

Bimodal Distribution Pattern Associated With The PCR Cycle Threshold (C_t) And Implications In COVID-19 Infections

Doris Yang

Oregon Health & Science University

Donna E. Hansel

Oregon Health & Science University

Marcel E. Curlin

Oregon Health & Science University

John M. Townes

Oregon Health & Science University

William B. Messer

Oregon Health & Science University

Guang Fan

Oregon Health & Science University

Xuan Qin (✉ qinxu@ohsu.edu)

Oregon Health & Science University

Research Article

Keywords:

Posted Date: March 7th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1340499/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

SARS-CoV-2 is notable for its extremely high level of viral replication in respiratory epithelial cells, relative to other cell types. This may partially explain the high transmissibility and rapid global dissemination observed during the COVID-19 pandemic. Polymerase chain reaction (PCR) cycle threshold (C_t) number has been widely used as a proxy for viral load based on the inverse relationship between C_t number and amplifiable genome copies present in a sample. We examined two PCR platforms (Centers for Disease Control and Prevention 2019-nCoV Real-time RT-PCR, Integrated DNA Technologies; and TaqPath COVID-19 multi-plex combination kit, ThermoFisher Scientific) for their performance characteristics and C_t distribution patterns based on results generated from 208,947 clinical samples obtained between October 2020 and September 2021. From 14,231 positive tests, C_t values ranged from 8 to 39 and displayed a pronounced bimodal distribution. The bimodal distribution persisted when stratified by gender, age, and time period of sample collection during which different viral variants circulated. This finding may be a result of heterogeneity in disease progression or host response to infection irrespective of age, gender, or viral variants. Quantification of respiratory mucosal viral load may provide additional insight into transmission and clinical indicators helpful for infection control.

Introduction

Since the onset of the COVID-19 pandemic in late 2019, SARS-CoV-2 has been distinguished by its unprecedented transmissibility compared to related coronaviruses such as SARS-CoV and MERS CoV. By March 11, 2020 when the World Health Organization declared COVID-19 a pandemic, there had already been over 118,000 cases in 114 countries and 4,291 deaths (WHO)¹; these tallies have now ballooned to cumulative totals of approximately 376 million cases and 5.6 million deaths worldwide. The amount of virus produced at the respiratory epithelium is considered to be a critical element in disease ^{2,3}, though not the only factor in determining SARS-CoV-2 transmissibility ⁴. Viral RNA load has also been investigated as a possible correlate of severity of illness ^{5,6}, host cell type or specific anatomic site of intense viral replication ⁷, viral replication dynamics during the course of clinical illness ⁸, and inoculum effect or infective dose. Polymerase chain reaction (PCR) cycle threshold (C_t) values have also been widely referenced as both epidemiological indicators and clinical indicators of disease burden and outcomes ^{9,10}.

A common technique for measuring SARS-CoV-2 viral load is through quantitative analysis of viral RNA genomic copy numbers. Quantitative viral RNA studies have generally treated the inverse polymerase chain reaction (PCR) cycle threshold (C_t) as a proxy for relative levels of viral genomes or viral load ⁶. However, many different PCR platforms have been developed for detection of SARS-CoV-2 and most were not originally intended to be fully quantitative. The degree of analytical variability associated with these assays can be minimized in the context of high throughput testing under a robust quality management system, and for this reason clinical laboratories involved in the pandemic response commonly validate multiple PCR platforms with correlation studies on a regular basis ¹¹.

In this study, we compared the performance characteristics of two leading PCR-based methods for quantification of SARS-CoV-2 targeting multiple viral genomic regions, and sought to characterize the C_t value distribution of SARS-CoV-2 RNA over a large sample size and date range in order to develop reporting strategies to improve effectiveness of infection prevention. We show that the PCR C_t value distribution was extremely wide and clearly departed from normality. The observed C_t distribution pattern was independent of patient age, gender, and time period of sample collection, during which a number of different variants were predominant.

Results

Test performance and assay dynamic range

Nucleic acid amplification was performed using either the 2019-nCoV CDC EUA Kit (IDT Integrated DNA Technologies, Inc., ie “CDC platform”) or the TaqPath™ Multiplex RT-PCR COVID-19 kit (Thermo Fisher Scientific, Inc., ie “Fisher platform”) (see “Methods”). We performed 208,947 PCR tests for SARS-COV-2 during 52 weeks between October 2020 and September 2021 using one of two platforms available (Table 1). In total, 14,231 (6.8%) positive tests were resulted and associated with 13,553 individuals. Patients contributing to the positive dataset ranged in age from 2 hours to 103 years; 48% were female. Most (97%) of the positive tests resulted from persons presenting for clinical care. Positive tests also included 185 (0.5%) of 37,908 pre-operation screening tests, and 589 (2.5%) of 23,262 healthcare staff tests. The prevailing variants of concern evolved over the time period of the study, as reflected by GISAID data summary for Oregon (Fig. 1). In addition, we included 878 test results between December 6, 2021 through January 22, 2022 that were positive but failed to amplify the S gene (S gene target failure or SGTF) with the Fisher PCR platform, in order to include likely omicron variants.

We performed an intra-assay C_t comparison between the N1 and N2 targets in the CDC platform, and the N and ORF1ab targets in the Fisher platform. Highly linear relationships between C_t values from two independent viral targets employed was observed in both platforms (Fig. 2) with coefficients of determination (R^2) of 0.983 and 0.919 respectively. Given the slightly higher correlation performance of the CDC PCR platform, we proceeded to an analysis of C_t distribution using samples tested with this method.

The viral load range observed in this analysis was very broad with C_t values ranging from < 8 to 39. To reduce signal noise for pattern analysis, we removed all extreme values beyond the expected linear range ($C_t \leq 9$, or ≥ 39), including 353 Fisher samples and 214 CDC samples. The resulting C_t span of 10 to 38 corresponds to absolute viral genomic copies ranging from approximately 1.5×10^{10} – 1.5×10^2 , where every 3 cycles constitutes roughly a one log viral titer change

(https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/926410/Understanding_Cycle_Threshold__Ct_in_SARS_CoV-2_RT-PCR_.pdf).

Analysis of bimodal distribution of C_t values

Initial plotting of the *Ct* values for both the N1 and N2 targets of the 5214 CDC PCR samples revealed a non-normal distribution (Fig. 3.), thus rendering traditional summary statistics ill-suited and insufficient for describing the distribution. In this analysis, Hartigan's Dip Test (<https://CRAN.R-project.org/package=diptest>) was used to confirm non-normal distribution or multimodality and the Bimodality Coefficient Test could however, specifically capture the observed bimodality. The p-value of Hartigan's Dip Test was $< 2.2e-16$ for both platforms regardless N1 and N2, or N and ORF1ab, in the 52-week period evaluation. This test alone confirmed the alternative hypothesis of a non-unimodal *Ct* number distribution for both platforms. The Bimodality Coefficient was 0.557 for the N1 target and 0.565 for the N2 target; both were above the critical value of 0.555, indicating a bimodal distribution for the CDC platform (Table 1). While *Ct* numbers generated from samples tested by Fisher platform also appeared to form two peaks, both N and ORF1ab targets failed Bimodality Coefficient tests (Table 1 and Fig. 3). The main difference between the two platforms was the PCR design by CDC single-plex versus Fisher multiplex. Therein, the examination of bimodality against viral or host factors was carried out using samples tested by CDC platform only.

The bimodal *Ct* distribution pattern was further examined by viral variant and demographic factors. To evaluate the impact of variants and vaccinations, we compared the *Ct* distributions separated by time frame: 2020 versus 2021 coinciding with the repeated emergence of SARS-CoV-2 variants of concern in 2021 (Table 1, Figs. 1 and 4). The bimodality pattern associated with either 2020- or 2021-time frame remained stable, passing both Hartigan's Dip Test and Bimodality Coefficient requirements with the exception of N1 falling below the benchmark of 0.555 in 2020 (Table 1). In a sub-analysis, we included 878 Fisher PCR results from December 6, 2021 to January 15, 2022 that showed the suspected omicron pattern of SGTF. Although only the first 109 samples of the 878 were confirmed omicron by full genomic sequencing at the time of manuscript writing, we again observed two peaks in *Ct* distribution for both N and ORF1ab which were confirmed by Hartigan's Dip Test, but not by the Bimodality Coefficient test (Table 1 and Suppl Fig. 1). Finally, we evaluated the impact of age and gender (Table 1 and Fig. 5–6, Supplemental Fig. 2). The only bimodal exceptions that completely failed Bimodality Coefficient test were observed when age brackets were set at < 12 and < 17 years, although neither were normal unimodal distribution by Hartigan's Dip Test (Table 1 and Suppl Fig. 2). In general, more sample groups generated by the CDC platform passed bimodality examinations by Hartigan's Dip Test as most of all the p-values remained below 0.05 (Table 1). Notably, the Bimodality Coefficient test appeared to be more stringent than Hartigan's Dip test, indicative of its specificity for bimodality confirmation.

Analysis of sample size that comprised of viral carriers under the first *Ct* peak

After establishing the bimodality of the distribution, we quantitatively measured the two peaks by fitting a mixture of two normal distributions. This yielded a mean of ~ 20 for the first *Ct* peak and a mean of ~ 33 for the second *Ct* peak (Table 2). The mixing proportions, or lambdas, show that approximately 30% of samples fall under the first peak. Taking the upper bounds of 0.5 or 1.0 standard deviation range of the first peak, approximately 47% of the samples fall under the *Ct* of 22 and 56% fall under the *Ct* of 24, respectively (Table 2). It was possible that our fitting method may have introduced a bias to the peaks as the first peak appeared to be right shifted and the second peak to be left shifted (Fig. 3).

Discussion

In this analysis, two mainstream platforms for amplification and quantification of SARS-CoV-2 displayed similar analytic characteristics. Both the CDC and the Fisher platforms produced highly linear *Ct* correlations with coefficient of determination close to 1 between their corresponding two viral targets (N1 and N2, or N and ORF1ab) used. Moreover, this analysis of 14,231 individual positive test results confirmed previous findings showing a wide spectrum of PCR *Ct* distribution in nasopharyngeal swab samples (*Ct* range 8–39, Fig. 1). The corresponding viral titers can range from few copies to billions of copies. When *Ct* distribution patterns were examined, samples appeared to form two peaks along the *Ct* gradients produced by either of the platforms (Fig. 3). However, the separation of the two peaks was less pronounced with the Fisher platform, presumably influenced by the multiplex chemistry owing to the self-inhibition among sets of primers and reduced amplification efficiency 12.

Notably, the distribution of *Ct* numbers observed in our series was bimodal (Table 1 and Fig. 3–5), suggesting contribution from two distinct subsets of samples. This effect is likely not an artifact of sample quality or preparation. The potential contributory factor pertaining to sample quality variation to the bimodality *Ct* distribution is ruled out as CDC platform included host *RNase P* as an internal control 13. Previous studies have demonstrated that the viral titer can be associated with inoculum size, tropism or replication in specific tissue or cell types, and risk of onward transmission 7. On the other hand, viral titers are not correlated with age or disease severity 14-15. Importantly, high levels of viral shedding may occur in asymptomatic hosts, posing substantial challenges to infection control efforts 16-17. However, there is currently little published information on COVID-19 *Ct* value distribution patterns or their significance to virus-host interactions in SARS-CoV-2 infection. A few studies that did note *Ct* distribution properties outside of normality did not analyze its significance in microbial and host relations distinctively associated with SARS-CoV-2 8-18. We explored whether the pattern of viral levels at the population level could provide insight into the nature of SARS-CoV-2 shedding difference potentially useful for infection prevention.

When the *Ct* distribution pattern was examined by age groups, the heterogeneous non-unimodal distribution was evident. For age group of < 5 , < 21 (by N2 only), 21–64, and 65+ years, their *Ct* distributions have met the bimodality coefficient criteria (Table 1). However, the non-bimodal nor unimodal *Ct* distribution pattern associated with the age groups of < 12 and < 17 years remains puzzling, when that of age group < 5 years was clearly bimodal. This result suggests there may be underlying differences between viral replication in very young patient's vs teen's. Otherwise, the bimodal nature of the *Ct* distribution was unaffected by gender, or calendar time-period, during which several different variants predominated. Notably, the *Ct* distribution of the 878 omicron samples appeared to show the two positive peaks skew closer into each other (Supplemental Fig. 1). The putative Omicron *Ct* distribution curves failed Bimodality Coefficient test. It is again possible that the Fisher multiplex chemistry suppressed the expression of bimodality as seen in the overall 52-week analysis (Table 1). However, at this point we believe there is still sufficient evidence to support the finding of this dichotomous distribution of viral replication pattern in the host population. More studies using other test platforms are needed to confirm this finding.

Host factors must play a role in heterogeneous viral replication properties. SARS-CoV-2 cell entry is mediated by human angiotensin-converting enzyme II (*ACE2*) and *ACE2* polymorphisms, which may affect the risk for SARS-CoV-2 infection and the course of COVID-19. In a multivariable analysis by Nikiforuk *et al.*, the researchers showed that the greatest viral RNA loads were observed in participants with high transmembrane *ACE2* transcription, while transcription of the soluble isoform appears to protect against high viral RNA load in the upper respiratory tract. It is possible that multiple host genetic factors, innate and adaptive immunity, and respiratory microbiota may all play roles in viral titers and disease outcomes.

Our data support the notion that high viral load carriers may contribute most to new transmissions in the community. An operational categorization separating high from low/moderate viral shedding could therefore be relevant to isolation requirements after infection, and infection control efforts. Using a cutoff value of $Ct < 22-24$, corresponding to the upper bound of 0.5 SD – 1 SD of the first peak, representing 47% – 56% of individuals in this cohort, could be used as indicators separating levels of respiratory tract viral shedding potentials. Ideally any categorization would be tested against presence of culturable virus and risk of transmission in clinical studies.

There are several limitations of this analysis. This study did not include host information on clinical course, vaccination status, or immune responses at the time samples were collected. We are therefore unable to explicitly relate *Ct* values with these clinical factors. In addition, we do not know the identities of viral strain or variant associated with most of the *Ct* values obtained, and this information would be helpful in formally evaluating the role of infecting variant on *Ct* values. With the data set size, it is likely that the samples collectively represent a random distribution along the clinical course of the viral infection. This dichotomous *Ct* distribution can be deployed in test reports as a relatively simple indicator that can be useful for the management of infected patients.

It has been well acknowledged that the PCR *Ct* values or their associated viral titers do not correlate well with the intensity of symptoms during SARS-CoV-2 infection, nor are they predictive of disease outcome; thus, they are currently not routinely used in clinical management. (<https://www.aphl.org/programs/preparedness/Crisis-Management/Documents/APHL-COVID19-Ct-Values.pdf> and <https://www.idsociety.org/globalassets/idsa/public-health/covid-19/idsa-amp-statement.pdf>). However, it has been suggested that those presenting with higher *Ct* values may require shorter periods of isolation to prevent onward transmission. Since the implications of super-spreading events are well known, we thus cautiously propose that in samples with a very low *Ct* finding ($Ct < 22-24$ or viral titers predicated to be in millions), a notation in test report could be considered. Clinical recommendations surrounding clinical standard actions such as isolation and period of quarantine with repeat test assurances for deisolation have to be developed before a reporting notation can be implemented. With the emergence of iterations highly transmissible SARS-CoV-2 variants of concern, it is prudent to empower existing testing and reporting strategies to reduce ongoing community transmission in order to control the case growth rate, healthcare burden, and workforce preservation.

Materials And Methods

PCR Cycle threshold (*Ct*) data on SARS-CoV-2 positive specimens.

We performed 208,947 tests between October 4, 2020 and September 30, 2021. The inclusion of additional 878 suspected Omicron samples was later included by taking the advantage of Fisher PCR platform using S gene target failure (SGTF) as a surrogate marker after September 30, in late 2021. All samples were collected by qualified healthcare professionals. The specimen type included primarily nasopharyngeal (NP) swabs (> 99.9%), and a small number of laboratory-validated bronchoalveolar lavage samples, tracheal aspirates, nasal swabs and throat swab specimens. All samples with positive PCR results were included in the positive dataset ($n = 14,231$), and included repeat testing in some individuals. Samples with “negative” or “inconclusive” PCR results were excluded from the dataset. Excel Microsoft 360 and Tableau 2021.1 were used for data analysis and visualization.

This study was approved by the Oregon Health & Science University Institutional Review Board (STUDY00021396: Collection and archiving of nasopharyngeal swab, sputum, and blood samples from persons tested for SARS-CoV-2 infection). This study protocol was exempt from informed consent as it does not allow any patient identifiable information to be used for data analysis or presentation. All methods used for the analysis were in accordance with the guidelines put forth by the STUDY00021396.

Viral titers inferred by PCR cycle threshold analysis

RNA extraction was performed by two different methods including MagNA Pure 96 and KingFisher Flex (ThermoFisher Scientific) for this study period, according to manufacturer instructions. To expand test capacity after September, 2020, RNA samples were tested by two PCR platforms: the 2019-nCoV CDC EUA Kit containing N1, N2 and human *RNase P* (RP) primer/probe mix (IDT Integrated DNA Technologies, Inc.) and the TaqPath™ Multiplex RT-PCR COVID-19 kit containing N gene, S gene, ORF1ab primers, and MS2 Phage control (ThermoFisher Scientific). Inter-analyzer correlation studies with regard to RNA extraction and PCR platforms were carried out at the time of initial test validation of RNA, then once every 6 months regularly. Each sample was tested by one or the other platform, but not both. All PCR amplifications were performed using QuantStudio5 thermocyclers (ThermoFisher Scientific) at the limit of detection ~ 15 copies per PCR reaction for all mixed use of RNA extraction and PCR platforms. For data clarity, all PCR *Ct* records in this study were rounded to 0.5 from their original records containing two decimal points. The corresponding *Ct* distributions of each test were measured by linearity and R-squared goodness-fit test.

A standard curve relating *Ct* values to viral copies ranging from 15–500,000 copies per PCR reaction was developed using serial dilutions of a synthetic SARS-CoV-2 RNA control provided by the CDC (data not shown). This confirms that a 3-point change in *Ct* value is roughly equal to a 10-fold change in the quantity of the template viral material

(https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/926410/Understanding_Cycle_Threshold__Ct__in_SARS

CoV-2_RT-PCR_.pdf). Similar to other studies 627, at viral titers near the limit-of-detection (LOD) using 10 PCR reaction replicates, the *Ct* values are often spread around 33–38, indicative of analytical stochasticity and loss of linearity (data not shown).

Statistical analysis

We compared performance characteristics between the two PCR test methods using corresponding obtained *Ct* values. We compared 5212 *Ct* values produced by CDC PCR platform to 8460 *Ct* values produced by Fisher PCR platform. The data collection period included several months when the viral variant B.1.1.7 was highly prevalent. B.1.1.7 samples were observed to be associated with S gene target failure when tested using the Fisher PCR platform, and therefore, *Ct* values generated on S gene amplification by multiplex PCR chemistry were not included in analysis. We did not have access to age data for 97 individuals and gender data for 356 individuals, and therefore these samples were excluded from any relevant analyses.

The reliability of the CDC versus Fisher platforms was compared through a linear regression analysis using R.4.1.0 (R Core Team, 2021, <https://www.R-project.org/>). The bimodality of the CT value distribution was evaluated in two ways: (1) Bimodality Coefficient 28 and (2) p-value for Hartigan's Dip Test (<https://CRAN.R-project.org/package=diptest>). To calculate the Bimodality Coefficient, skewness (m_3) and kurtosis (m_4) were first found using the *e1071* R package 29 (v1.7-7; <https://CRAN.R-project.org/package=e1071>). These values, along with sample size (*n*) were substituted into the formula

$$BC = \frac{m_3^2 + 1}{m_4 + 3 \cdot \frac{(n-1)^2}{(n-2)(n-3)}}$$

28. The p-value for Hartigan's Dip Test was calculated using the *diptest* (v0.76-0; Maechler, 2021) R package. When p-value indicates a statistical significance (e.g. $p < 0.05$) by Hartigan's test, the distribution is anything but unimodal. Normal mixture modelling was done using the *mclust* (v5.4.9, 2021 R package) 30. R, Microsoft Excel, and Tableau were also used as tools for visualization.

Table 1

Bimodality Coefficients and Hartigan's Dip Test p-values of PCR *Ct* values by age and gender of patients and date of collection (values shaded in grey do not meet the threshold for Bimodality Coefficient or multimodality by Hartigan's Dip Test). *S gene target failure by Fisher multiplex PCR chemistry.

		Hartigan's Dip Test		Bimodality Coefficient		
		p-value				
		N1	N2	N1	N2	
CDC PCR Chemistry	All CDC (n=5212)	<2.20E-16	<2.20E-16	0.5574	0.5725	
	Oct 2020-Sep 2021					
	Age	65+ (n=350)	0.006868	0.002952	0.5916	0.6079
	(Oct 2020-Sep 2021)	21-64 (n=3739)	<2.20E-16	<2.20E-16	0.5631	0.5756
		<21 (n=1028)	9.39E-06	1.35E-05	0.5403	0.5602
		<17 (n=655)	0.000314	0.000982	0.5213	0.5402
		<12 (n=376)	0.0386	0.02934	0.5238	0.5392
		<5 (n=100)	0.05607	0.01073	0.5560	0.5783
	Gender	Female (n=2338)	0.000331	0.00036	0.5557	0.5716
	(Oct 2020-Sep 2021)	Male (n=2520)	<2.20E-16	<2.20E-16	0.5695	0.5842
Date Range	Oct-Dec 2020 (n=3149)	2.07E-05	4.25E-05	0.5497	0.5687	
	Jan-Sep 2021 (n=2063)	4.94E-05	1.01E-07	0.5719	0.5782	
Fisher PCR Chemistry		N	ORF1ab	N	ORF1ab	
		Omicron (n=878) *	0.009973	0.0201	0.4851	0.5156
	All Fisher (n=8460)	<2.20E-16	<2.20E-16	0.5067	0.5267	
	Oct 2020-Sep 2021					

Table 2: Statistics for fitting of two-component normal mixture distribution for PCR *Ct* values by CDC Chemistry.

PCR target		Peak 1 (% of samples, Ct < upper bound of SD)	Peak 2
N1	Mixing Proportion	0.6791	0.3209
	Mean	20.06	32.28
	SD	4.18	3.16
	0.5 SD Range	17.97-22.15 (47%, Ct <22)	30.70-33.86
	1 SD Range	15.87-24.24 (56%, Ct <24)	29.12-35.44
	2 SD Range	11.69-28.42 (70%, Ct <28)	25.95-38.61
N2	Mixing Proportion	0.6718	0.3282
	Mean	19.91	33.00
	SD	4.36	3.37
	0.5 SD Range	17.73-22.09 (47%, Ct <22)	31.32-34.68
	1 SD Range	15.55-24.27 (55%, Ct <24)	29.64-36.37
	2 SD Range	11.18-28.64 (70%, Ct <29)	26.27-39.73

Figure 1. SARS-CoV-2 Variant Percentages in Oregon, September 2020 to January 2022.

Figure 2. Correlation of PCR Ct distribution between two viral targets associated with CDC or Fisher chemistry.

Figure 3. PCR Ct distribution of N1 and N2 targets by CDC chemistry and N and ORF1ab targets by Fisher chemistry.

Figure 4. PCR Ct distribution of N1 and N2 targets from samples collected in 2020 and 2021.

Figure 5. PCR Ct distribution of N1 and N2 associated with age groups: (a) age < 5 years, (b) age groups of 21–64 years and > 65 years.

Figure 6. PCR Ct distribution of N1 and N2 targets from samples associated with female and male patients.

Supplemental Fig. 1. PCR Ct distribution of N and ORF1ab targets associated with samples predicated to be omicron with SGTF by Fisher PCR multiplex chemistry.

Supplemental Fig. 2. (a) PCR Ct distribution of N1 and N2 targets by CDC chemistry collected from patients of age less than 12 years (b) PCR Ct distribution of N1 and N2 targets by CDC chemistry collected from patients of age less than 17 years (c) PCR Ct distribution of N1 and N2 targets by CDC chemistry collected from patients of age less than 21 years.

Declarations

Acknowledgements: SARS-CoV-2 genome sequencing was funded in part by the Oregon Public Health Authority (OHA Agreement Number 170015). We thank OHSU clinical laboratory scientists for their high-quality performance and documentation of SARS-CoV-2 test results. We thank the research group of Dr. Brian O’Roak for its sustained contributions to OHSU’s COVID-19 sequencing operations. Doris Yang was an undergraduate student attending Harvard College while working as a student internist on Weekly Oregon State Genome Dashboard for OHSU.

Author contributions: XQ collected data and conceptualized the study. DYY contributed to data analysis by identifying and executing appropriate statistical formula to address questions raised by the PCR Ct data distribution patterns. DYY and XQ contributed to initial manuscript drafting and editing. GF and DEH established the high throughput COVID diagnostic PCR operations at OHSU and manuscript editing. DEH has led and maintained a biorepository IRB which covered the research activities associated with this study. MEC, JMT, and WBM all have invested their valuable thoughts and time on iterations of revisions of the manuscript.

Data availability statement: Supplemental information 1 and Supplemental information 2.

Ethics Oversight: This study was approved by the Oregon Health & Science University Institutional Review Board (STUDY00021396: Collection and archiving of nasopharyngeal swab, sputum, and blood samples from persons tested for SARS-CoV-2 infection).

Competing Interests: The authors declare no competing interests.

References

1. Lewis, D. Superspreading drives the COVID pandemic - and could help to tame it. *Nature* **590**, 544–546, doi:10.1038/d41586-021-00460-x (2021).
2. Marks, M. *et al.* Transmission of COVID-19 in 282 clusters in Catalonia, Spain: a cohort study. *Lancet Infect Dis* **21**, 629–636, doi:10.1016/S1473-3099(20)30985-3 (2021).
3. Spinelli, M. A., Rutherford, G. & Gandhi, M. Lowering SARS-CoV-2 viral load might affect transmission but not disease severity in secondary cases – Authors' reply. *The Lancet Infectious Diseases* **21**, 915–916, doi:https://doi.org/10.1016/S1473-3099(21)00210-3 (2021).
4. Puhach, O. *et al.* Infectious viral load in unvaccinated and vaccinated patients infected with SARS-CoV-2 WT, Delta and Omicron. *medRxiv*, 2022.2001.2010.22269010, doi:10.1101/2022.01.10.22269010 (2022).
5. Salto-Alejandre, S. *et al.* SARS-CoV-2 viral load in nasopharyngeal swabs is not an independent predictor of unfavorable outcome. *Scientific Reports* **11**, 12931, doi:10.1038/s41598-021-92400-y (2021).
6. Bland, J., Kavanaugh, A., Hong, L. K. & Kadkol, S. S. Development and validation of viral load assays to quantitate SARS-CoV-2. *J Virol Methods* **291**, 114100, doi:10.1016/j.jviromet.2021.114100 (2021).
7. Liu, J. *et al.* SARS-CoV-2 cell tropism and multiorgan infection. *Cell Discov* **7**, 17, doi:10.1038/s41421-021-00249-2 (2021).
8. Jones, T. C. *et al.* Estimating infectiousness throughout SARS-CoV-2 infection course. *Science* **373**, doi:10.1126/science.abi5273 (2021).
9. van Kampen, J. J. A. *et al.* Duration and key determinants of infectious virus shedding in hospitalized patients with coronavirus disease-2019 (COVID-19). *Nat Commun* **12**, 267, doi:10.1038/s41467-020-20568-4 (2021).
10. Yin, N. *et al.* Leveraging of SARS-CoV-2 PCR Cycle Thresholds Values to Forecast COVID-19 Trends. *Frontiers in Medicine* **8**, doi:10.3389/fmed.2021.743988 (2021).
11. Binnicker, M. J. Challenges and Controversies to Testing for COVID-19. *J Clin Microbiol* **58**, doi:10.1128/JCM.01695-20 (2020).
12. Chifiriuc, M. C. *et al.* in *Food Preservation* (ed Alexandru Mihai Grumezescu) 645–669 (Academic Press, 2017).
13. Kanji, J. N. *et al.* False negative rate of COVID-19 PCR testing: a discordant testing analysis. *Virol J* **18**, 13, doi:10.1186/s12985-021-01489-0 (2021).
14. Jacot, D., Greub, G., Jatou, K. & Opo, O. Viral load of SARS-CoV-2 across patients and compared to other respiratory viruses. *Microbes Infect* **22**, 617–621, doi:10.1016/j.micinf.2020.08.004 (2020).
15. Glenet, M. *et al.* Asymptomatic COVID-19 Adult Outpatients identified as Significant Viable SARS-CoV-2 Shedders. *Scientific Reports* **11**, 20615, doi:10.1038/s41598-021-00142-8 (2021).
16. Beldomenico, P. M. Do superspreaders generate new superspreaders? A hypothesis to explain the propagation pattern of COVID-19. *Int J Infect Dis* **96**, 461–463, doi:10.1016/j.ijid.2020.05.025 (2020).
17. Yang, Q. *et al.* Just 2% of SARS-CoV-2-positive individuals carry 90% of the virus circulating in communities. *Proc Natl Acad Sci U S A* **118**, doi:10.1073/pnas.2104547118 (2021).
18. Young, R. M. *et al.* Smartphone screen testing, a novel pre-diagnostic method to identify SARS-CoV-2 infectious individuals. *Elife* **10**, doi:10.7554/eLife.70333 (2021).
19. Mohlendick, B. *et al.* ACE2 polymorphism and susceptibility for SARS-CoV-2 infection and severity of COVID-19. *Pharmacogenet Genomics* **31**, 165–171, doi:10.1097/FPC.0000000000000436 (2021).
20. Nikiforuk, A. M. *et al.* The contrasting role of nasopharyngeal angiotensin converting enzyme 2 (ACE2) transcription in SARS-CoV-2 infection: A cross-sectional study of people tested for COVID-19 in British Columbia, Canada. *EBioMedicine* **66**, 103316, doi:10.1016/j.ebiom.2021.103316 (2021).
21. Wang, R. *et al.* Genetic Screens Identify Host Factors for SARS-CoV-2 and Common Cold Coronaviruses. *Cell* **184**, 106–119 e114, doi:10.1016/j.cell.2020.12.004 (2021).
22. Amodio, E. *et al.* SARS-CoV-2 Viral Load, IFNlambda Polymorphisms and the Course of COVID-19: An Observational Study. *J Clin Med* **9**, doi:10.3390/jcm9103315 (2020).
23. Trypsteen, W., Van Cleemput, J., Snippenberg, W. V., Gerlo, S. & Vandekerckhove, L. On the whereabouts of SARS-CoV-2 in the human body: A systematic review. *PLoS Pathog* **16**, e1009037, doi:10.1371/journal.ppat.1009037 (2020).
24. Aranha, C., Patel, V., Bhor, V. & Gogoi, D. Cycle threshold values in RT-PCR to determine dynamics of SARS-CoV-2 viral load: An approach to reduce the isolation period for COVID-19 patients. *J Med Virol* **93**, 6794–6797, doi:10.1002/jmv.27206 (2021).
25. Team, C. C.-R. SARS-CoV-2 B.1.1.529 (Omicron) Variant - United States, December 1–8, 2021. *MMWR Morb Mortal Wkly Rep* **70**, 1731–1734, doi:10.15585/mmwr.mm7050e1 (2021).
26. Fan, G., Qin, X., Strelow, D. N., Hoyos, C. M. & Hansel, D. E. Comparison of SARS-CoV-2 PCR-Based Detection Using Saliva or Nasopharyngeal Swab Specimens in Asymptomatic Populations. *Microbiol Spectr* **9**, e0006221, doi:10.1128/Spectrum.00062-21 (2021).
27. Chu, D. K. W. *et al.* Molecular Diagnosis of a Novel Coronavirus (2019-nCoV) Causing an Outbreak of Pneumonia. *Clin Chem* **66**, 549–555, doi:10.1093/clinchem/hvaa029 (2020).
28. Pfister, R., Schwarz, K. A., Janczyk, M., Dale, R. & Freeman, J. B. Good things peak in pairs: a note on the bimodality coefficient. *Front Psychol* **4**, 700, doi:10.3389/fpsyg.2013.00700 (2013).
29. Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A. & Leisch, F.

Figures

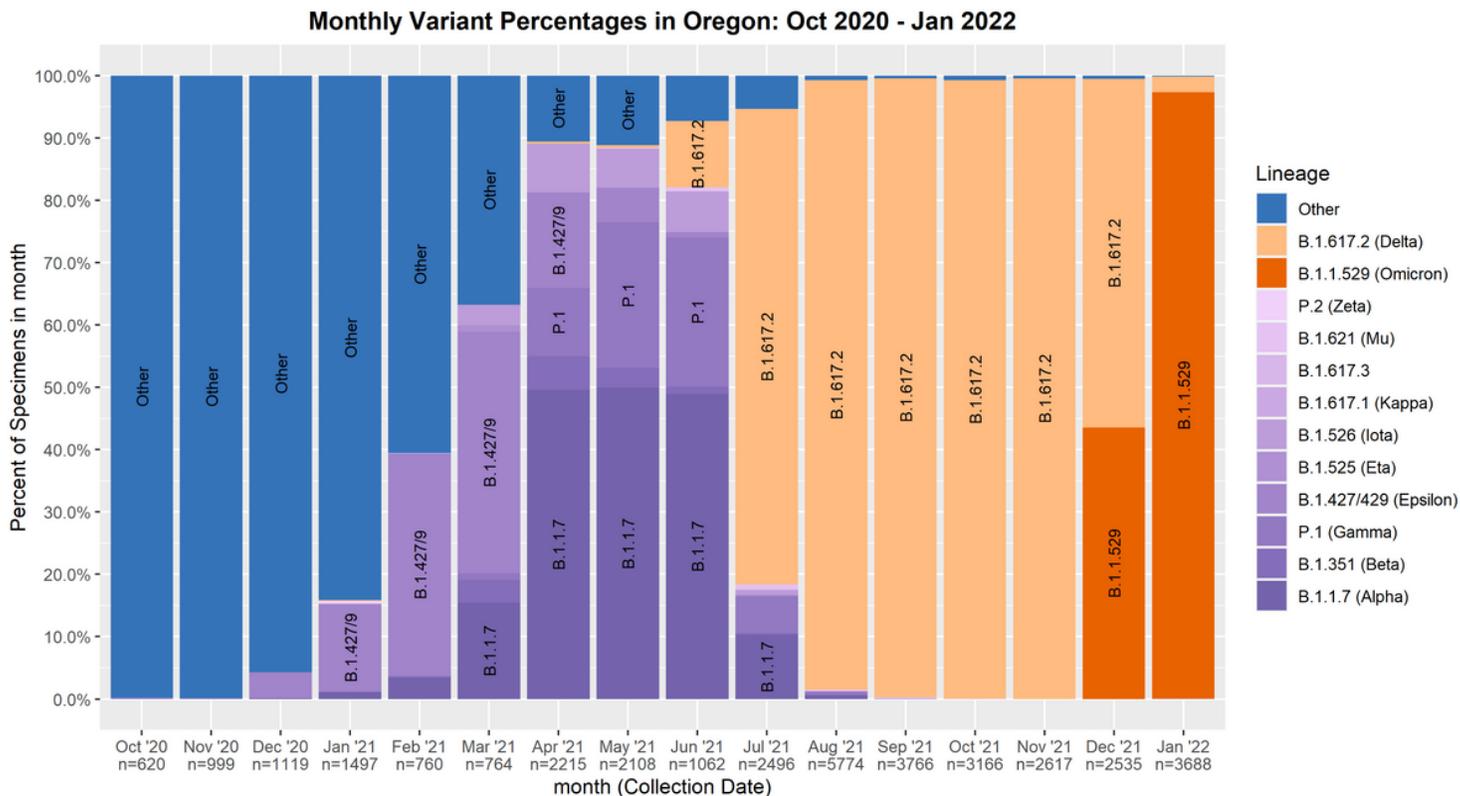


Figure 1

SARS-CoV-2 Variant Percentages in Oregon, September 2020 to January 2022.

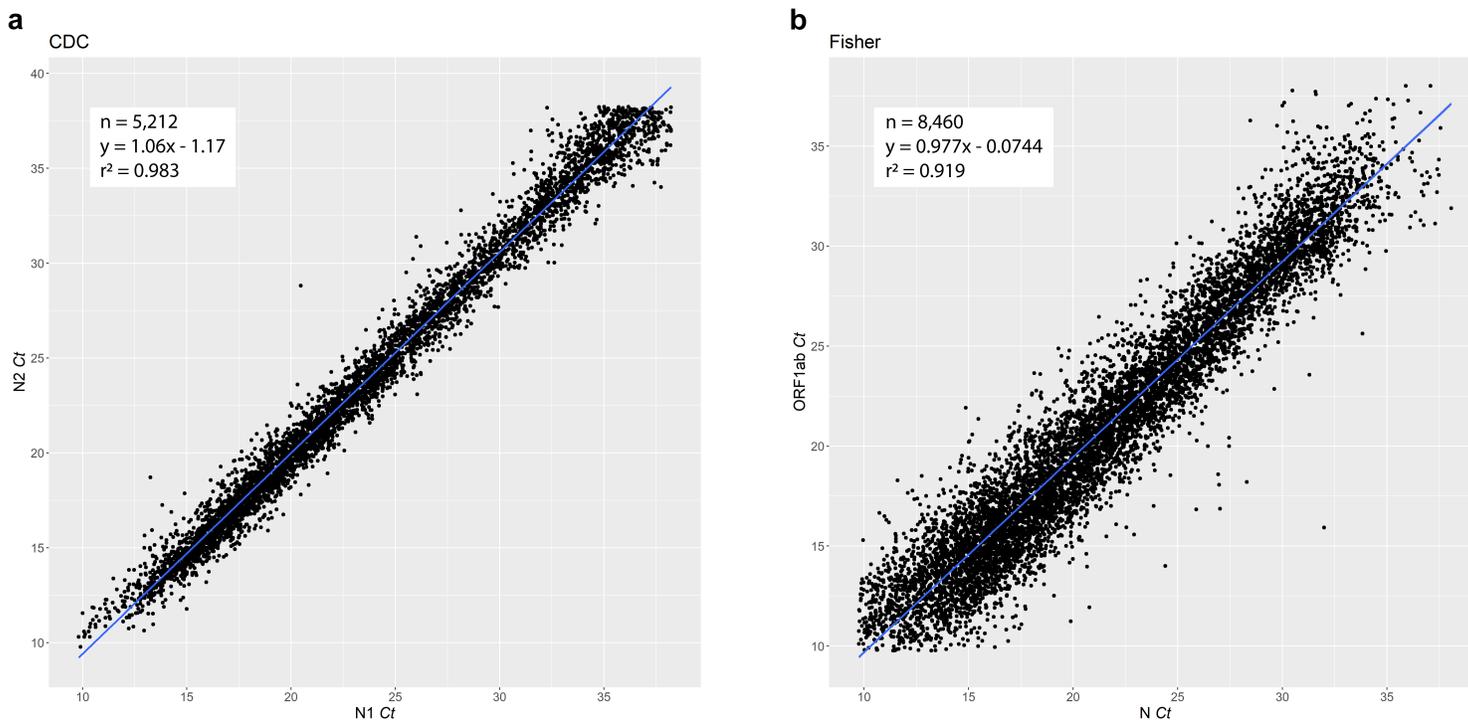


Figure 2

Correlation of PCR Ct distribution between two viral targets associated with CDC or Fisher chemistry.

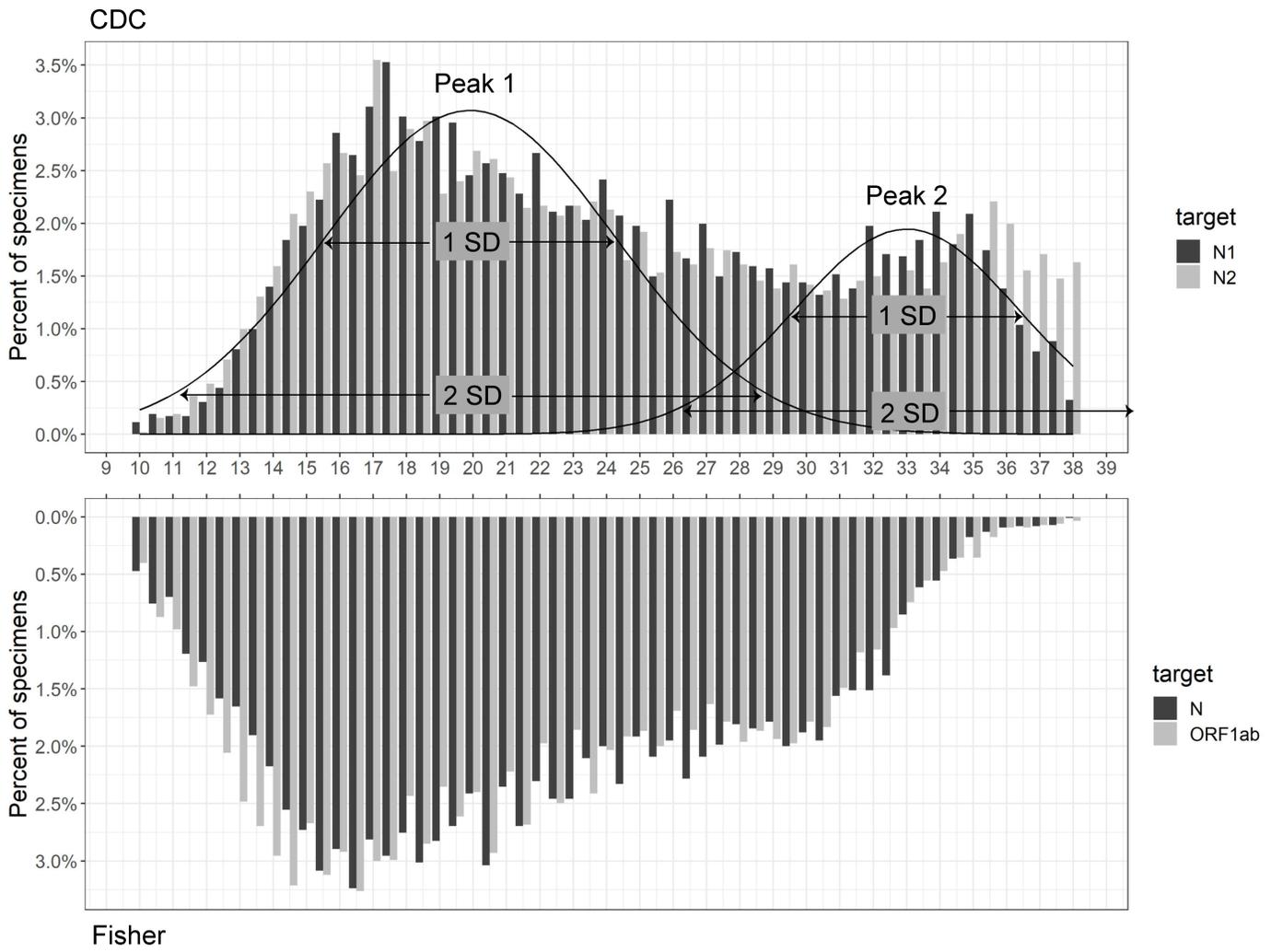
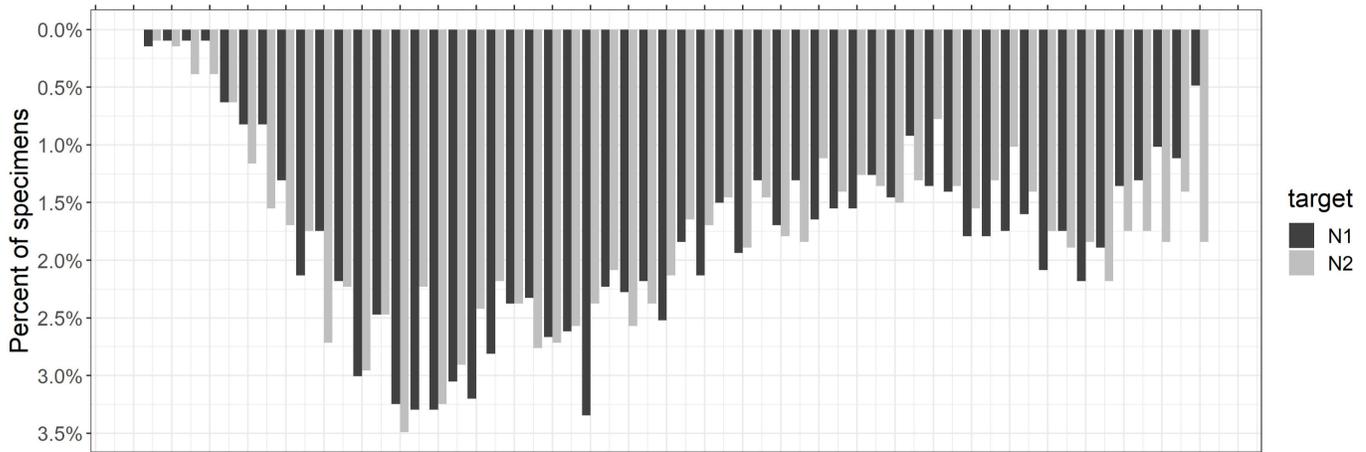
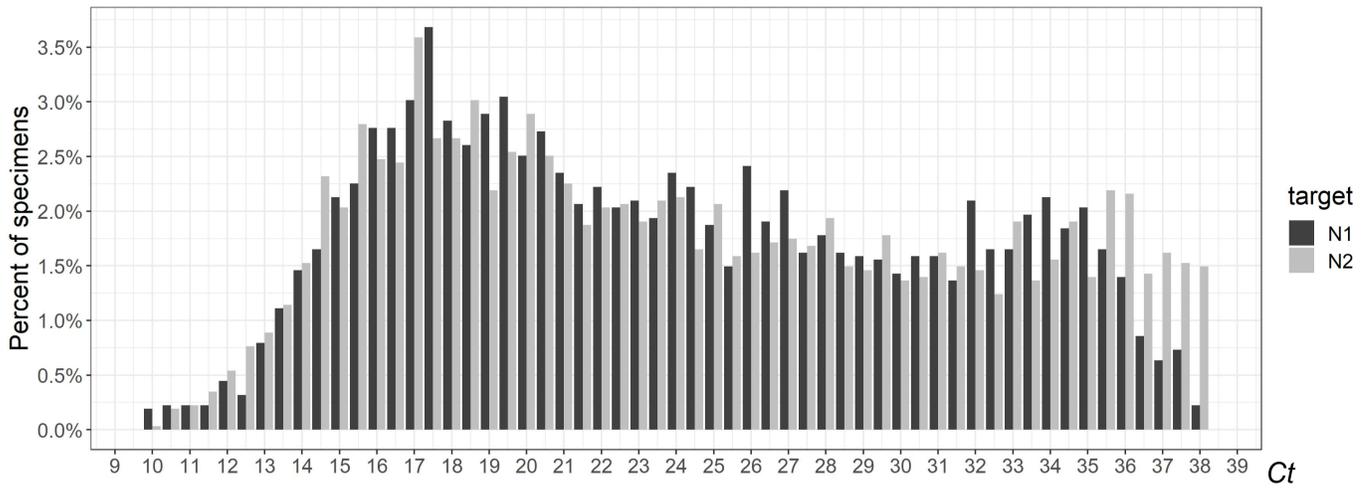


Figure 3

PCR Ct distribution of N1 and N2 targets by CDC chemistry and N and ORF1ab targets by Fisher chemistry.

Sep. 2020 - Dec. 2020



Jan. 2021 - Oct. 2021

Figure 4

PCR Ct distribution of N1 and N2 targets from samples collected in 2020 and 2021.

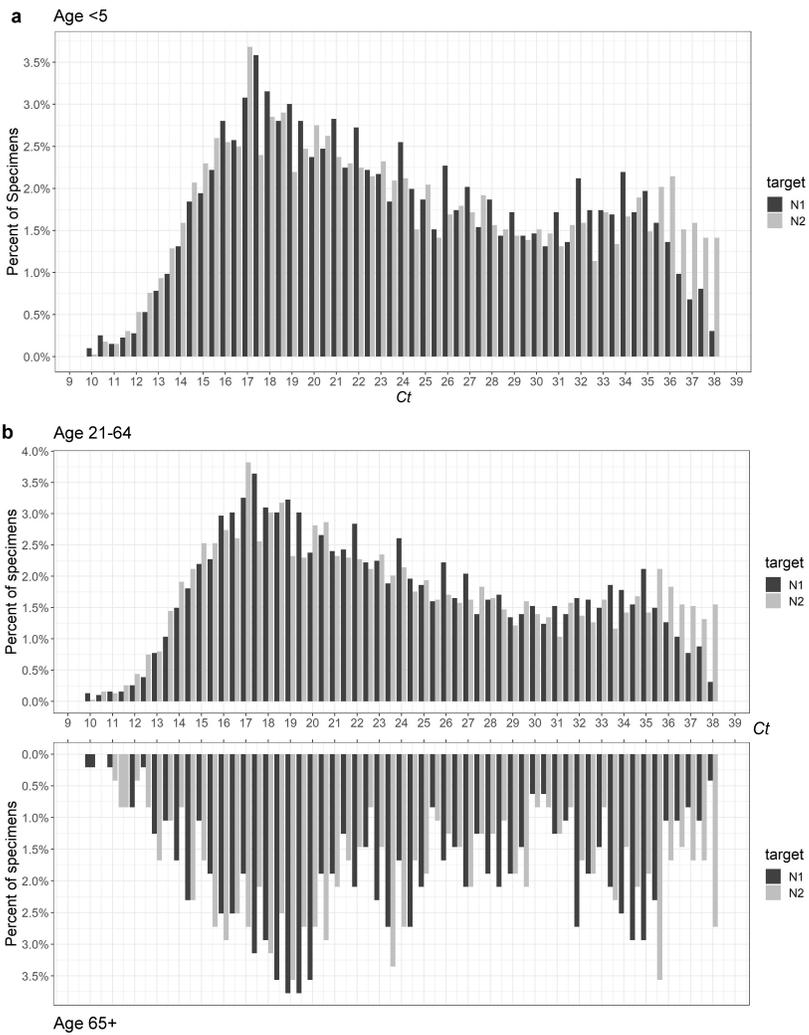


Figure 5

PCR Ct distribution of N1 and N2 associated with age groups: (a) age <5 years, (b) age groups of 21-64 years and >65 years.

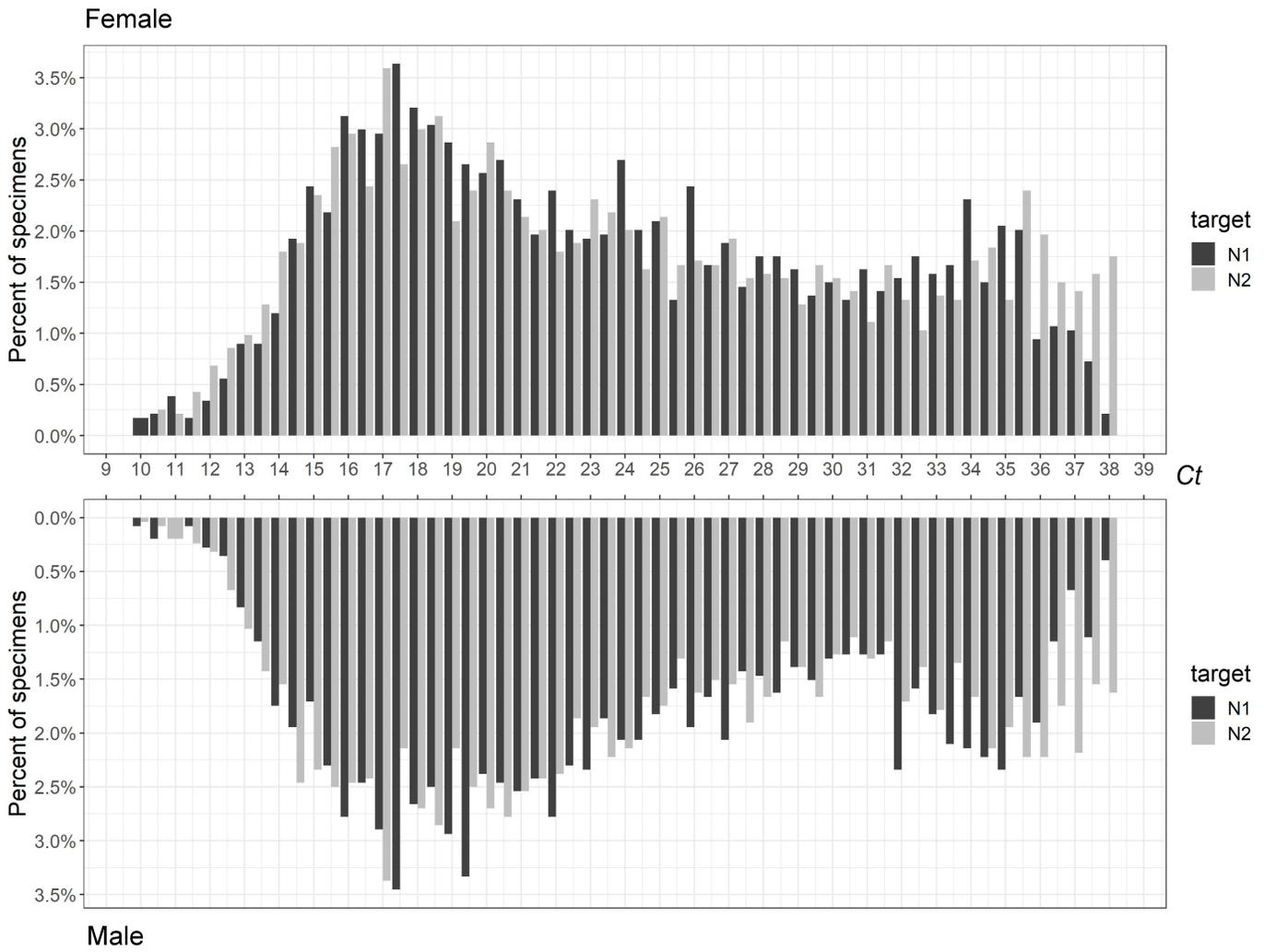


Figure 6

PCR Ct distribution of N1 and N2 targets from samples associated with female and male patients.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Bimodaldraft282022.Suppl.pdf](#)