

1 **Viral dynamics of SARS-CoV-2 infection and the predictive value of repeat testing**

2  
3 Stephen M. Kissler<sup>\*1</sup>, Joseph R. Fauver<sup>\*2</sup>, Christina Mack<sup>\*3,4</sup>, Caroline Tai<sup>3</sup>, Kristin Y. Shiue<sup>3,4</sup>,  
4 Chaney C. Kalinich<sup>2</sup>, Sarah Jednak<sup>5</sup>, Isabel M. Ott<sup>2</sup>, Chantal B.F. Vogels<sup>2</sup>, Jay Wohlgemuth<sup>6</sup>,  
5 James Weisberger<sup>7</sup>, John DiFiori<sup>8</sup>, Deverick J. Anderson<sup>9</sup>, Jimmie Mancell<sup>10</sup>, David D. Ho<sup>11</sup>,  
6 Nathan D. Grubaugh<sup>†2</sup>, Yonatan H. Grad<sup>†1</sup>

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

10 <sup>2</sup> Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven,  
11 CT

12 <sup>3</sup> IQVIA, Real World Solutions, Durham, NC

13 <sup>4</sup> Department of Epidemiology, University of North Carolina-Chapel Hill, Chapel Hill, NC

14 <sup>5</sup> Department of Health Management and Policy, University of Michigan School of Public  
15 Health, Ann Arbor, MI

16 <sup>6</sup> Quest Diagnostics, San Juan Capistrano, CA

17 <sup>7</sup> Bioreference Laboratories, Elmwood Park, NJ

18 <sup>8</sup> Hospital for Special Surgery, and the National Basketball Association, New York, NY

19 <sup>9</sup> Duke Center for Antimicrobial Stewardship and Infection Prevention, Durham, NC

20 <sup>10</sup> Department of Medicine, University of Tennessee Health Science Center, Memphis, TN

21 <sup>11</sup> Aaron Diamond AIDS Research Center, Columbia University Vagelos College of Physicians  
22 and Surgeons, New York, NY

23  
24 \* denotes equal contribution

25 † denotes co-senior authorship

27 **Abstract**

28 SARS-CoV-2 diagnostics that report viral RNA concentrations can be used to determine a  
29 patient's stage of infection, but this potential has not yet been realized due to a lack of  
30 prospective longitudinal data to calibrate such inferences. Here, we report the viral RNA  
31 trajectories for 68 individuals using quantitative PCR testing. On average, symptomatic and  
32 asymptomatic individuals reached similar peak viral RNA concentrations (22.2 Ct, 95% credible  
33 interval [19.1, 25.1] vs. 22.4 Ct [20.2, 24.5]) within similar amounts of time (2.9 days [0.7, 4.7] vs.  
34 3.0 days [1.3, 4.3]), but acute shedding lasted longer for symptomatic individuals (10.5 days [6.5,  
35 14.0] vs. 6.7 days [3.2, 9.2]). A second test within 2 days after an initial positive PCR result reliably  
36 indicated whether viral RNA concentration was increasing, decreasing, or in a low-level  
37 persistent phase. Quantitative viral RNA assessment, informed by viral trajectory, can improve  
38 algorithms for clinical and public health management.

39 **Main text.**

40  
41 As mortality from the COVID-19 pandemic surpasses one million, SARS-CoV-2 continues to  
42 cause hundreds of thousands of daily new infections (1). A critical strategy to curb the spread of  
43 the virus without imposing widespread lockdowns is to rapidly identify and isolate infectious  
44 individuals. Since symptoms are an unreliable indicator of infectiousness and infections are  
45 frequently asymptomatic (2), diagnostic tests are key to determining whether a person is infected  
46 and may be contagious.

47  
48 Real time quantitative reverse transcriptase polymerase chain reaction (RT-qPCR) tests are the  
49 gold standard for detecting SARS-CoV-2 infection. Normally, these tests yield a binary  
50 positive/negative diagnosis based on detection of viral RNA. However, they can also inform on  
51 the viral titer via the cycle threshold (Ct). The Ct is the number of thermal cycles needed to amplify  
52 sampled viral RNA to a detectable level: the higher the sampled viral RNA concentration, the  
53 lower the Ct. This inverse correlation between Ct and viral concentration makes RT-qPCR tests  
54 far more valuable than a binary diagnostic, as they can be used to reveal a person's progress  
55 through key stages of infection (3), assisting with clinical and public health decision-making.  
56 However, this potential has not yet been realized due to a lack of data describing complete Ct  
57 trajectories for individuals infected with SARS-CoV-2. The dynamics of the Ct during the earliest  
58 stages of infection, when contagiousness is rapidly increasing, have been especially unclear  
59 since viral testing is usually performed after the onset of symptoms, after viral RNA concentration  
60 has peaked and already begun to decline, and using only a single specimen (4,5).

61  
62 Here, we present the findings from an analysis of prospective longitudinal SARS-CoV-2 testing  
63 performed for players, staff and vendors participating in the occupational health program  
64 implemented as part of the resumption of the 2019-20 National Basketball Association (NBA)  
65 season. We report the results of 2,411 RT-qPCR tests from 68 individuals who provided at least  
66 one sample with a Ct value that was within the limit of detection.

67  
68 Using the Ct values and associated clinical metadata, we identified 46 individuals with probable  
69 new SARS-CoV-2 infections. The remaining individuals were assumed to be persistently  
70 shedding SARS-CoV-2 RNA due to an infection that occurred prior to the study period. This  
71 persistent RNA shedding can last for weeks after an acute infection and seems more likely to  
72 represent non-infectious RNA than infectious virus (6). For the 46 acute infections, we estimated  
73 the peak Ct value, the time from onset of infection to peak, and the time from peak to conclusion  
74 of acute viral shedding using a Bayesian statistical model. Based on these inferences, we used  
75 the full pool of 68 individuals to estimate the probability that a given Ct value was associated  
76 with the acute vs. persistent stages of viral shedding. Within the acute stage, we also estimated  
77 whether a given Ct value was associated with the initial period of viral proliferation (increasing  
78 viral titers) or the subsequent period of viral clearance (decreasing viral titers) (**Figure 1**). We  
79 found that a second test within two days of an initial positive can help determine whether a  
80 person is acutely shedding viral RNA and can substantially clarify whether the person is in the  
81 proliferation or the clearance stage. These findings underscore the potential value of integrating  
82 viral Ct trajectory data into viral diagnostics and surveillance algorithms to inform clinical and  
83 public health decision-making.

85 The study population consisted of NBA players, staff, and vendors associated with the NBA's  
86 season restart. The study period began in teams' local cities from June 23<sup>rd</sup> through July 9<sup>th</sup>,  
87 2020, and testing continued for all teams as they transitioned to Orlando, Florida through  
88 September 7<sup>th</sup>, 2020. We report data from 68 individuals (90% male) who were tested at least  
89 five times during the study period and recorded at least one positive Ct value (<40). Due to a  
90 lack of new infections among players and team staff after clearing quarantine in Orlando, all  
91 players and team staff included in the results pre-date the Orlando phase of the restart. Of the  
92 individuals included in the study, 27 of the 46 with active infections and 40 of the 68 overall were  
93 from staff and vendors. The median number of tests administered to each of the 68 individuals  
94 was 41 (IQR [14, 51]; Range [5, 70]). The median number of positive Ct values recorded for each  
95 person was 3 (IQR [2, 4]; Range [1, 9]). A diagnosis of "acute" or "persistent" infection was  
96 abstracted from physician records. "Acute" denoted a likely new infection. "Persistent" indicated  
97 the presence of virus in a clinically recovered individual, likely due to infection that developed  
98 prior to the onset of the study. The Ct trajectories for all 68 individuals included in the analysis  
99 with their designations of acute or persistent infection are depicted in **Supplemental Figures 1–**  
100 **4**. If an individual reported symptoms on the day of the initial positive test, these were recorded.  
101

102 Clinical samples were obtained by combined swabs of the anterior nares and oropharynx  
103 administered by a trained provider. The samples were initially tested by either Quest Diagnostics  
104 (while teams were in local markets using the Quest SARS-CoV-2 RT-qPCR (7)) or BioReference  
105 Laboratories (while teams were in Orlando using the cobas SARS-CoV-2 test (8)). Viral transport  
106 media from positive samples were sent to Yale University for subsequent RT-qPCR testing using  
107 a multiplexed version of the assay from the US Centers for Disease Control and Prevention (9)  
108 to normalize Ct values across testing platforms. A total of 234 samples from BioReference and  
109 128 from Quest were tested at Yale; 49 positive samples had Ct values assigned on first testing  
110 but did not undergo repeat testing at the Yale laboratory. To account for the different calibration  
111 of the testing instruments, we used a linear conversion (**Supplemental Figures 5–7**,  
112 **Supplemental Methods**) to adjust these samples to the Yale laboratory scale. Subsequent  
113 analysis is based on the N1 Ct value from the Yale multiplex assay and on the adjusted Roche  
114 cobas target 1 assay.

115  
116 We used a Bayesian statistical model to infer the peak Ct value and the durations of the  
117 proliferation and clearance stages for the 46 acute infections with at least one Ct value below 35  
118 (**Figure 1**). We assumed that the proliferation and clearance stages could respectively be  
119 characterized by periods of exponential growth and exponential decay in viral RNA concentration  
120 (10). The Ct is roughly proportional to the negative logarithm of the viral RNA concentration (3),  
121 and thus Ct values are expected to decrease linearly to a peak and then increase linearly until  
122 the resolution of the acute infection (**Supplemental Figure 8**). Resolution may be followed by  
123 intermittent positive Ct values near the limit of detection (6), corresponding to persistent  
124 shedding of virus or viral fragments. We used piecewise-linear regression to fit the Ct trajectories,  
125 accounting for rare false negatives during the acute infection period. Full details are provided in  
126 the **Supplemental Methods** and code may be accessed online [<https://github.com/gradlab/CtTrajectories>].  
127

128  
129 Of the 46 individuals with acute infections, 13 reported symptoms at the time of diagnosis. The  
130 mean peak Ct for symptomatic individuals was 22.2 (95% credible interval [19.1, 25.1]), the mean  
131 duration of the proliferation phase was 2.9 days [0.7, 4.7], and the mean duration of clearance

132 was 10.5 days [6.5, 14.0] (**Figure 2**). This compares with 22.4 Ct [20.2, 24.5], 3.0 days [1.3, 4.3],  
133 and 6.7 days [3.2, 9.2], respectively, for individuals who did not report symptoms at the time of  
134 diagnosis (**Figure 2**). In summary, there was little difference in mean Ct and proliferation duration  
135 between individuals with known symptoms vs. no known symptoms, but individuals with  
136 symptoms had a relatively longer clearance phase than individuals with no known symptoms.  
137 This yielded a slightly longer overall duration of acute infection for individuals who reported  
138 symptoms (13.4 days [9.3, 17.1]) vs. those who did not (9.7 days [6.0, 12.5]) (**Figure 2D-E**). For  
139 all individuals regardless of symptoms, the mean peak Ct value, proliferation duration, clearance  
140 duration, and duration of acute shedding were 22.4 Ct [20.6, 24.1], 2.7 days [1.2, 3.8], 7.4 days  
141 [3.9, 9.6], and 10.1 days [6.5, 12.6] (**Supplemental Figure 9**). There was a substantial amount of  
142 individual-level variation in the peak Ct and the proliferation and clearance stage durations  
143 (**Supplemental Figures 10-13**). Combined posterior distributions for these quantities at the  
144 individual level are depicted in **Supplemental Figure 14** with best-fit normal and gamma  
145 distributions as appropriate.

146

147 Using the full dataset of 68 individuals, we estimated the probability that a positive Ct was  
148 associated with an acute infection (*i.e.*, the proliferation or clearance phase, but not the  
149 persistent phase), and if so, the probability that it was associated with just the proliferation stage.  
150 First, we assigned to each positive sample the probability that it was collected during each of  
151 the three stages of infection. To do so, we began with the positive samples from the 46  
152 individuals with acute infections and calculated the frequency with which each sample sat within  
153 the proliferation stage, the clearance stage, or the persistent stage (*i.e.*, neither of the previous  
154 two stages) across 10,000 posterior parameter draws for that person. For the remaining 22  
155 individuals, all positive samples were assigned to the persistent stage. Next, we calculated the  
156 probability that a Ct value falling within a 5-unit window corresponded to an active infection (*i.e.*,  
157 either the proliferation or the clearance stage) by summing the proliferation and clearance  
158 probabilities for all positive samples with that window and dividing by the total number of positive  
159 samples in the window. We considered windows with midpoints spanning from Ct = 37.5 to Ct  
160 = 15.5 (**Figure 3A**). We performed a similar calculation to determine the probability that a Ct  
161 falling within a given 5-unit window corresponded to just the proliferation phase, assuming it had  
162 already been determined that the sample fell within an active infection (**Figure 3B**). The  
163 probability of active infection increased rapidly with decreasing Ct (increasing viral  
164 concentration), but the Ct provided little information about whether the infection was in the  
165 proliferation or the clearance stage.

166

167 Next, we assessed whether a second test within two days of the first could improve these  
168 predictions. We collected all positive samples with a subsequent sample (positive or negative)  
169 that was taken within two days and repeated the above calculations, calculating the probability  
170 that a Ct value within a five-unit window, followed by a second test with higher/lower Ct,  
171 corresponded to an active vs. persistent infection and to the proliferation vs. clearance stages.  
172 A positive test followed by a second test with lower Ct (higher viral RNA concentration) was  
173 slightly more likely to be associated with an active infection than a positive test alone (**Figure**  
174 **3A**). Similarly, a positive test followed by a second test with lower Ct (higher viral RNA  
175 concentration) was much more likely to be associated with the proliferation phase than with the  
176 clearance phase (**Figure 3B**).

177

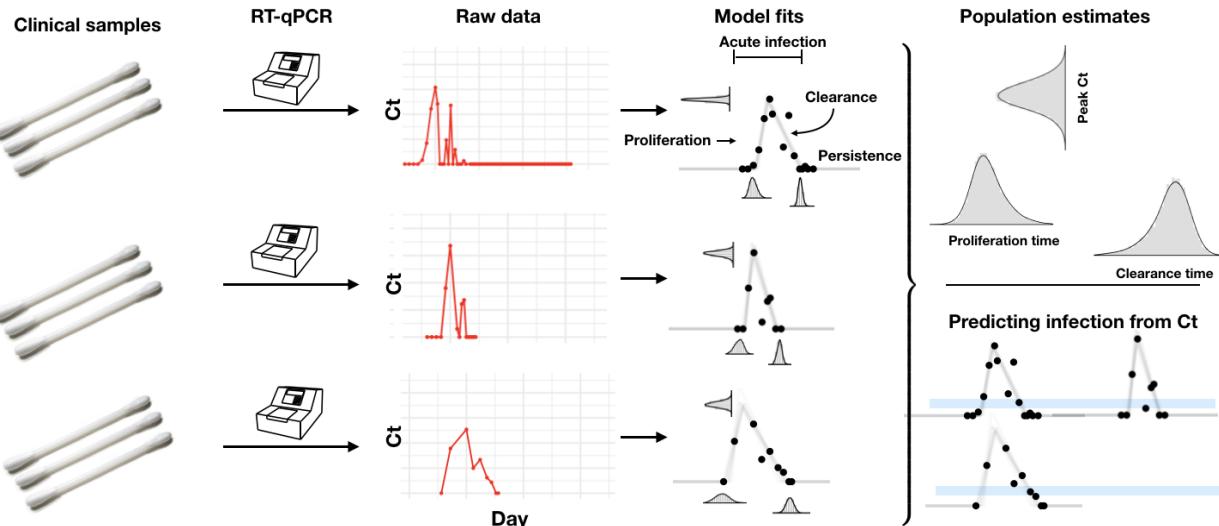
178 This report provides first comprehensive data on the early-infection RT-qPCR Ct dynamics  
179 associated with SARS-CoV-2 infection. Our findings highlight that repeated PCR tests can be  
180 used to infer the stage of a patient's infection. While a single test can inform on whether a patient  
181 is in the acute or persistent viral RNA shedding stages, a subsequent test can help identify  
182 whether viral RNA concentrations are increasing or decreasing, thus informing clinical care.  
183 Moreover, as contagiousness varies over the course of an infection (3), our results also indicate  
184 how Ct values and dynamics can inform on the nature and duration of interventions needed to  
185 reduce the risk of onward transmission.

186  
187 Our findings on the duration of SARS-CoV-2 viral RNA shedding expand on and agree with  
188 previous studies (11) and with observations that peak Ct does not differ substantially between  
189 symptomatic and asymptomatic individuals (4). While previous studies have largely relied on  
190 serial sampling of admitted hospital patients, our study used prospective sampling of ambulatory  
191 infected individuals to characterize complete viral dynamics for the presymptomatic stage and  
192 for individuals who did not report symptoms. This allowed us to assess differences between the  
193 viral RNA proliferation and clearance stages for individuals with and without reported symptoms.  
194 The similarity in the early-infection viral RNA dynamics for both symptomatic and asymptomatic  
195 individuals underscores the need for SARS-CoV-2 screening regardless of symptoms. The rapid  
196 progression from a negative test to a peak Ct value 2-4 days later provides empirical support for  
197 screening and surveillance strategies that employ frequent rapid testing to identify potentially  
198 infectious individuals (12, 13). Taken together, the dynamics of viral RNA shedding substantiate  
199 the need for frequent population-level SARS-CoV-2 screening and a greater availability of  
200 diagnostic tests.

201  
202 Our findings are limited for several reasons. The cohort does not constitute a representative  
203 sample from the population, as it was predominantly male and included professional athletes.  
204 Some of the trajectories were sparsely sampled, limiting the precision of our posterior estimates.  
205 The reporting of symptoms was imperfect, as follow-up to identify individuals who developed  
206 symptoms after initial evaluation was not systematic. As with all predictive tests, the probabilities  
207 that link Ct values with infection stages (**Figure 3**) pertain to the population from which they were  
208 calibrated and do not necessarily generalize to other populations for which the prevalence of  
209 infection and testing protocols may differ. Still, we anticipate that the central patterns will hold  
210 across populations: first, that low Cts (<30) strongly predict acute infection, and second, that a  
211 follow-up test collected within two days of an initial positive test can substantially help to discern  
212 whether a person is closer to the beginning or the end of their infection. Our study did not test  
213 for the presence of infectious virus, though previous studies have documented a close inverse  
214 correlation between Ct values and culturable virus (14).

215  
216 Prospective longitudinal studies using larger samples of representative populations will be  
217 valuable in affirming the generalizability of our findings. We demonstrated the value of Ct values  
218 from paired PCR tests, focusing on two tests because of the real-world challenges to performing  
219 multiple tests; future studies could further specify the precise predictive probabilities of a given  
220 sequence of Ct values, establishing the number and sequence of testing that would make for a  
221 clinically useful decision tool. Such studies should be carried out in demographically diverse  
222 populations with varying levels of SARS-CoV-2 prevalence, as the predictive probabilities of Ct  
223 sequences will vary according to individual characteristics (e.g. age) and the level of circulation  
224 in the population (3).

225  
226 To manage the spread of SARS-CoV-2, we must develop novel technologies and find new ways  
227 to extract more value from the tools that are already available. Our results, building on the first  
228 reports of the dynamics of the proliferation stage, suggest that integrating the quantitative viral  
229 RNA trajectory into algorithms for clinical management could offer benefits. The ability to chart  
230 a patient's progress through their infection underpins our ability to provide appropriate clinical  
231 care and to institute effective measures to reduce the risk of onward transmission. Marginally  
232 more sophisticated diagnostic and screening algorithms may greatly enhance our ability to  
233 manage the spread of SARS-CoV-2 using tests that are already available.

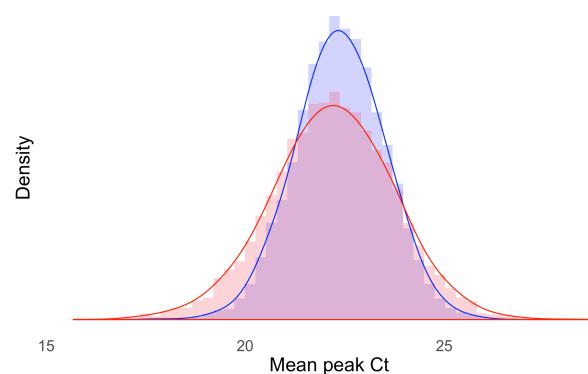


234  
235

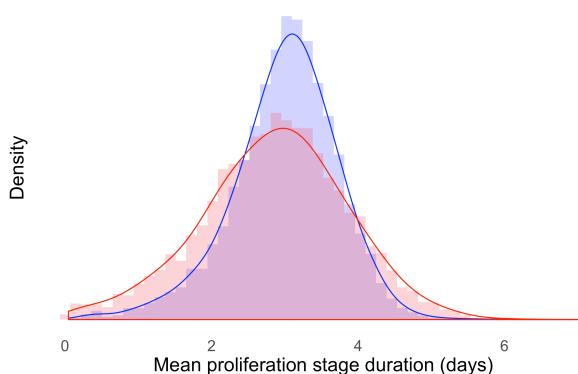
236 **Figure 1. Illustration of the analysis pipeline.** Combined anterior nares and oropharyngeal swabs were tested using  
237 a RT-qPCR assay to generate longitudinal Ct values ('Raw data', red points) for each person. Using a statistical model,  
238 we estimated Ct trajectories consistent with the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

244

A)



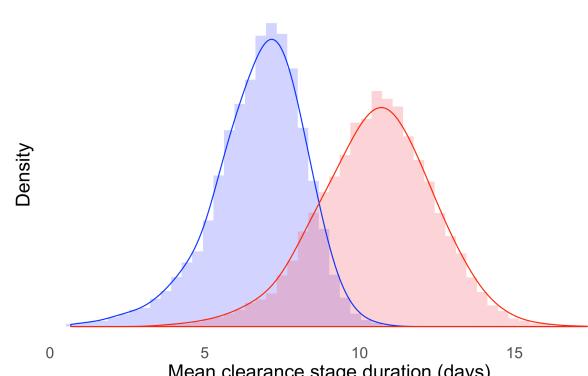
B)



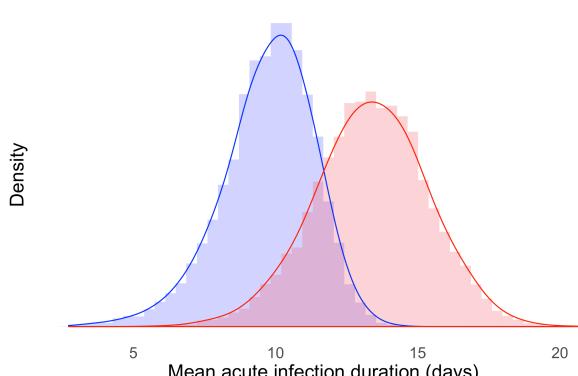
245

246

C)



D)

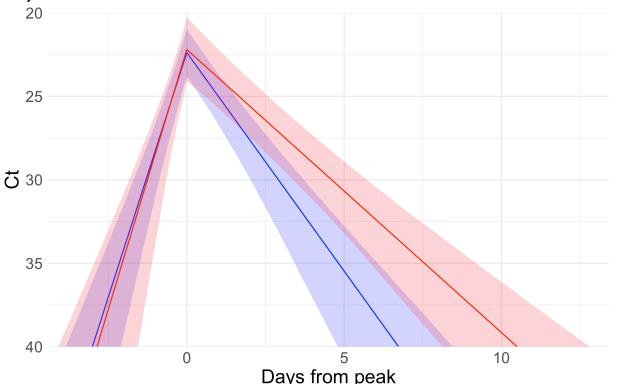


247

248

249

E)



250

251

252

253

254

255

256

257

258

259

260

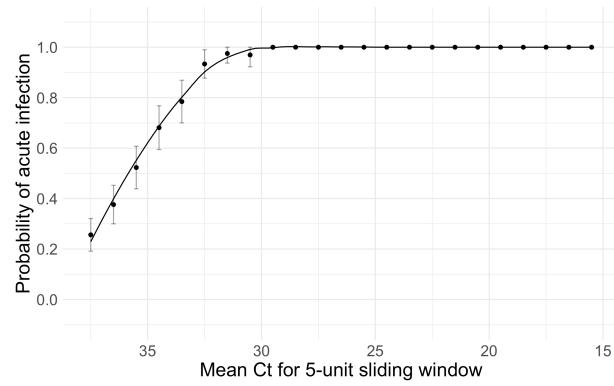
261

262

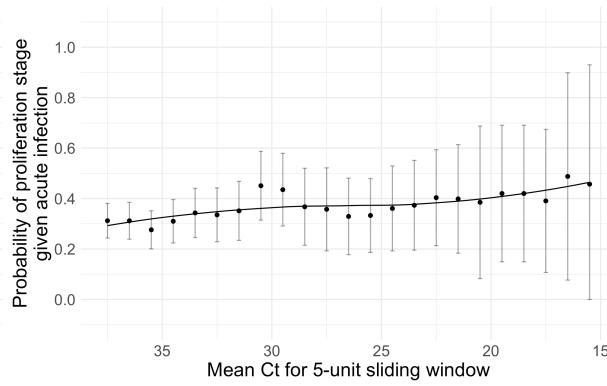
**Figure 2. Peak Ct and infection stage duration distributions according to symptoms reported at time of diagnosis.** Histograms (colored bars) of 10,000 simulated draws from the posterior distributions for mean peak Ct value (A), mean duration of the proliferation stage (infection detection to peak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

263

A)



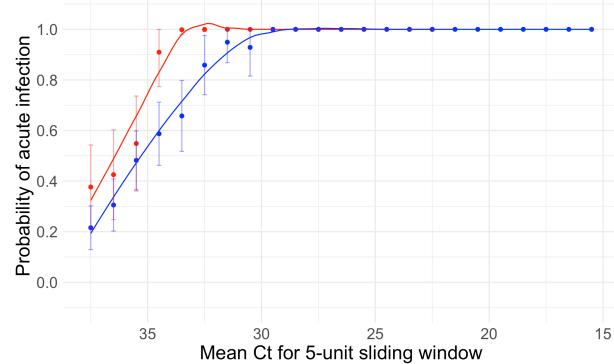
B)



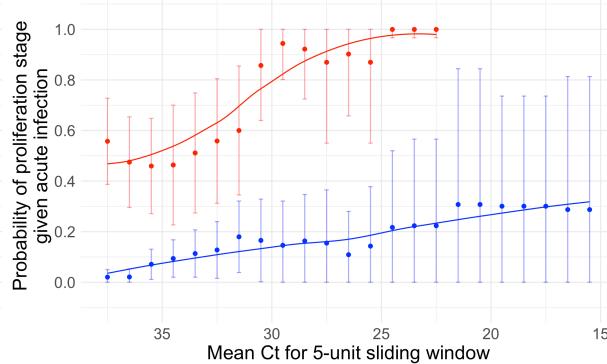
264

265

C)



D)



266

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

267

268

269

270

271

272

273

274

275

**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 positive Ct, while sub-figures C and D depict the predictive probabilities for a positive Ct 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.

## 276 References

- 278 1. World Health Organization. 2020 Coronavirus disease (COVID-19) Situation Report.
- 279 2. Furukawa NW, Brooks JT, Sobel J. 2020 Evidence Supporting Transmission of Severe  
280 Acute Respiratory Syndrome Coronavirus 2 While Presymptomatic or Asymptomatic.  
*Emerg. Infect. Dis.* **26**, (doi:10.3201/eid2607.201595)
- 281 3. Tom MR, Mina MJ. 2020 To Interpret the SARS-CoV-2 Test, Consider the Cycle  
282 Threshold Value. *Clin. Infect. Dis.* **02115**, 1–3. (doi:10.1093/cid/ciaa619)
- 283 4. Walsh KA et al. 2020 SARS-CoV-2 detection, viral load and infectivity over the course of  
284 an infection. *J. Infect.* **81**, 357–371. (doi:10.1016/j.jinf.2020.06.067)
- 285 5. Wyllie AL et al. 2020 Saliva or Nasopharyngeal Swab Specimens for Detection of SARS-  
286 CoV-2. *N. Engl. J. Med.* **February**, NEJMc2016359. (doi:10.1056/NEJMc2016359)
- 287 6. Xiao AT, Tong YX, Zhang S. 2020 Profile of RT-PCR for SARS-CoV-2: A Preliminary  
288 Study From 56 COVID-19 Patients. *Clin. Infect. Dis.* (doi:10.1093/cid/ciaa460)
- 289 7. U.S. Food and Drug Administration. 2020 Quest Diagnostics Infectious Disease, Inc.  
290 (“Quest Diagnostics”) SARS-CoV-2 RNA Qualitative Real-Time RT-PCR Emergency Use  
291 Authorization.
- 292 8. U.S. Food and Drug Administration. 2020 Roche Molecular Systems, Inc. cobas SARS-  
293 CoV-2 Emergency Use Authorization.
- 294 9. Kudo E et al. 2020 Detection of SARS-CoV-2 RNA by multiplex RT-qPCR. *PLOS Biol.*  
295 **18**, e3000867. (doi:10.1371/journal.pbio.3000867)
- 296 10. Cleary B, Hay JA, Blumenstiel B, Gabriel S, Regev A, Mina MJ. 2020 Efficient prevalence  
297 estimation and infected sample identification with group testing for SARS-CoV-2.  
298 *medRxiv*
- 299 11. Cevik M, Tate M, Lloyd O, Maraolo AE, Schafers J, Ho A. 2020 SARS-CoV-2, SARS-  
300 CoV-1 and MERS-CoV viral load dynamics, duration of viral shedding and  
301 infectiousness: a living systematic review and meta-analysis. *medRxiv*
- 302 12. Larremore DB, Wilder B, Lester E, Shehata S, Burke JM, Hay JA, Tambe M, Mina MJ,  
303 Parker R. 2020 Test sensitivity is secondary to frequency and turnaround time for  
304 COVID-19 surveillance. *medRxiv*, 2020.06.22.20136309.  
(doi:10.1101/2020.06.22.20136309)
- 305 13. Paltiel AD, Zheng A, Walensky RP. 2020 Assessment of SARS-CoV-2 Screening  
306 Strategies to Permit the Safe Reopening of College Campuses in the United States.  
307 *JAMA Netw. Open* **3**, e2016818. (doi:10.1001/jamanetworkopen.2020.16818)
- 308 14. Singanayagam A, Patel M, Charlett A, Lopez Bernal J, Saliba V, Ellis J, Ladhani S,  
309 Zambon M, Gopal R. 2020 Duration of infectiousness and correlation with RT-PCR cycle  
310 threshold values in cases of COVID-19, England, January to May 2020. *Euro Surveill.* **25**,  
311 1–5. (doi:10.2807/1560-7917.ES.2020.25.32.2001483)
- 312 15. Ott IM, Vogels C, Grubaugh N, Wyllie AL. 2020 Saliva Collection and RNA Extraction for  
313 SARS-CoV-2 Detection V.2.
- 314 16. Carpenter B et al. 2017 Stan: A Probabilistic Programming Language. *J. Stat. Softw.* **76**.  
(doi:10.18637/jss.v076.i01)
- 315 17. R Development Core Team R. 2011 R: A Language and Environment for Statistical  
316 Computing. *R Found. Stat. Comput.* **1**, 409. (doi:10.1007/978-3-540-74686-7)

321 **Acknowledgements.** We thank the NBA, National Basketball Players Association (NBPA), and  
322 all of the study participants who are committed to applying what they learned from sports  
323 towards enhancing public health. In particular, we thank D. Weiss of the NBA for his continuous  
324 support and leadership. We are appreciative of the discussions from the COVID-19 Sports and  
325 Society Working Group. We also thank J. Hay and R. Niehus for helpful suggestions on the  
326 statistical approach and P. Jack and S. Taylor for laboratory support.  
327

328 **Funding.** This study was funded by the NWO Rubicon 019.181EN.004 (CBFV), a clinical research  
329 agreement with the NBA and NBPA (NDG), the Huffman Family Donor Advised Fund (NDG), Fast  
330 Grant funding support from the Emergent Ventures at the Mercatus Center, George Mason  
331 University (NDG), and the Morris-Singer Fund for the Center for Communicable Disease  
332 Dynamics at the Harvard T.H. Chan School of Public Health (YHG).  
333

334 **Author contributions.** SMK conceived of the study, conducted the statistical analysis, and  
335 wrote the manuscript. JRF conceived of the study, conducted the laboratory analysis, and wrote  
336 the manuscript. CM conceived of the study, collected the data, and wrote the manuscript. CT  
337 analyzed the data and edited the manuscript. KYS analyzed the data and edited the manuscript.  
338 CCK conducted the laboratory analysis and edited the manuscript. SJ conducted the laboratory  
339 analysis and edited the manuscript. IMO conducted the laboratory analysis. CBFV conducted  
340 the laboratory analysis. JW conducted laboratory analysis and edited the manuscript. JW  
341 conducted laboratory analysis and edited the manuscript. JD conceived of the study and edited  
342 the manuscript. DJA contributed to data analysis and edited the manuscript. JM contributed to  
343 data analysis and edited the manuscript. DDH conceived of the study and edited the manuscript.  
344 NDG conceived of the study, oversaw the study, and wrote the manuscript. YHG conceived of  
345 the study, oversaw the study, and wrote the manuscript.  
346

347 **Competing interests.**

348 JW is an employee of Quest Diagnostics. JW is an employee of Bioreference Laboratories.

349 **Supplement.**

350 Ethics.

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

357

358 Additional testing protocol details.

359 Residual viral transport media (VTM) from Quest Diagnostics or BioReference Laboratories were  
360 shipped overnight to Yale on dry ice. VTM was thawed on ice and 300 µL was used for RNA  
361 extraction using the MagMAX Viral/Pathogen Nucleic Acid Isolation Kit and the KingFisher Flex  
362 robot (Thermo Fisher Scientific, Waltham, MA (15)). Total nucleic acid was eluted into 75ul of  
363 elution buffer and SARS-CoV-2 RNA was quantified from 5 µL of extracted total RNA using a  
364 multiplexed version of the CDC RT-qPCR assay that contains the 2019-nCoV\_N1 (N1), 2019-  
365 nCoV-N2 (N2), and human RNase P (RP) primer-probe sets (9). The RT-qPCR was performed  
366 using the Luna Universal Probe One-Step RT-qPCR Kit (New England Biolabs, Ipswich, MA, US)  
367 and the following thermocycler conditions: (1) reverse transcription for 10 minutes at 55°C, (2)  
368 initial denaturation for 1 min at 95°C, and PCR for 45 cycles of 10 seconds at 95°C and 30  
369 seconds at 55°C on the CFX96 qPCR machine (Bio-Rad, Hercules, CA, US).

370

371 Converting Ct values. Most ( $n = 226$ ) of the 312 positive samples in the raw dataset underwent  
372 RT-qPCR at the Yale laboratory. We used the Yale Ct value whenever it was available. Still, 86  
373 samples underwent initial diagnostic testing at BioReference Laboratories but not confirmatory  
374 testing at the Yale laboratory. Both platforms rely on a multiplex RT-qPCR strategy. The two  
375 testing platforms yield slightly different Ct values, as evidenced by the 94 samples that underwent  
376 RT-qPCR at both facilities (**Supplemental Figure 5**). For comparison between platforms, target  
377 1 from the Roche cobas assay, which is specific to SARS-CoV-2, and the N1 target from the  
378 Yale multiplex assay were used. For the 86 samples that were not processed at the Yale  
379 laboratory, we adjusted the Ct values using the best-fit (minimum sum of squares) linear  
380 regression between the initial Ct value and the Yale Ct value for the samples that were processed  
381 in both facilities. To do so, we estimated the coefficients  $\beta_0$  and  $\beta_1$  in the following regression  
382 equation:

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

383

384

385 Here,  $y_i$  denotes the  $i^{th}$  Ct value from Yale,  $x_i$  denotes the  $i^{th}$  Ct value from the initial test, and  $\epsilon_i$  is  
386 an error term with mean 0 and constant variance across all samples. The resulting fit  
387 (**Supplemental Figure 5**) was strong ( $R^2 = 0.86$ ) with homoscedastic residuals (**Supplemental**  
388 **Figure 6**) that are approximately normally distributed, as evidenced by a Q-Q plot  
389 (**Supplemental Figure 7**).

390

391 Data parsing. The raw dataset included 3,207 test results for 102 individuals. We excluded 21  
392 individuals who had 5 or fewer tests, since the data for these individuals were too sparse to  
393 reliably infer a Ct trajectory. We also excluded 13 individuals who did not record any Ct values  
394 that surpassed the RT-qPCR limit of detection (40). We removed 146 entries for which the test  
395 result was recorded as ‘positive’ but there was no associated Ct value; these tests were initially

396 conducted on an instrument that provided only a binary diagnosis and the samples were not  
397 available for confirmatory testing. This left 2,411 total tests for 68 individuals for the main  
398 analysis. We trivially shifted the date indices so that date 0 corresponded to the time of the  
399 minimum Ct. We set the Ct value for negative tests equal to the limit of detection. For the  
400 statistical analysis, we removed any sequences of 3 or more consecutive negative tests to avoid  
401 overfitting to these trivial values.  
402

403 **Model fitting.** We assumed that the viral concentration trajectories consisted of a proliferation  
404 phase, with exponential growth in viral RNA concentration, followed by a clearance phase  
405 characterized by exponential decay in viral RNA concentration. Since Ct values are roughly  
406 proportional to the negative logarithm of viral concentration, this corresponds to a linear  
407 decrease in Ct followed by a linear increase. We therefore constructed a piecewise-linear  
408 regression model to estimate the peak Ct, the time from infection onset to peak (i.e. the duration  
409 of the proliferation stage), and the time from peak to infection resolution (i.e. the duration of the  
410 clearance stage). This idealized trajectory is depicted in **Supplemental Figure 8**. The trajectory  
411 may be represented by the equation  
412

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

413  
414 Here,  $E[Ct(t)]$  represents the expected value of the Ct at time  $t$ , “l.o.d” represents the RT-qPCR  
415 limit of detection,  $\delta$  is the absolute difference in Ct between the limit of detection and the peak  
416 (lowest) Ct, and  $t_o$ ,  $t_p$ , and  $t_r$  are the onset, peak, and recovery times, respectively.  
417  
418

419 Before fitting, we re-parametrized the model using the following definitions:  
420

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

425  
426 We constrained  $0 \leq \omega_p \leq 14$  days and  $0 \leq \omega_c \leq 30$  days to prevent inferring unrealistically large  
427 values for these parameters for trajectories that were missing data prior to the peak and after the  
428 peak, respectively. We also constrained  $0 \leq \delta \leq 40$  as Ct values can only take values between 0  
429 and the limit of detection (40).

430  
431 We next assumed that the observed  $\Delta Ct(t)$  could be described the following mixture model:  
432

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

433  
434 where  $E[\Delta Ct(t)] = \text{l.o.d.} - E[Ct(t)]$  and  $\lambda$  is the sensitivity of the q-PCR test, which we fixed at 0.99.  
435 The bracket term on the right-hand side of the equation denotes that the distribution was  
436

437 truncated to ensure Ct values between 0 and the limit of detection. This model captures the  
438 scenario where most observed Ct values are normally distributed around the expected trajectory  
439 with standard deviation  $\sigma(t)$ , yet there is a small (1%) probability of an exponentially-distributed  
440 false negative near the limit of detection. The log(10) rate of the exponential distribution was  
441 chosen so that 90% of the mass of the distribution sat below 1 Ct unit and 99% of the distribution  
442 sat below 2 Ct units, ensuring that the distribution captures values distributed at or near the limit  
443 of detection. We did not estimate values for  $\lambda$  or the exponential rate because they were not of  
444 interest in this study; we simply needed to include them to account for some small probability  
445 mass that persisted near the limit of detection to allow for the possibility of false negatives.  
446

447 For the 86 samples that were not tested in the Yale laboratory, we included additional uncertainty  
448 in the observed Ct value by inflating  $\sigma(t)$ , such that  
449

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

450 Here,  $\sigma(\text{tilde})$  is a constant,  $\epsilon$  is the standard deviation of the residuals from the linear fit between  
451 the initial test and the Yale laboratory test, and  $I_{adj}$  is an indicator variable that is 1 if the sample  
452 at time  $t$  was adjusted and 0 otherwise.  
453

454 We used a hierarchical structure to describe the distributions of  $\omega_p$ ,  $\omega_r$ , and  $\delta$  for each individual  
455 based on their respective population means  $\mu_{\omega_p}$ ,  $\mu_{\omega_r}$ , and  $\mu_\delta$  and population standard deviations  
456  $\sigma_{\omega_p}$ ,  $\sigma_{\omega_r}$ , and  $\sigma_\delta$  such that  
457

$$\begin{aligned} \omega_p &\sim \text{Normal}(\mu_{\omega_p}, \sigma_{\omega_p}) \\ \omega_r &\sim \text{Normal}(\mu_{\omega_r}, \sigma_{\omega_r}) \\ \delta &\sim \text{Normal}(\mu_\delta, \sigma_\delta) \end{aligned}$$

458 We inferred separate population means ( $\mu$ ) for symptomatic and asymptomatic individuals. We  
459 used a Hamiltonian Monte Carlo fitting procedure implemented in Stan (version 2.24) (16) and R  
460 (version 3.6.2) (17) to estimate the individual-level parameters  $\omega_p$ ,  $\omega_r$ ,  $\delta$ , and  $t_p$  as well as the  
461 population-level parameters  $\sigma(\text{tilde})$ ,  $\mu_{\omega_p}$ ,  $\mu_{\omega_r}$ ,  $\mu_\delta$ ,  $\sigma_{\omega_p}$ ,  $\sigma_{\omega_r}$ , and  $\sigma_\delta$ . We used the following priors:  
462

463 *Hyperparameters:*  
464

$$\sigma(\text{tilde}) \sim \text{Cauchy}(0, 5) [0, \infty]$$

$$\begin{aligned} \mu_{\omega_p} &\sim \text{Normal}(14/2, 14/6) [0, 14] \\ \mu_{\omega_r} &\sim \text{Normal}(30/2, 30/6) [0, 30] \\ \mu_\delta &\sim \text{Normal}(40/2, 40/6) [0, 40] \end{aligned}$$

$$\sigma_{\omega_p} \sim \text{Cauchy}(0, 14/\tan(\pi(0.95-0.5))) [0, \infty]$$

$$\sigma_{\omega_r} \sim \text{Cauchy}(0, 30/\tan(\pi(0.95-0.5))) [0, \infty]$$

$$\sigma_\delta \sim \text{Cauchy}(0, 40/\tan(\pi(0.95-0.5))) [0, \infty]$$

481 *Individual-level parameters:*

482  $\omega_p \sim \text{Normal}(\mu_{\omega_p}, \sigma_{\omega_p}) [0, 14]$

483  $\omega_r \sim \text{Normal}(\mu_{\omega_r}, \sigma_{\omega_r}) [0, 30]$

484  $\delta \sim \text{Normal}(\mu_\delta, \sigma_\delta) [0, 40]$

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

486

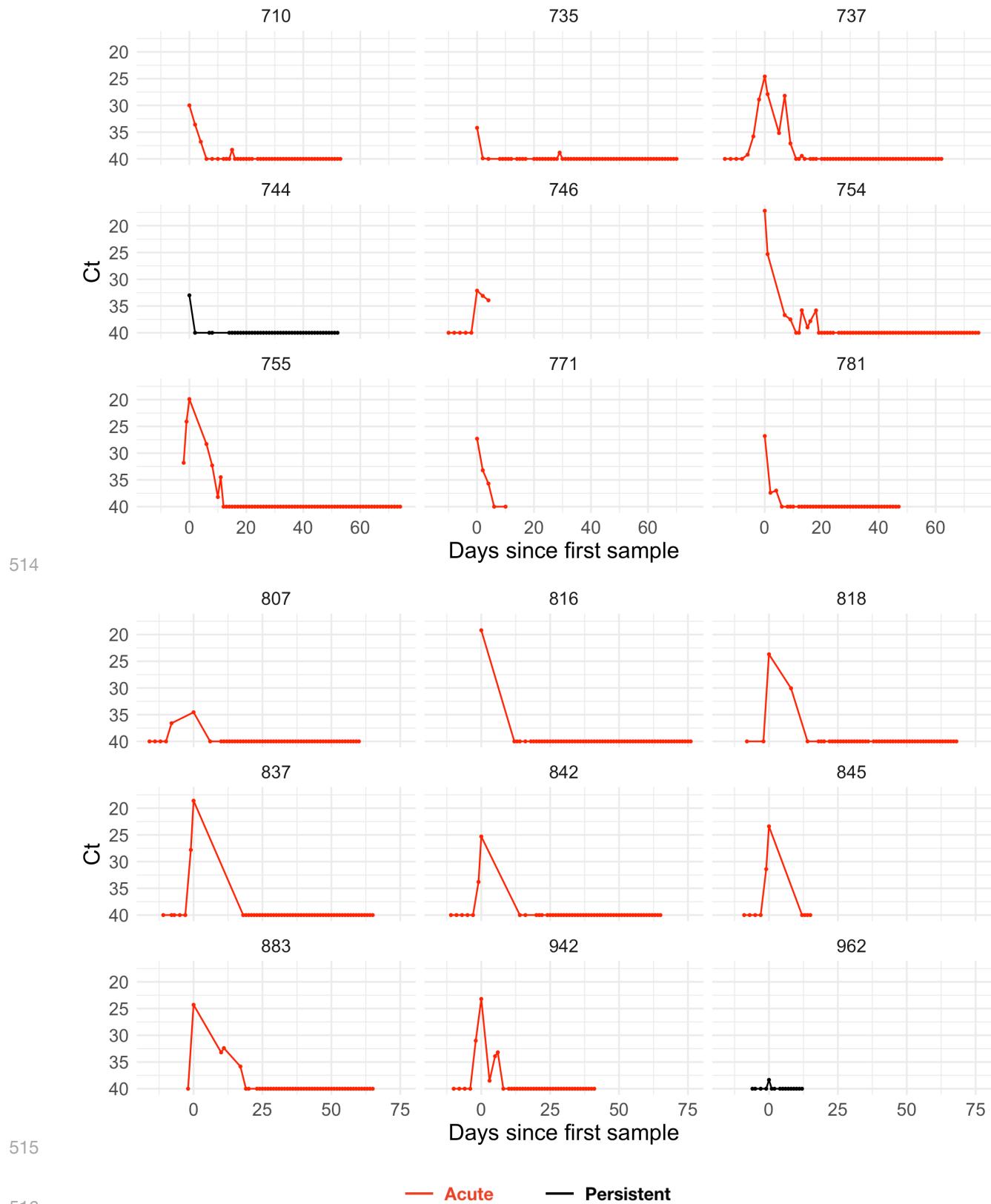
487 The values in square brackets denote truncation bounds for the distributions. We chose a vague  
488 half-Cauchy prior with scale 5 for the observation variance  $\sigma(\tilde{\sigma})$ . The priors for the population  
489 mean values ( $\mu$ ) are normally-distributed priors spanning the range of allowable values for that  
490 parameter; this prior is vague but expresses a mild preference for values near the center of the  
491 allowable range. The priors for the population standard deviations ( $\sigma$ ) are half Cauchy-distributed  
492 with scale chosen so that 90% of the distribution sits below the maximum value for that  
493 parameter; this prior is vague but expresses a mild preference for standard deviations close to  
494 0.

495

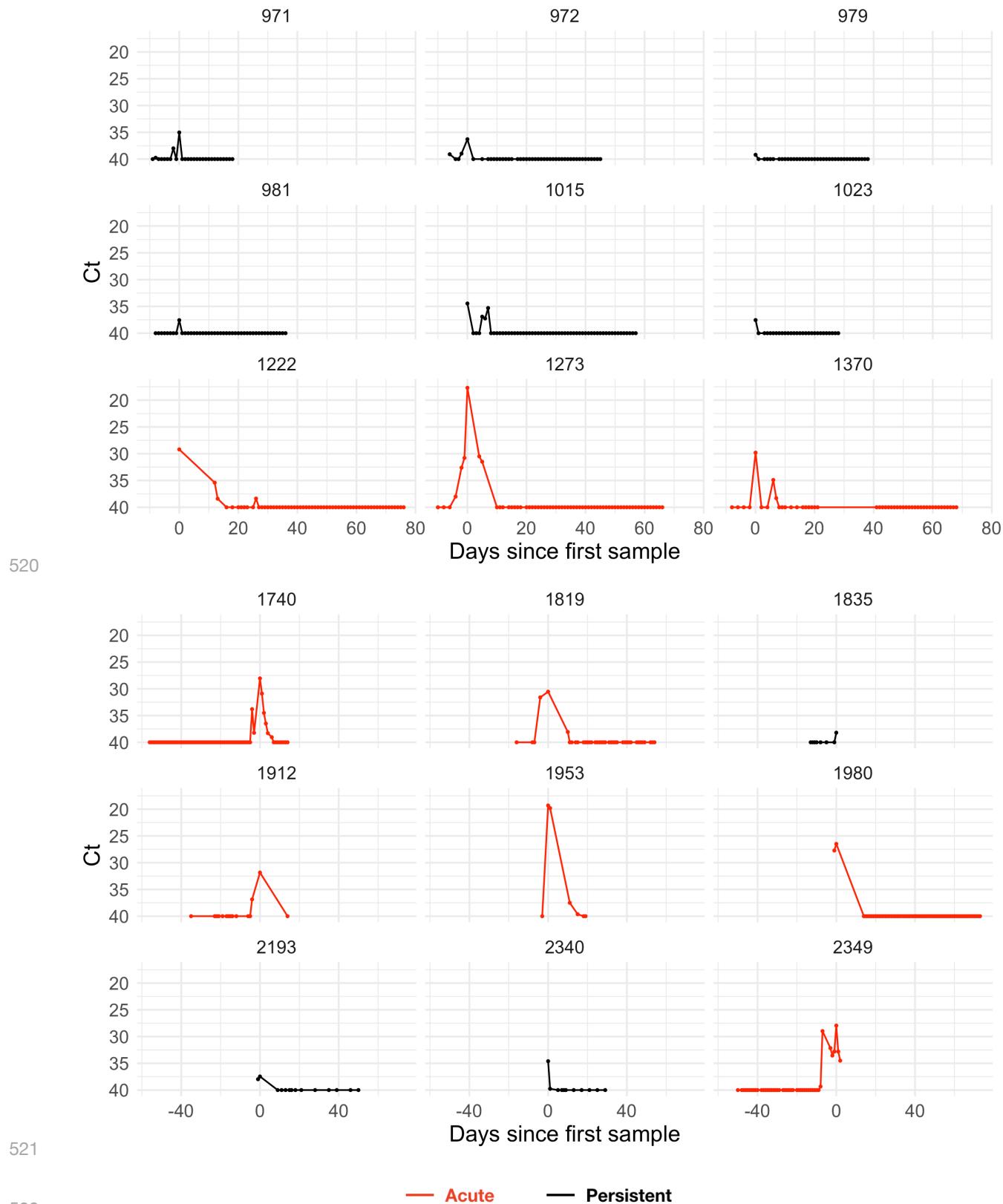
496 We ran four MCMC chains for 5,000 iterations each with a target average proposal acceptance  
497 probability of 0.99. The first half of each chain was discarded as the warm-up. The Gelman R-  
498 hat statistic was less than 1.1 for all parameters except for the  $t_p$  and  $\omega_r$  associated with individual  
499 1370, as the posterior distributions for those parameters were multi-modal (see **Supplemental**  
500 **Figures 12-13**). This indicates good overall mixing of the chains. There were fewer than 10  
501 divergent iterations (<0.1% of the transitions after warm-up), indicating good exploration of the  
502 parameter space.

503

504 The posterior distributions for  $\mu_\delta$ ,  $\mu_{\omega_p}$ , and  $\mu_{\omega_r}$ , estimated separately for symptomatic and  
505 asymptomatic individuals, are reported in **Figure 2** (main text). We fit a second model that did  
506 not distinguish between symptomatic and asymptomatic individuals. The posterior distributions  
507 for these same parameters under this model are depicted in **Supplemental Figure S9**. The  
508 posterior distributions for the individual-level parameters  $\omega_p$ ,  $\omega_r$ , and  $\delta$  are depicted in  
509 **Supplemental Figures S10-12**, with 500 sampled trajectories from these posterior distributions  
510 for each individual depicted in **Supplemental Figure 13**. The overall combined posterior  
511 distributions for the individual-level parameters  $\omega_p$ ,  $\omega_r$ , and  $\delta$  are depicted in **Supplemental**  
512 **Figure 14**. We estimated the best-fit normal (for  $\delta$ ) and gamma (for  $\omega_p$  and  $\omega_r$ ) distributions using  
513 the ‘fitdistrplus’ package implemented in R (version 3.6.2) (17).

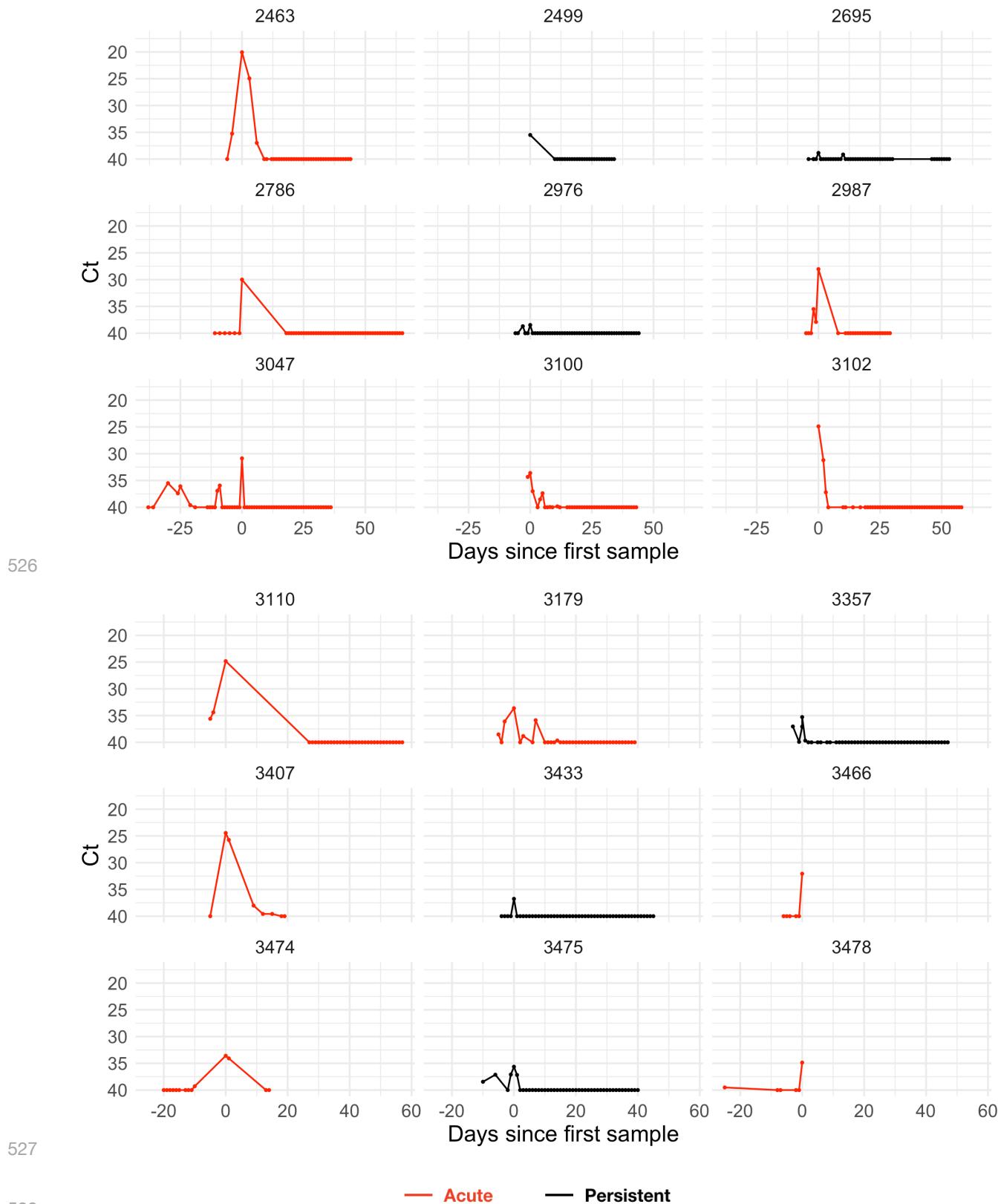


517 **Supplemental Figure 1. Observed Ct values from the study participants (1/4).** Points depict observed Ct values,  
518 which are connected with lines to better visualize trends. Individuals with presumed acute infections are marked in  
519 red. All others are in black.

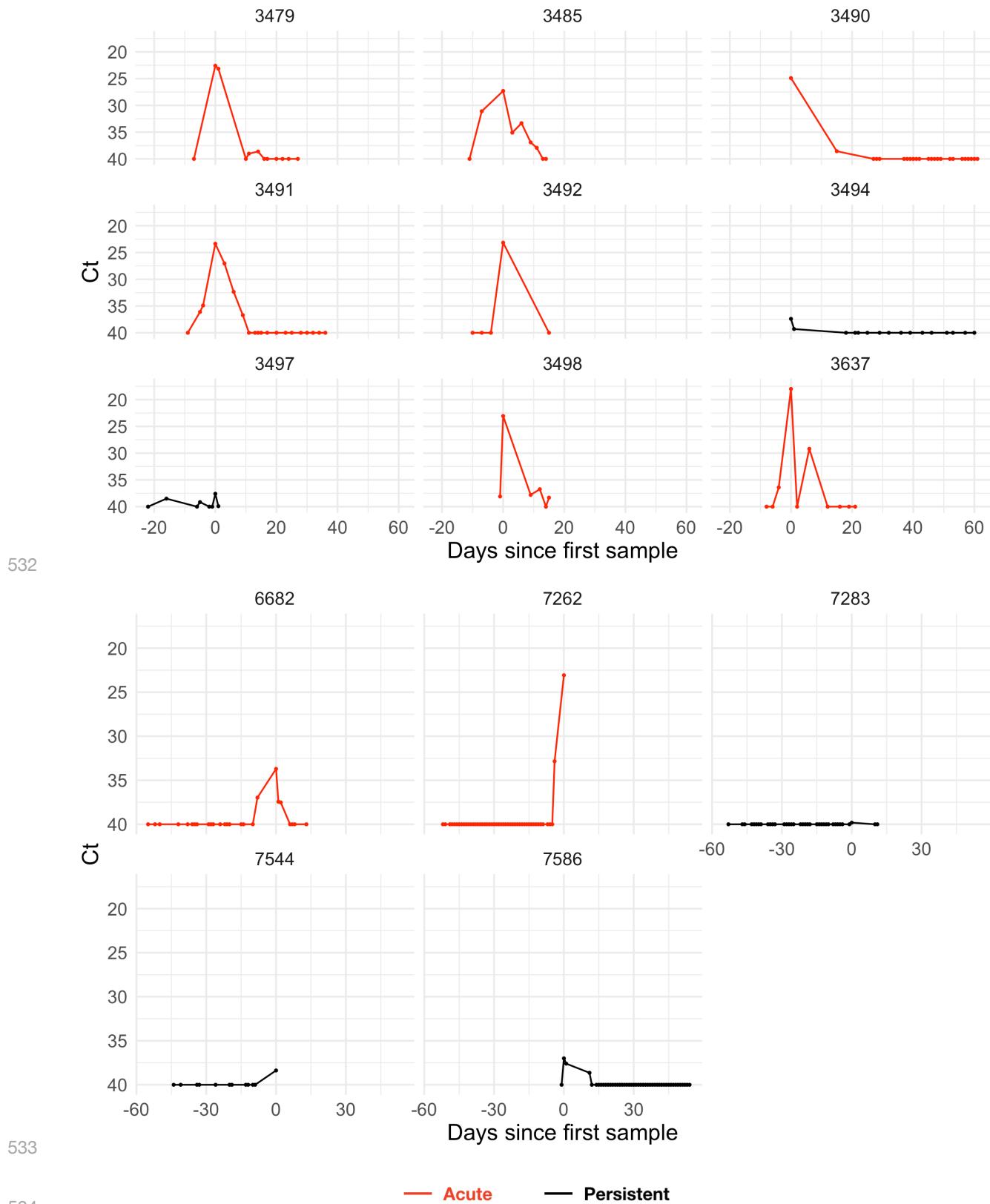


— Acute      — Persistent

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



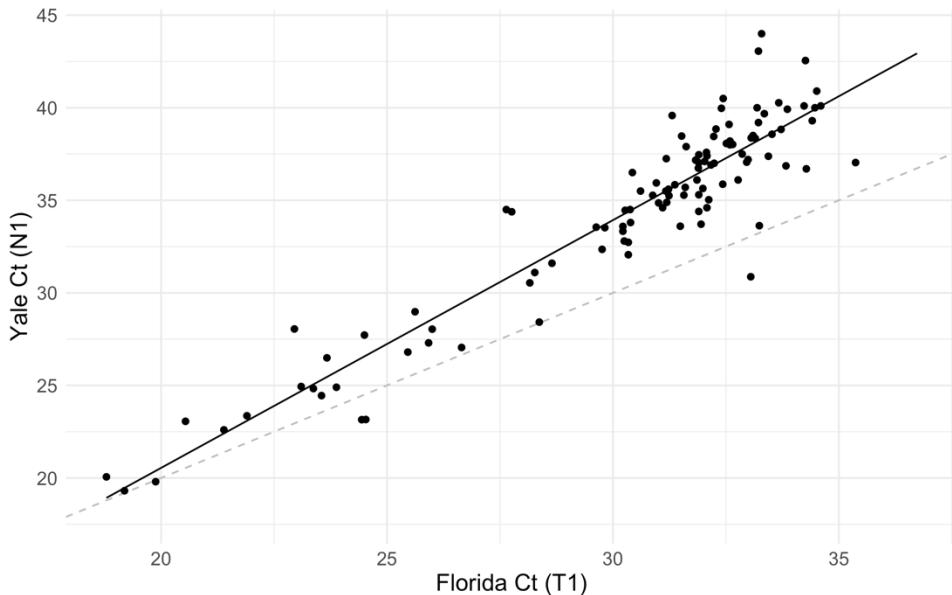
529 **Supplemental Figure 3. Observed Ct values from the study participants (3/4).** Points depict observed Ct values,  
530 which are connected with lines to better visualize trends. Individuals with presumed acute infections are marked in  
531 red. All others are in black.



— Acute      — Persistent

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

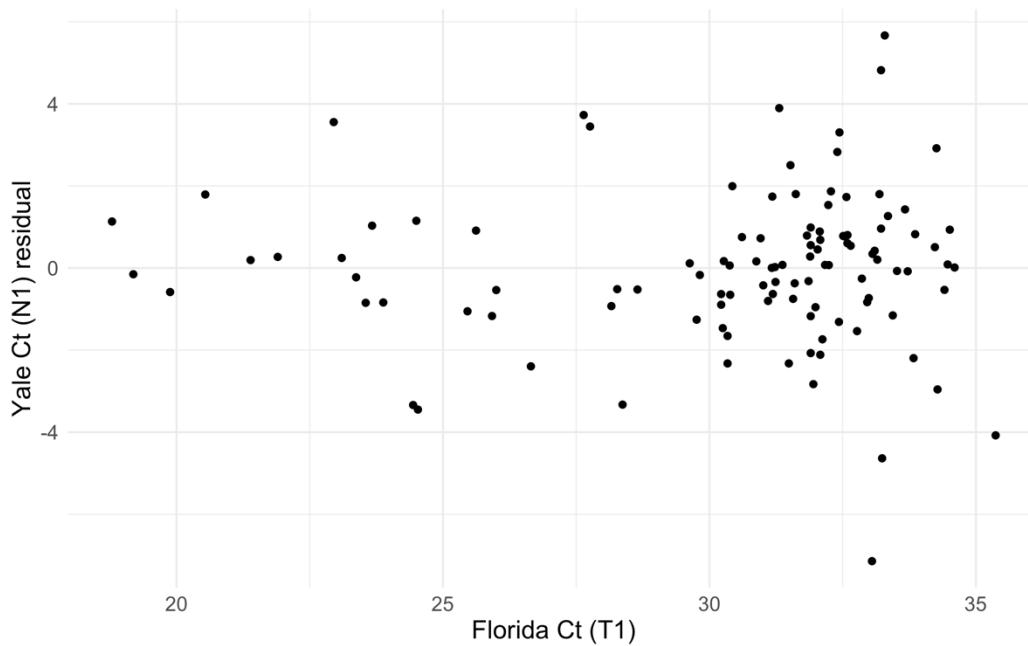
538



539

540 **Supplemental Figure 5. Ct values from the Yale and Florida labs.** Points depict the Ct values for SARS-CoV-2  
541 nasal swab samples that were tested in both Florida and Yale labs. Ct values from Florida represent Target 1 (ORF1ab)  
542 on the Roche cobas system, and Ct values from Yale represent N1 in the Yale multiplex assay. The solid black line  
543 depicts the best-fit linear regression (intercept = -6.25, slope = 1.34,  $R^2 = 0.86$ ). The dashed black line marks the 1-1  
544 line where the points would be expected to fall if the two labs were identical.

545

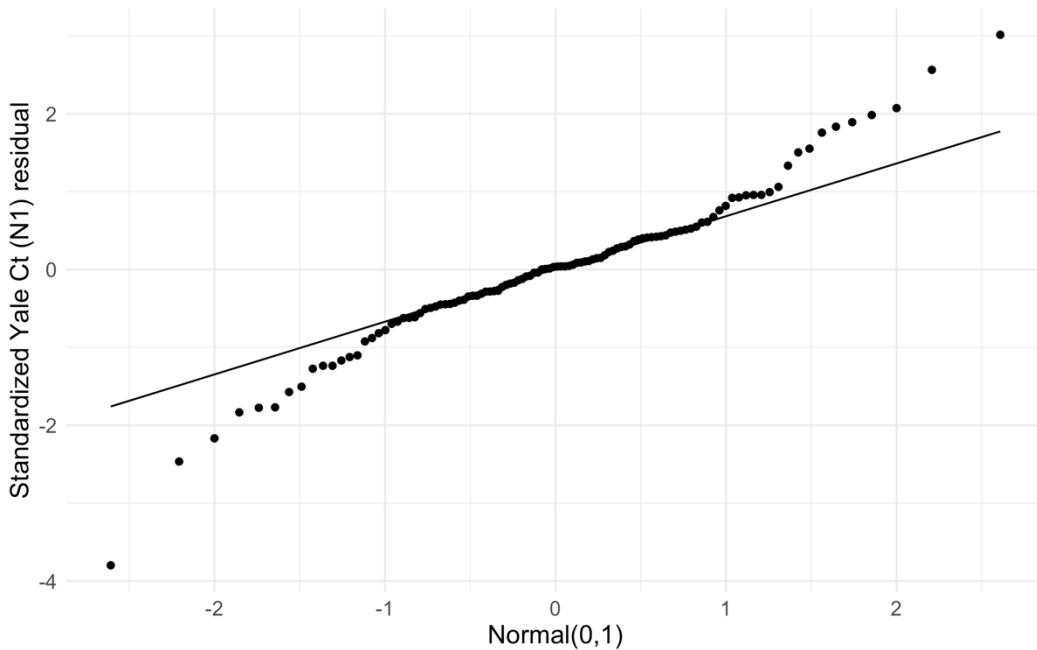


546

547 **Supplemental Figure 6. Residuals from the Yale/Florida Ct regression.** Points depict the residual after removing  
548 the best-fit linear trend in the relationship between the Yale and Florida Ct values.

549

550

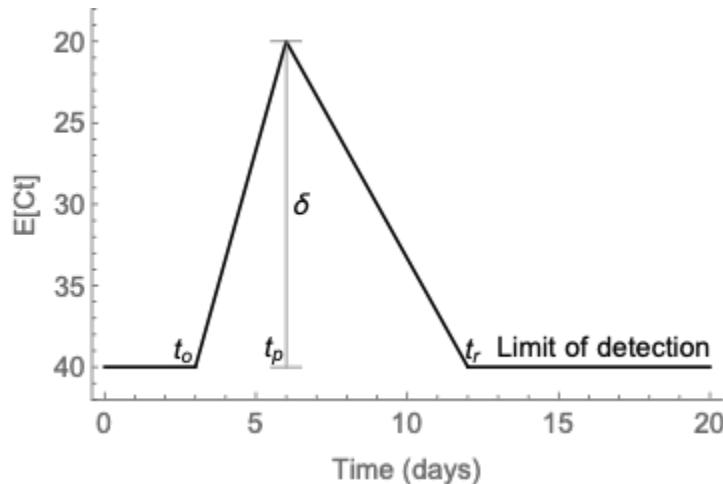


551

552  
553  
554  
555

**Supplemental Figure 7. QQ plot of the residuals from the Yale/Florida Ct regression.** The residuals were standardized (subtracted the mean and divided by the standard deviation) before comparing with the theoretical quantiles of a normal distribution with mean 0 and standard deviation 1. The points depict the empirical quantiles of the data points and the line depicts the where the points would be expected to fall if they were drawn from a standard normal distribution.

556



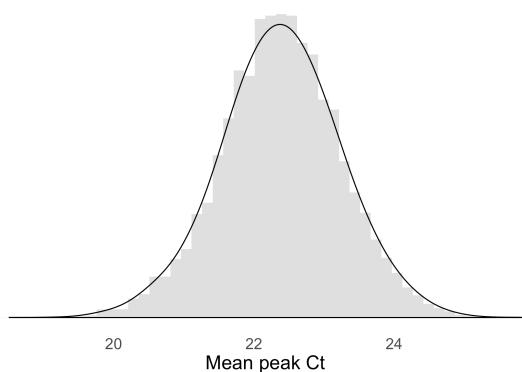
557

558 **Supplemental Figure 8. A theoretical Ct trajectory.**  $E[Ct]$  is the expected Ct value on a given day. The Ct begins at  
559 the limit of detection, then declines from the time of infection ( $t_o$ ) to the peak at  $\delta$  cycles below the limit of detection  
560 after  $t_p$  days. The Ct then rises again to the limit of detection after  $t_r$  days. The model incorporating these parameter  
561 values used to generate this piecewise curve is given in Equation S1 (**Supplemental Methods**).

562

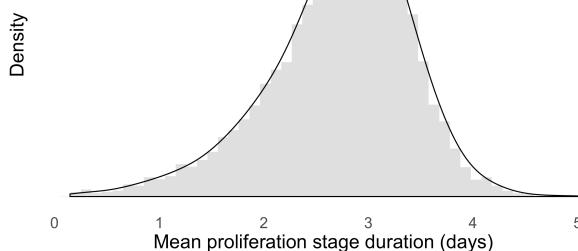
A)

Density



B)

Density

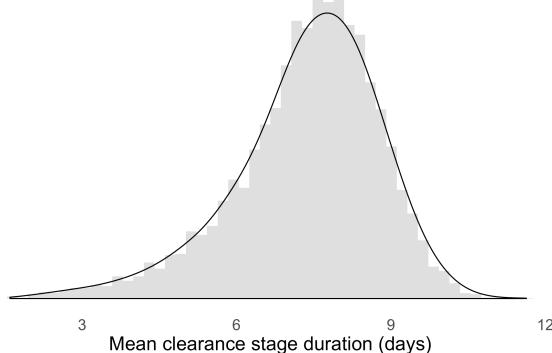


563

564

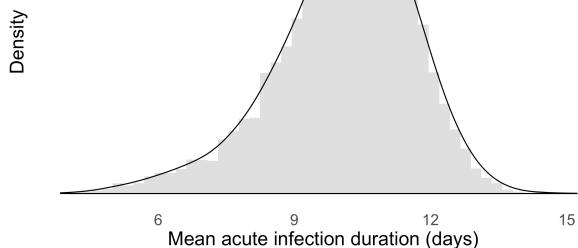
C)

Density



D)

Density



565

566

567

E)

20

25

30

35

40

45

50

55

60

65

70

75

80

85

90

-3

Days from peak

568

569

570

571

572

573

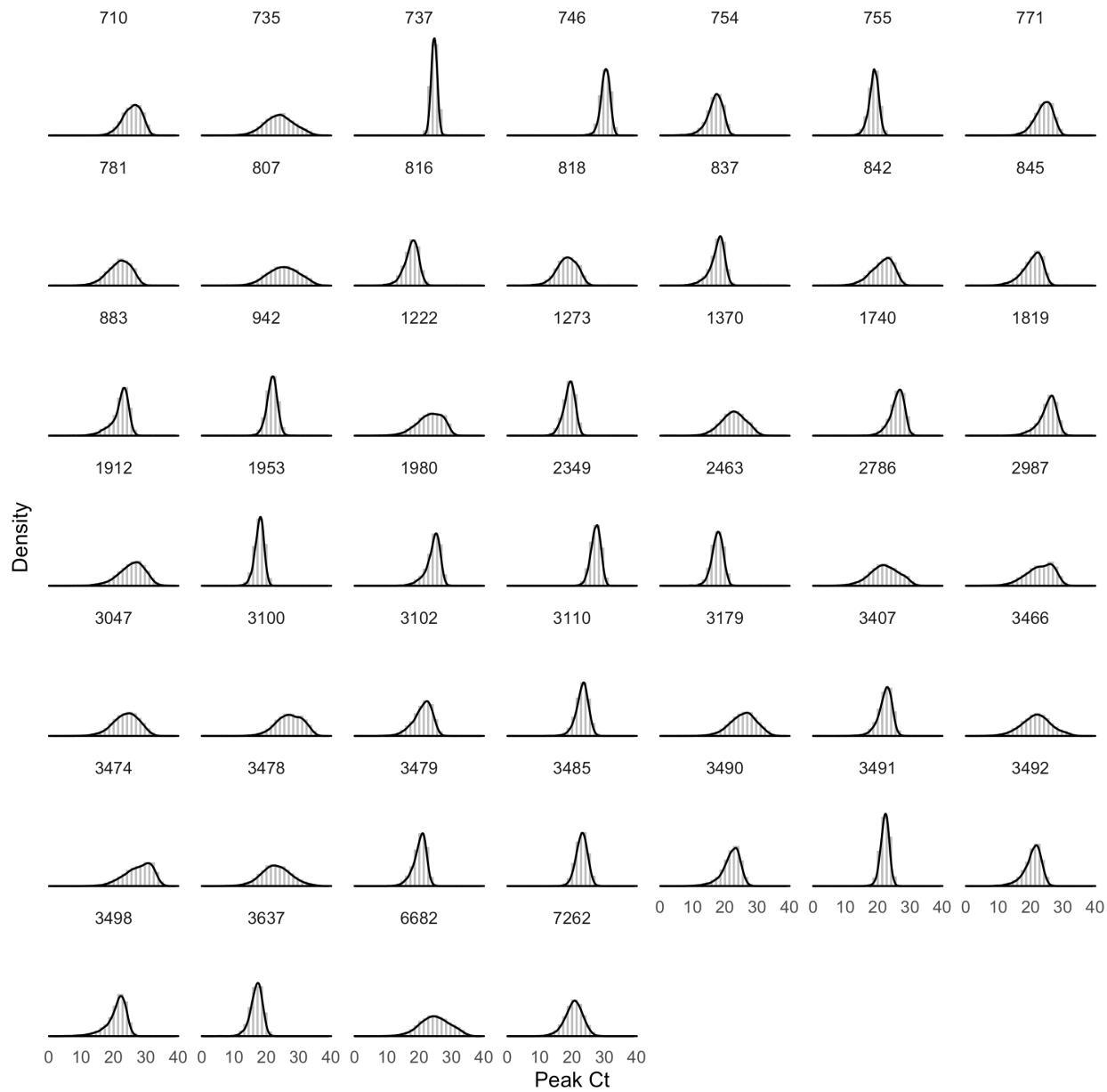
574

575

576

**Figure S9. Mean peak Ct and waiting time distributions for individuals with acute infections.** Histograms (colored bars) of 10,000 posterior draws from the distributions for peak Ct value (A), duration of the proliferation stage (infection detection to peak Ct, B), duration of the clearance stage (peak Ct to resolution of acute RNA shedding, C), and total duration of acute shedding (D) across the 46 individuals with a verified infection. The curves are kernel density estimators for the histograms to assist with visualizing the shapes of the histograms. The mean Ct trajectory corresponding to the mean values for peak Ct, proliferation duration, and clearance duration is depicted in (E) (solid lines), where shading depicts the 90% credible interval.

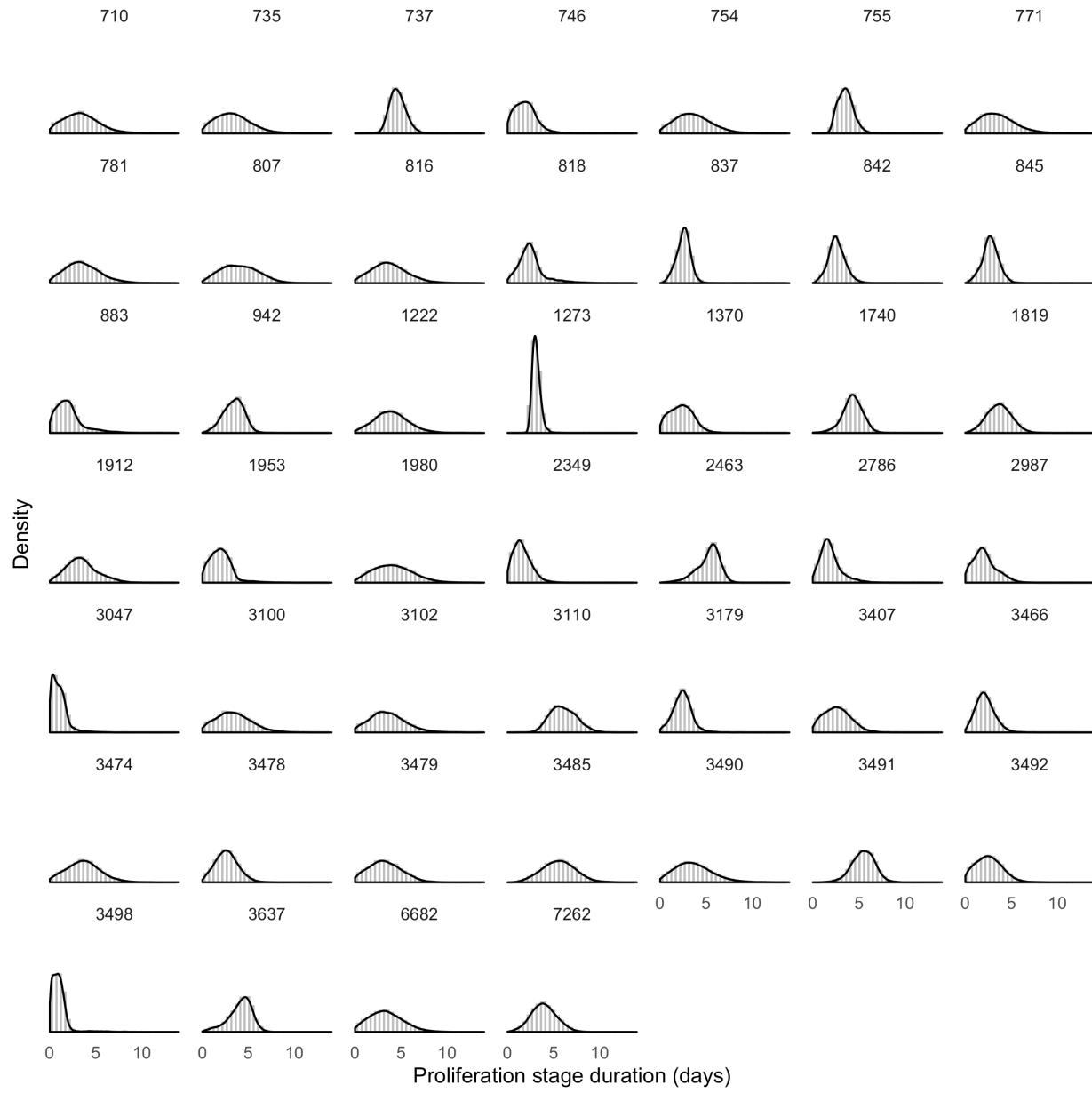
577



578

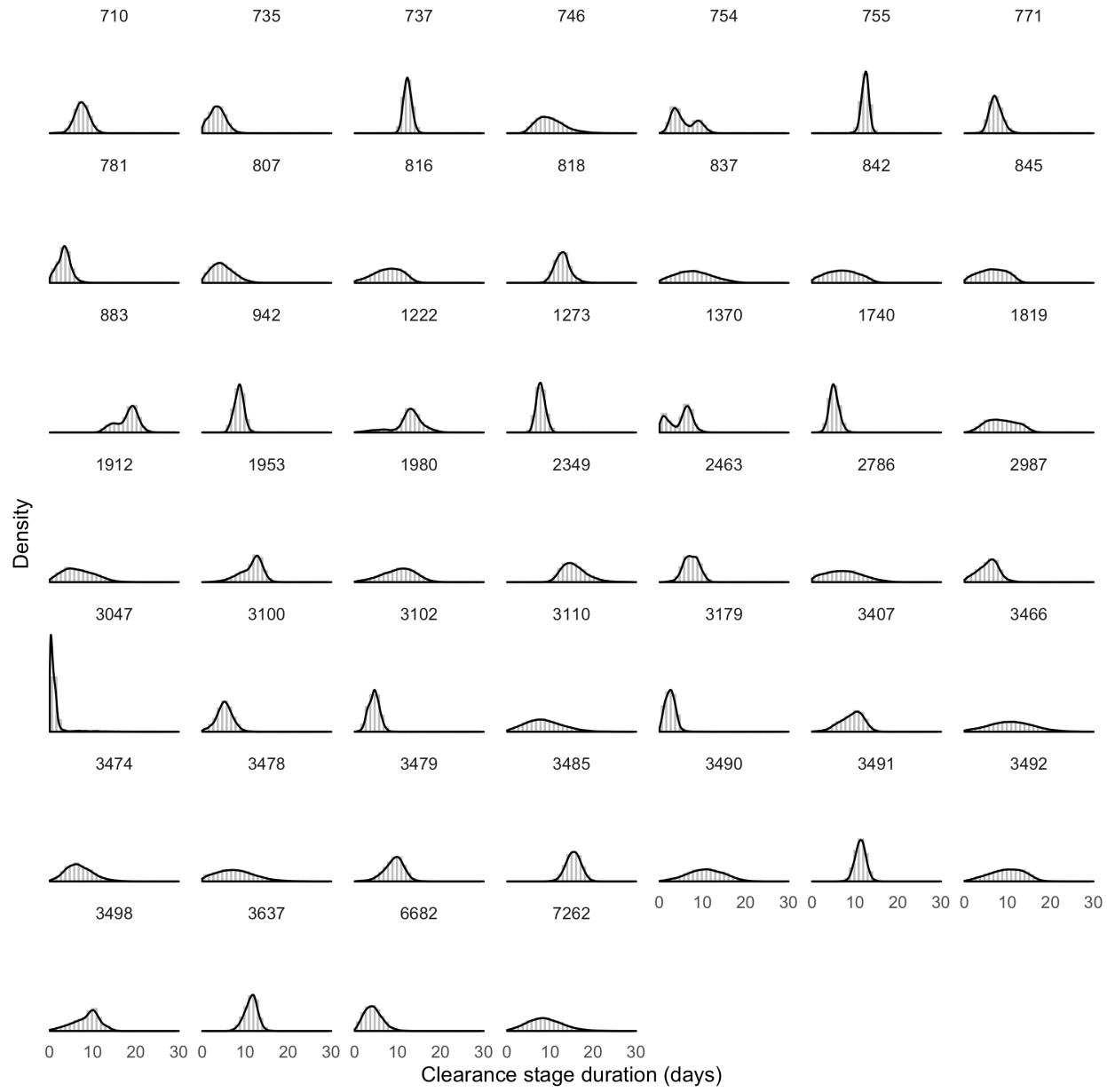
579

**Supplemental Figure 10. Posterior peak Ct distributions for the 46 individuals with acute infections.**



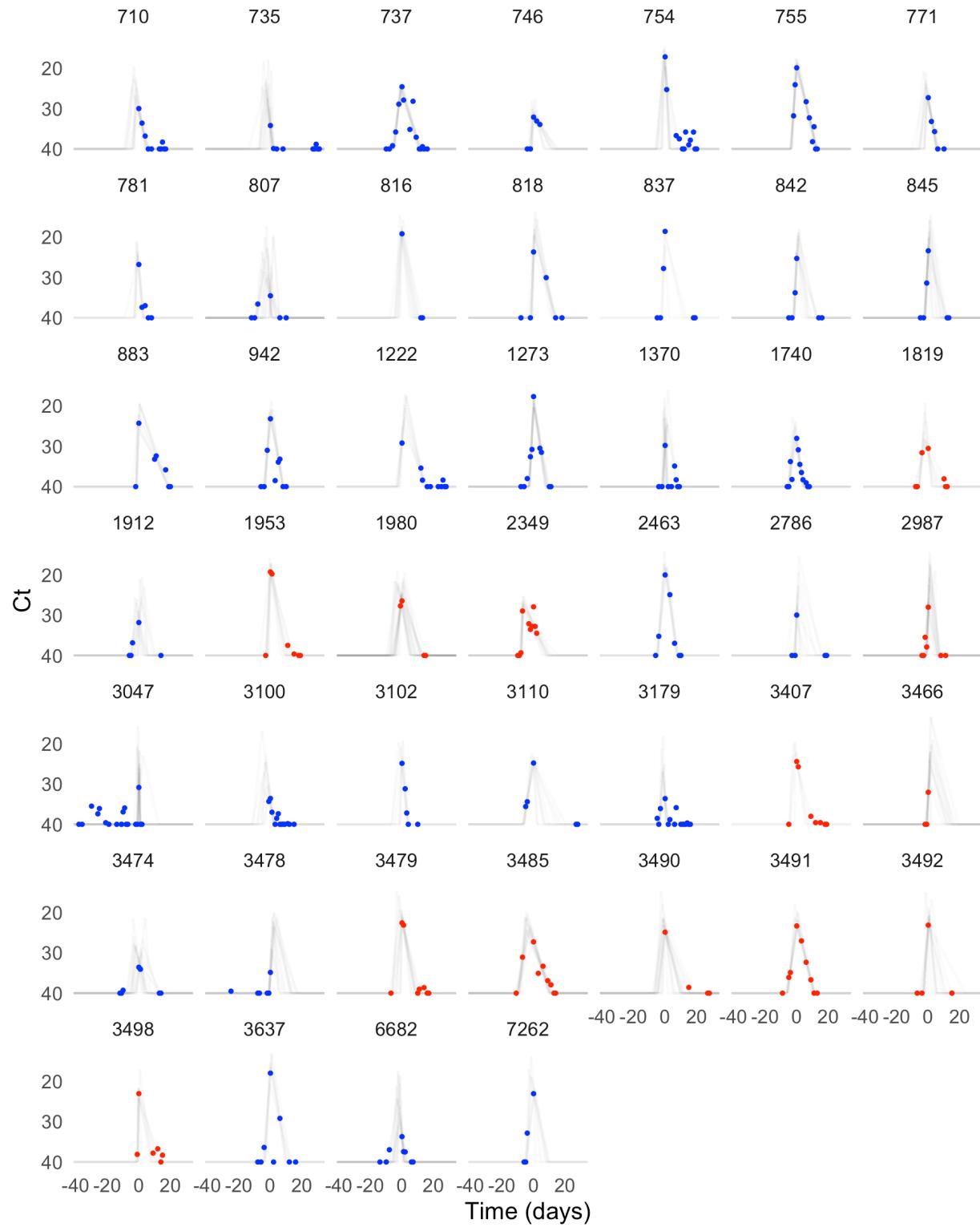
580  
581  
582

**Supplemental Figure 11. Posterior distributions for the duration of the proliferation stage for 46 individuals with acute infections.**



583  
584  
585

**Supplemental Figure 12. Posterior distributions for the clearance stage duration for 46 individuals with acute infections.**



586

● Symptoms reported   ● Symptoms not reported

587

588

589

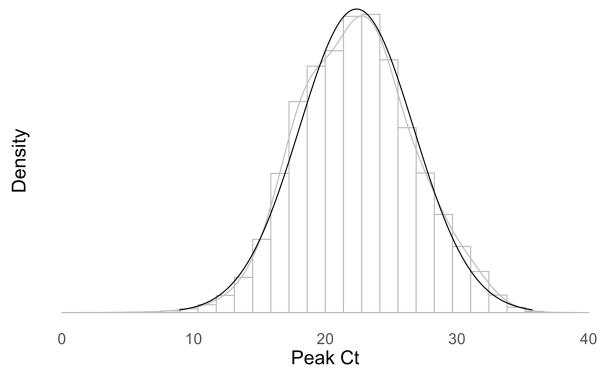
590

**Supplemental Figure 13. 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.

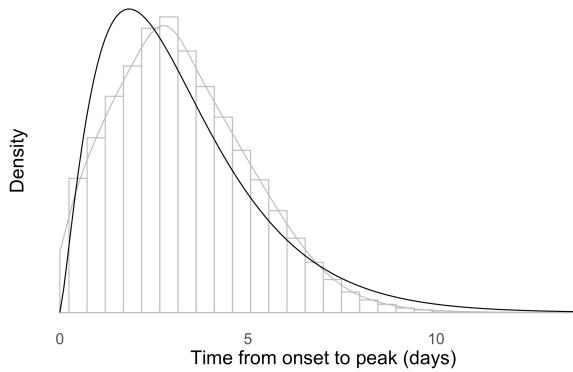
591

592

A)



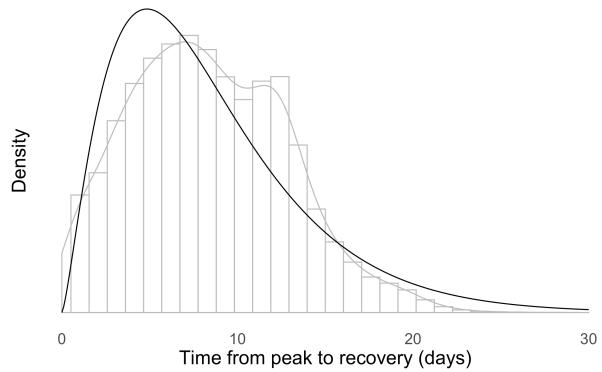
B)



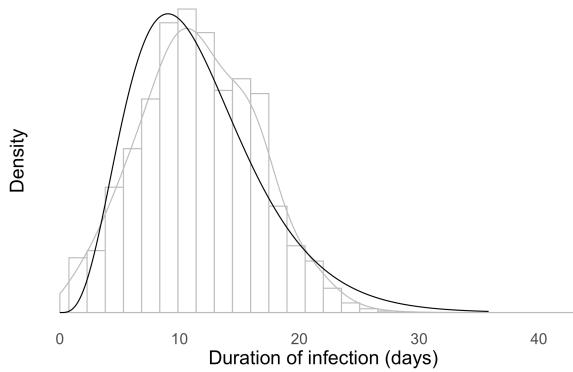
593

594

C)



D)



595

596

597 **Supplemental Figure 14. Individual-level peak Ct and waiting time distributions.** Histograms (grey bars) of 10,000  
598 posterior draws from the distributions for peak Ct value (A), time from onset to peak (B),  
599 time from peak to recovery (C), and total duration of infection (D) across the 46 individuals with an acute infection.  
600 Grey curves are kernel density estimators to more clearly exhibit the shape of the histogram.  
601 Black curves represent the best-fit normal (A) or gamma (B, C, D) distributions to the histograms.  
602 The duration of infection is the sum of the time from onset to peak and the  
603 time from peak to recovery. The best-fit normal distribution to the posterior peak Ct distribution had mean 22.3 and  
604 standard deviation 4.2; the best-fit gamma distribution to the proliferation stage duration had shape parameter 2.3  
605 and inverse scale parameter 0.7; the best-fit gamma distribution to the clearance stage duration had shape parameter  
606 2.4 and inverse scale parameter 0.3; and the best-fit gamma distribution to the total duration of infection had shape  
parameter 4.3 and inverse scale parameter 0.4.