Yale represent N1 in the Yale multiplex assay. The solid black line depicts the best-fit linear regression
772 (intercept = -6.25, slope = 1.34, $R^2 = 0.86$). The dashed black line marks the 1-1 line where the points would be expected to fall if
773 the two labs were identical.

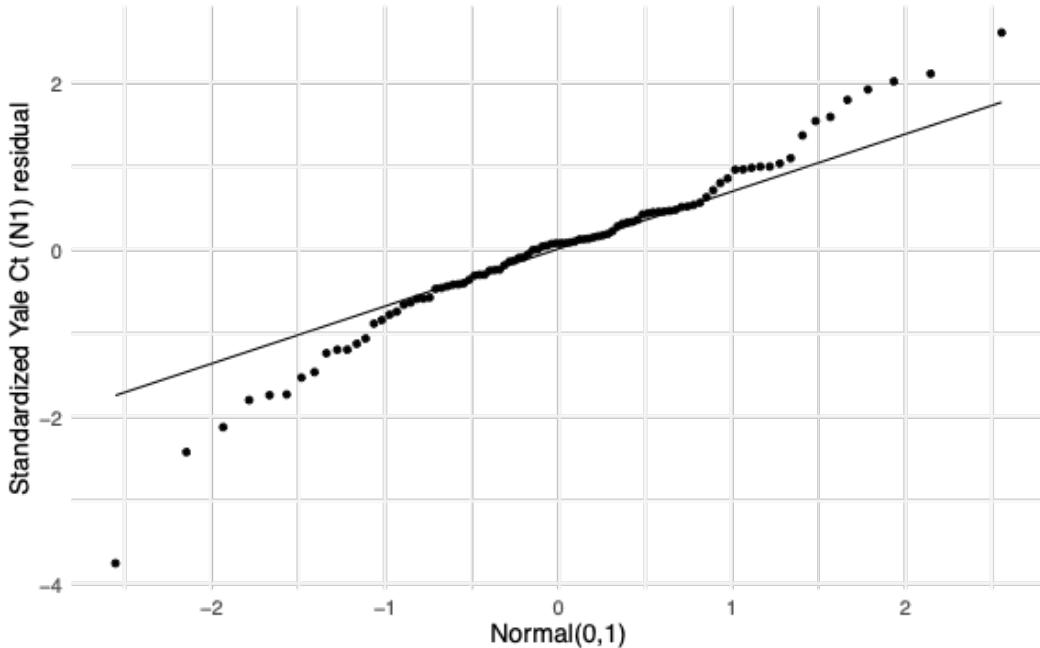
774



775

776 **Supplemental Figure 13. Residuals from the Yale/Florida Ct regression.** Points depict the residual after removing the best-fit
777 linear trend in the relationship between the Yale and Florida Ct values.

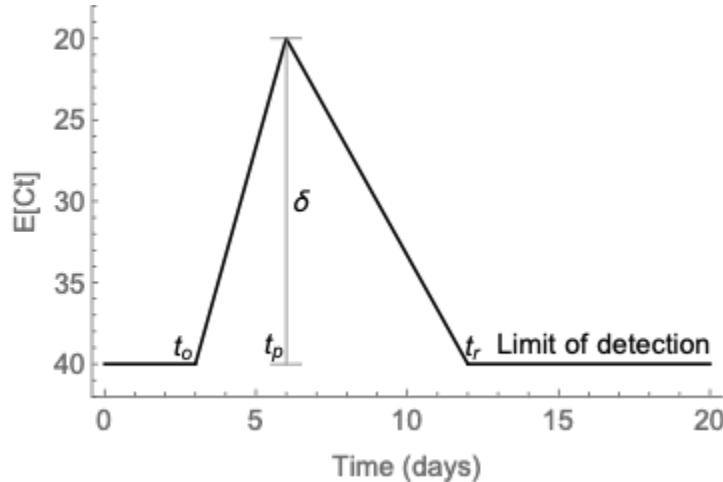
778



779

780 **Supplemental Figure 14. QQ plot of the residuals from the Yale/Florida Ct regression.** The residuals were standardized
781 (subtracted the mean and divided by the standard deviation) before comparing with the theoretical quantiles of a normal distribution
782 with mean 0 and standard deviation 1. The points depict the empirical quantiles of the data points and the line depicts the where
783 the points would be expected to fall if they were drawn from a standard normal distribution.

784



785

786

787

788

789

790

Supplemental Figure 15. A theoretical Ct trajectory. $E[Ct]$ is the expected Ct value on a given day. The Ct begins at the limit of detection, then declines from the time of infection (t_o) to the peak at δ cycles below the limit of detection at time t_p . The Ct then rises again to the limit of detection after t_r days. The model incorporating these parameter values used to generate this piecewise curve is given in Equation S1 (Methods).