Effect of SARS-CoV-2 digital droplet RT-PCR assay sensitivity on COVID-19 wastewater based epidemiology

Sooyeol Kim[1], Marlene K. Wolfe[2], Craig S. Criddle[1], Dorothea H. Duong[3], Vikram Chan-Herur[3], Bradley J. White[3], Alexandria B. Boehm[1*]

1. Dept of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305, United States of America
2. Rollins School of Public Health, Emory University, Atlanta, GA, 30329, United States of America
3. Verily Life Sciences, South San Francisco, CA, 94080, United States of America

* Corresponding author: aboehm@stanford.edu

Abstract

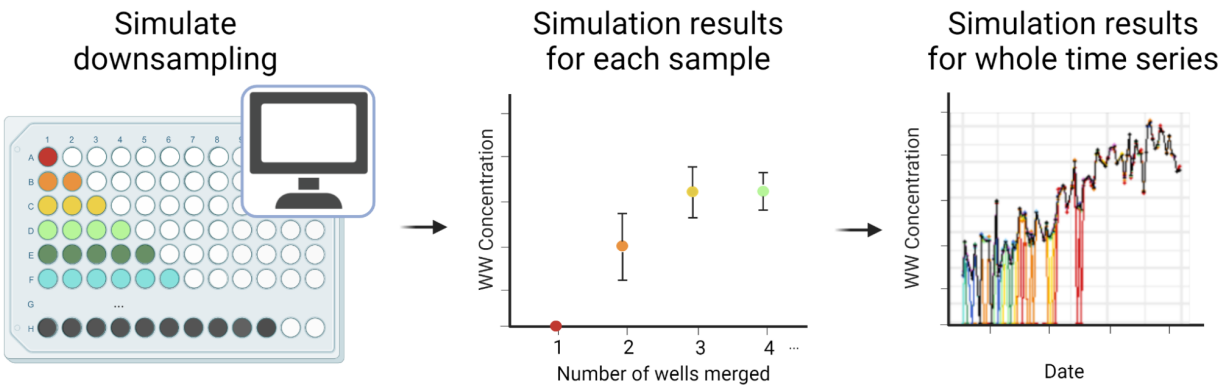
We developed and implemented a framework for examining how molecular assay sensitivity for a viral RNA genome target affects its utility for wastewater-based epidemiology. We applied this framework to digital droplet RT-PCR measurements of SARS-CoV-2 and Pepper Mild Mottle Virus genes made using 10 replicate wells, and determined how using fewer wells affected assay sensitivity and its performance for wastewater-based epidemiology applications. We used a computational, downsampling approach. When percent of positive droplets was between 0.024% and 0.5% (as was the case for SARS-CoV-2 genes during the Delta surge), measurements obtained with 3 or more wells were similar to those obtained using 10. When percent of positive droplets was less than 0.024%, then 6 or more wells were needed to obtain similar results as those obtained using 10 wells. When COVID-19 incidence is low, as it was before the Delta surge and SARS-CoV-2 gene concentrations are $<10^4$ cp/g, using 6 wells will yield a detectable concentration 90% of the time. Overall, results support an adaptive approach where assay sensitivity is increased by running 6 or more wells during periods of low SARS-

CoV-2 gene concentrations, and 3 or more wells during periods of high SARS-CoV-2 gene concentrations.

Keywords: SARS-CoV-2, assay sensitivity, replicate wells, digital PCR, wastewater

Synopsis: Adaptive approaches developed with assay sensitivity in consideration may reduce cost and increase sensitivity for wastewater-based epidemiology.

Abstract Art



Introduction

Wastewater-based epidemiology for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is becoming an increasingly important tool in monitoring coronavirus disease 2019 (COVID-19) incidence rates in communities. Monitoring programs that use samples from publicly owned treatment works (POTWs)[1,2] and from sewer conveyances[3] have actively been implemented to supplement clinical testing data. Wastewater surveillance has the advantage of

providing insights into population health by overcoming limitations of clinical testing, such as test seeking behavior and test availability. Wastewater surveillance has also been used to gain insight into the epidemiology of other respiratory viruses such as influenza A virus[4] and respiratory syncytial virus[5] (RSV), as well as gastrointestinal pathogens such as hepatitis A[6] and *Salmonella*[7]. Therefore, wastewater surveillance is likely to become increasingly useful for assessing various aspects of community health beyond COVID-19.

SARS-CoV-2 RNA concentrations in wastewater, whether measured in the solid or liquid phase or using quantitative or digital reverse-transcription polymerase chain reaction (RT-PCR), correlate to COVID-19 laboratory incidence rates in the population contributing to the wastewater[8–10]. The lower detection limit, or the sensitivity, of any method for detection of SARS-CoV-2 RNA, or any other disease target, will dictate the lowest levels of disease occurrence that can be detected using wastewater.

Digital (RT-)PCR can be a sensitive method for detecting disease targets in wastewater. Digital (RT-)PCR methods divide the entire (RT-)PCR solution (master mix, primers, probes and template) into a large number of partitions (droplets or physical partitions in a plate) such that each partition likely contains only one copy of template nucleic acid. By increasing the number of partitions (assuming their volume remains constant) and the associated reaction volume, the analytical sensitivity of the measurement can be increased. For digital droplet RT-PCR (ddRT-PCR), the number of partitions can be increased by increasing the number of wells used to analyze samples; the results from droplets generated by all replicate wells are merged to compute the final measurement. An analogous approach can be applied to other methods of digital partitioning. However, increasing the number of partitions used for each sample to improve sensitivity increases project reagent costs. In a previous study, we compared the lowest measurable concentration of SARS-CoV-2 RNA reported by different laboratories using

different pre-analytical methods and ddRT-PCR, and the detection limit decreased as the number of replicate wells used in the method increased[11]. Authors have reported using between one to ten merged wells without full justification on how the number of wells merged was selected[2,11,12].

Given the ongoing importance of wastewater surveillance for COVID-19 disease monitoring and its potential use for other disease surveillance, and increasing use of digital RT-PCR, it is important to better understand how analytical measurement sensitivity is controlled by increasing the number of partitions. To achieve this aim, we measured SARS-CoV-2 RNA and Pepper Mild Mottle Virus (PMMoV) RNA in wastewater solids samples using ddRT-PCR using 10 replicate wells. We then computationally down sampled the wells to investigate how the number of wells, and thus partitions, affects the lowest measurable concentrations of SARS-CoV-2 genes and associations between SARS-CoV-2 gene measurements and disease incidence rates. The framework developed herein for examining how molecular assay sensitivity for a viral RNA genome target affects its utility for wastewater-based epidemiology is generalizable to other infectious agents and other analytical approaches for measuring molecular targets.

Materials and Methods

The Institutional Review Board of Stanford University determined that this project does not meet the definition of human subject research as defined in federal regulations 45 CFR 46.102 or 21 CFR 50.3 and indicated that no formal IRB review is required.

POTWs and data collection

Data used in this study was obtained from an on-going SARS-CoV-2 wastewater monitoring program in California, USA described by Wolfe et al.[2] Samples were collected and processed daily between June 1, 2021 to August 31, 2021 from four POTWs: City of Davis Wastewater Treatment Plant (Dav) in Davis, South County Regional Wastewater Authority Wastewater Treatment Plant (Gil) in Gilroy, Oceanside Water Pollution Control Plant (Ocean) in San Francisco, and San Jose-Santa Clara Regional Wastewater Facility (SJ) in San Jose (listed in the order of size from smallest to largest). Further details on the POTWs and sampling procedures can be found in Wolfe et al.[2] and in Table S1. The data reported herein has not been previously published.

The solids samples were processed within 24 hours of collection exactly according to the methods described by Wolfe et al.[2] and are summarized in the SI. We measured SARS-CoV-2 N gene and PMMoV gene concentrations, and also recovery of spiked-in bovine coronavirus using ddRT-PCR using 10 replicate wells. EMMI guidelines were followed[13].

Data for each individual well was downloaded from QuantaSoft Analysis Pro software (BioRad, CA, version 1.0.596). Samples for which all wells did not have at least 10 000 generated droplets in each well were eliminated from our analysis. This eliminated a total of 41 SARS-CoV-2 N gene measurements (12 from Dav, 9 from Gil, 17 from Ocean, and 3 from SJ), resulting in 327 measurements for further analysis (89 from Dav, 80 from Gil, 83 from Ocean, 75 from SJ).

COVID-19 epidemiology data

Laboratory confirmed incident cases of COVID-19 as a function of episode date was obtained as described previously[2]; see SI for details.

Downsampling simulation

In order to estimate the SARS-CoV-2 N gene and PMMoV RNA concentration we would have obtained if we had run a smaller number of wells (X = 1 - 9), we randomly selected X wells from the 10 wells to calculate the resultant concentration:

$$Concentration\ (cp/\mu L)\ =\ -ln\left(\frac{number\ of\ negative\ droplets}{total\ accepted\ droplets}\right) / 0.00085\mu L$$

where 0.00085 µL is the volume of a single droplet[14]. If the total number of positive droplets was less than three, the concentration was denoted as not detected (ND).

A thousand simulations were conducted for each possible number of merged wells (X = 1 - 9) for each sample. The resulting concentrations were converted to units of cp/g dry weight using dimensional analysis [2]. From these thousand simulations, we calculated (1) the percent of the simulations that resulted in less than 3 positive droplets across merged wells and was designated as ND, (2) the median concentration, and (3) its interquartile range (25th and 75th percentiles). To calculate the median, we substituted zero for ND. Similar analyses were performed for ten randomly selected PMMoV measurements from each POTW; each measurement had at least 10 000 total droplets generated in each well. Simulations were conducted using RStudio (version 1.4.1106).

Statistical analysis

Statistics were computed using RStudio (version 1.4.1106). Shapiro-Wilk test was used to determine whether simulation outputs were normally distributed. The dispersion of the simulation outputs is defined by the interquartile range (IQR). The relative dispersion of the simulation outputs is described as the ratio of the median and the interquartile range. The dispersion and relative dispersion of the simulated concentrations were compared to the standard deviation of the concentration, or the standard deviation normalized by the

concentration, respectively, derived from the measurement obtained using 10 wells. The standard deviation of that measurement is the 68% confidence interval as defined by the total error from the instrument software which includes Poisson error and variation among wells; the total error formula is proprietary and not available from the vendor.

Nonparametric Kendall's tau was used to assess the association between the N gene concentrations in wastewater and laboratory confirmed COVID-19 incidence rates. Kendall's tau was calculated for both the entire time series and the low incidence month of June. For this analysis, zero was substituted for measurements considered NDs.

The theoretical lower measurement limit was calculated for each number of merged wells (X = 1 - 10) by: 1) calculating the concentration resulting from three positive droplets total across merged wells out of 20 000 total accepted droplets (average number of droplets generated) per each well merged, and 2) converting the concentration to units of cp/g dry weight using average solid content for samples from each POTW (Table S2).

Linear regression was used to derive an empirical relationship between $\log_{10}$-transformed COVID-19 laboratory-confirmed incidence rates and $\log_{10}$-transformed measured SARS-CoV-2 N gene concentrations using data obtained by merging ten wells; relationships were quantified for each POTW separately, and for the POTWs in aggregate. For the linear regression, NDs were substituted with half the theoretical lower measurement limit calculated for X = 10. Using the empirical relationship between incidence rate and SARS-CoV-2 RNA concentration for the associated POTW, the lowest detectable COVID-19 incidence rate was estimated based on the calculated theoretical lower measurement limits.

A logistic regression was used to model the fraction of samples that were assigned a concentration (versus assigned ND) for X = 1 - 9 as a function of the true concentration of the sample, defined as the concentration obtained using 10 wells. The concentration corresponding to a detection frequency of 0.5 ($C_{0.5}$) was calculated using the regression equation. For this analysis, half of the theoretical lower measurement limit was substituted for the 6 NDs for X =10.

Results

QA/QC

Negative and positive extraction and PCR controls were negative and positive, respectively. BCoV was used as a process control to verify that the extraction was successful and there was no gross inhibition in quantification. Samples that had less than 10% recovery of BCoV were rerun; no sample had less than 10% recovery. No further correction or analysis of BCoV recoveries are provided here given the complexities of interpreting recoveries of exogenous controls[15]. PMMoV concentrations across samples are similar to those measured and reported previously, also suggesting no gross issues with extraction or inhibition (Fig. S1). Data on N and PMMoV gene concentrations are available through the Stanford Digital Repository (https://purl.stanford.edu/km637ys9238).

Simulation output trends at high and low concentrations

For each measurement, a thousand simulations were conducted to sample each possible number of merged wells (X = 1 - 9) and the results are reported as concentration in units of cp/g dry weight (cp/g, hereafter). The resulting concentration distributions obtained for each X for each measurement were not normally distributed based on Shapiro-Wilk tests ($p < 0.05$);

205 therefore, medians and interquartile ranges are used to describe the results. Simulation outputs
206 of example measurements for SARS-CoV-2 N gene in samples collected during a period of low
207 COVID-19 incidence (June 1, 2021) and high COVID-19 incidence (August 31, 2021); as well as
208 example PMMoV gene measurements (June 6, 2021) are provided in Figure 1. The simulation
209 dispersion can be compared to the results obtained using 10 merged wells and its standard
210 deviation, defined by the total error as reported by the ddPCR instrument, which includes errors
211 associated with the Poisson distribution and variability among replicate wells.
212
213 There are several important insights to glean from these results. First, when the number of
214 positive droplets is less than 3 across the 10 wells (< 0.0015% of droplets positive), and the
215 measurement is deemed as ND (see Dav sample from June 1, 2021), the results obtained from
216 fewer wells agree with the results obtained from 10 wells. Second, when the number of positive
217 droplets is high (for example, for PMMoV where there are 10^4~10^5 positive droplets across 10
218 wells, 5 - 50% of droplets positive), then the dispersion as represented by the interquartile range
219 (IQR), in the simulations for X < 10 wells is similar to the standard deviation reported by the
220 instrument for X = 10 wells. Finally, when the number of positive droplets is intermediate to
221 these two regimes (fraction of positive droplets between 0.0015% and 0.5%), then the IQR
222 increases as X decreases and is often greater than the standard deviation from the X = 10 well
223 measurement, particularly when X < 3. Below SARS-CoV-2 concentration of 10^4 cp/g (where
224 the fraction of positive droplets is 0.0062% - 0.024% depending on POTW), the value of X
225 below which the relative dispersion, defined by IQR divided by the median, is larger than the
226 standard deviation normalized by the measured concentration for X = 10 scales inversely with
227 SARS-CoV-2 N gene concentration (Figure S2).
228
229 As PMMoV is present in such high concentrations and the measurements yielded high positive
230 droplet counts, resulting in similar concentrations across X = 1 - 10, the remainder of this

231 analysis will focus on the SARS-CoV-2 N gene concentrations, as those measurements yielded
232 low to intermediate positive droplet counts.
233
234 *Theoretical sensitivity*
235 An empirical relationship between the $\log_{10}$-transformed COVID-19 incidence rate and the $\log_{10}$-
236 transformed SARS-CoV-2 concentration using ten merged wells was derived with linear
237 regression. The regression showed that for 1 $\log_{10}$ increase in SARS-CoV-2 RNA cp/g, there
238 was between 0.50 and 0.88 $\log_{10}$ increase in laboratory-confirmed COVID-19 incidence rate
239 (Table S3), depending on the POTW. The data from all four POTWs appear to fall on a single
240 line when COVID-19 incidence rate is plotted against SARS-CoV-2 concentration (Figure S3);
241 when data from all POTW are combined, there was a 0.64 $\log_{10}$ increase in laboratory-
242 confirmed COVID-19 incidence rate for 1 $\log_{10}$ increase in SARS-CoV-2 RNA cp/g.
243
244 Theoretical lower measurement limit (Table S4) and the corresponding incidence rate lower limit
245 (Table S5) was calculated. The theoretical lower measurement limit for each POTW ranged
246 from 7500 (SJ) to 24000 (Gil) cp/g when using only one well and from 750 (SJ) to 2400 (Gil)
247 cp/g when using ten merged wells. Since this theoretical lower measurement limit was
248 calculated with average solid content of samples from each POTW by measuring the percent
249 weight of the dewatered solids and assuming a total of 20 000 generated droplets, the observed
250 lower measurement limit may be different. The corresponding incidence rate lower limit per 100
251 000, calculated using the empirical relationships in Table S3, ranged from 1.6 (SJ) to 6.9 (Gil)
252 when using one well and from 0.2 (SJ) to 1.4 (Gil) when using ten merged wells (Table S5).
253
254 *Association with clinical data*
255 Time series of median concentrations resulting from a thousand simulations for each
256 measurement for all possible numbers of wells are provided in Figure 2. Lines representing low

number of wells deviate from those representing higher number of wells during the low incidence rate period in June 2021 when SARS-CoV-2 N gene concentrations in wastewater were relatively low. When the entire study period of June 1, 2021 to August 31, 2021 was considered, SARS-CoV-2 N gene concentrations were positively and significantly correlated with 7-day smoothed COVID-19 incidence rates at all four POTWs regardless of X (Table S6, tau > 0.54, $p < 0.05$ for all); X did not have an effect on Kendall's tau when considering the entire time series (tau changed by < 0.05 as X varied). However, when considering the month of June alone when COVID-19 incidence rate was relatively low, SARS-CoV-2 N gene concentrations were positively and significantly correlated with 7-day smoothed COVID-19 incidence rates only when X > 1 ($\boldsymbol{\alpha} = 0.1$) and X did affect tau by as much as 0.15 (Table S7).

Detection frequency

Detection frequency for SARS-CoV-2 N gene across all POTWs was examined as a function of X and the true concentration of N gene, as defined by the concentration obtained using 10 merged wells (Figure 3). Logistic regressions were fit to the curves (Table S8) and the concentration at which the detection frequency is ≤ 0.5 ($C_{0.5}$) was calculated, as well as the corresponding incidence rate using the empirical relationship between log-transformed SARS-CoV-2 RNA N gene concentrations and COVID-19 incidence rate for all POTWs (Table 1). As there were not as many measurements that resulted in ND for X > 8, the confidence interval of the logistic regression is relatively large, especially for X = 9 and 10. Therefore $C_{0.5}$ derived for X = 9 and10 is an extrapolation of the data used in this study and may be more uncertain. $C_{0.5}$ scales with X according to the following equation $\log_{10}(C_{0.5}) = -0.13*X + 4.0$ ($R^2 = 0.90$, p-value $< 10^{-4}$). Percentage of samples that fall under $C_{0.5}$ for each X during the entire time series of three months and during the month of June was calculated. Most samples that fell under $C_{0.5}$ were collected in June. Figure 4 shows data exclusively in June 2021 to focus on this time

period; measurements obtained using X = 10 are shown with the concentration at which the detection frequency is ≤ 0.5 for X = 1, 3, 6, 10 is shown in dashed lines. During this low incidence rate period, for X =1, 81% (92 of 113) measurements fall below the $C_{0.5}$ across all POTWs. For X = 3, 38% (43 of 113) measurements, and for X = 6, 8.8% (10 of 113) fall below the corresponding $C_{0.5}$. For reference, when X = 10, 4.4% (5 of 113) measurements fall below the $C_{0.5}$ across all POTWs.

Discussion

We measured concentrations of SARS-CoV-2 and PMMoV genes daily in wastewater settled solids at four POTWs in California using ddRT-PCR during a period of time that included both low and high COVID-19 incidence. We used 10 merged wells for the measurements, and then determined how the measurement would have been affected by using fewer than 10 wells through a down-sampling scheme. Our findings indicate that when a large fraction of droplets are positive (> 5% positive), as was observed for PMMoV, a virus found in high quantity in human stool and wastewater[16], concentrations measured using just one well are similar to those obtained using ten when considering the variability associated with the measurements. On the other hand, when a smaller fraction of droplets are positive (< 0.5%), as was the case for the SARS-CoV-2 gene measurements, using fewer wells can result in measurements that may vary from those obtained using 10 wells, and more measurements characterized as non-detects.

For the SARS-CoV-2 gene measurements, variability in measurements increased as the number of wells decreased. Generally, we found that when the fraction of positive droplets was greater than 0.024% (corresponding to a conservative approximate concentration of 10^4 cp/g

dry weight), that the variability in the measurement resulting from using 3 or more wells was similar or smaller than the measurement total error obtained using 10 wells. In contrast, when the fraction of positive droplets was less than 0.024%, the variability in the measurements resulting from using 6 or more wells was similar or smaller than the measurement total error obtained using 10 wells. These results could guide adaptive analysis plans that use fewer wells to reduce costs when concentrations of SARS-CoV-2 are relatively high.

The probability of obtaining a non-detect increased as the SARS-CoV-2 gene concentration decreased and the number of wells used in ddRT-PCR decreased. Using logistic regression, we identified the concentration at which the detection frequency was less than 0.5 ($C_{0.5}$), and this value varies inversely with the number of wells; that is $C_{0.5}$ is higher when fewer wells are used for ddRT-PCR. This means that when low concentrations of SARS-CoV-2 genes are expected, using too few wells can result in a large number of non-detects. For example, during the low incidence period of June, if only 1 well had been used instead of 10, 92 of the 113 measurements across four POTWs would have been below $C_{0.5}$. We found that at least 6 wells were needed to achieve 90% of the measurements to be above $C_{0.5}$ for June 2021.

Consistent with other studies, the wastewater concentration showed positive and significant correlation with 7-day smoothed COVID-19 incidence rates. When there was variation in COVID-19 incidence rates within the time frame being investigated (here before and during the Delta variant surge), the number of wells being used for the analysis did not affect the magnitude or statistical significance of the correlation. There was a positive and significant correlation even when using only one well because there was enough variation in both variables, although the majority of June measurements were characterized as non-detects. This illustrates that finding a significant correlation between disease incidence and SARS-CoV-2 gene concentrations does not necessarily indicate good measurement sensitivity. It should be

noted that while we take the laboratory confirmed COVID-19 incidence rates to be reflective of the level of COVID-19 that the community is experiencing, they are likely an underestimate of incidence rates in the sewershed as the reported incidence rates are dependent on test-seeking behavior and test availability[17].

There are a few limitations of this analysis. First, in our analysis, we assumed that the measurement obtained using 10 wells is the "true concentration" and compared all results simulated with fewer than 10 wells to the true concentration and its error from the ddRT-PCR instrument. Second, the results presented herein regarding assay sensitivity, and in particular the $C_{0.5}$ values in Table 1 are specific to the methods applied in this study. The relationship between the number of wells used to the number of non-detects, and the lowest measurable concentration will be impacted by the pre-analytical and analytical processes used. While the specific values in Table 1 are not externally valid, unless others are using our exact methods (available on protocols.io[18–20]), the framework for examining the required sensitivity for wastewater surveillance is. That is, careful attention to how sensitivity affects the lowest measurable concentration and the number of non-detects, as well as the relationships between these values and laboratory confirmed COVID-19 incidence rates is needed to fully understand how decisions on assay implementation are made.

Conclusions

We developed and implemented a framework for examining how molecular assay sensitivity for a viral RNA genome target affects its utility for wastewater-based epidemiology. The framework involves understanding how assay sensitivity affects lowest measurable concentrations in units of copies per environmental matrix mass, and the detection probability of a target that is present; and how this change during periods of different disease occurrence can affect resultant

statistical associations between the viral target and measures of disease incidence. We applied this framework to digital droplet RT-PCR (ddRT-PCR) measurements of a SARS-CoV-2 gene made using 10 replicate wells, and determined how using fewer wells affected assay sensitivity and its performance for wastewater-based epidemiology applications. From a reagent cost savings perspective, we recommend an adaptive analytical approach where assay sensitivity is increased by running more replicate wells (6 or more) during periods of low SARS-CoV-2 gene concentrations (using our methods, $< 10^4$ cp/g) and COVID-19 incidence rate (< 3.5/100 000) and fewer replicate wells (3 or more) during periods of higher SARS-CoV-2 RNA concentrations and COVID-19 incidence. While the precise recommendations here are not generalizable, unless one is using the same pre-analytical and analytical protocols, the framework and the conclusion that adaptive approaches can reduce costs and increase sensitivity during periods of low disease incidence can be applied to other methods and other wastewater-based epidemiology targets.

Supporting Information

Additional experimental, POTW (Table S1-S2), and COVID-19 epidemiological data details; simulation output (Table S3-S8); and results (Fig S1-S3).

Conflict of interest

Dorothea H. Duong, Vikram Chan-Herur, and Bradley J. White are employees of Verily Life Sciences.

Acknowledgements

This study was supported by the CDC Foundation, NSF RAPID (CBET-2023057), and by the Epidemiology and Laboratory Capacity for Infectious Diseases Cooperative Agreement (no. 6NU50CK000539-03-02) from CDC. We thank the California Department of Public Health

COVID-19 Wastewater Surveillance, Epidemiology and Data teams for their help with COVID-19 incidence data. We thank Dr. Linlin Li, Michael Balliet, Dr. Pamela Stoddard and Dr. George Han at the County of Santa Clara Public Health Department for provision of case data. Numerous people contributed to sample collection including including Payak Sarkar (SJ), Noel Enoki (SJ), Amy Wong (SJ), Alexandre Miot (Ocean), Lily Chan (Ocean), the Oceanside plant operations personnel, Saeid Vaziry (Gil), Chris Vasquez (Gil), and Jeromy Miller (Dav).

References

(1) Kantor, R. S.; Greenwald, H. D.; Kennedy, L. C.; Hinkle, A.; Harris-Lovett, S.; Metzger, M.; Thornton, M. M.; Paluba, J. M.; Nelson, K. L. *Operationalizing a Routine Wastewater Monitoring Laboratory for SARS-CoV-2*; preprint; Epidemiology, 2021. https://doi.org/10.1101/2021.06.06.21258431.

(2) Wolfe, M. K.; Topol, A.; Knudson, A.; Simpson, A.; White, B.; Vugia, D. J.; Yu, A. T.; Li, L.; Balliet, M.; Stoddard, P.; Han, G. S.; Wigginton, K. R.; Boehm, A. B. High-Frequency, High-Throughput Quantification of SARS-CoV-2 RNA in Wastewater Settled Solids at Eight Publicly Owned Treatment Works in Northern California Shows Strong Association with COVID-19 Incidence. *mSystems* **2021**, *6* (5). https://doi.org/10.1128/mSystems.00829-21.

(3) Gibas, C.; Lambirth, K.; Mittal, N.; Juel, M. A. I.; Barua, V. B.; Roppolo Brazell, L.; Hinton, K.; Lontai, J.; Stark, N.; Young, I.; Quach, C.; Russ, M.; Kauer, J.; Nicolosi, B.; Chen, D.; Akella, S.; Tang, W.; Schlueter, J.; Munir, M. Implementing Building-Level SARS-CoV-2 Wastewater Surveillance on a University Campus. *Sci. Total Environ.* **2021**, *782*, 146749. https://doi.org/10.1016/j.scitotenv.2021.146749.

(4) Wolfe, M. K.; Duong, D.; Bakker, K. M.; Ammerman, M.; Mortenson, L.; Hughes, B.; Martin, E. T.; White, B. J.; Boehm, A. B.; Wigginton, K. R. *Wastewater-Based Detection of an Influenza Outbreak*; preprint; Public and Global Health, 2022. https://doi.org/10.1101/2022.02.15.22271027.

(5) Hughes, B.; Duong, D.; White, B. J.; Wigginton, K. R.; Chan, E. M. G.; Wolfe, M. K.; Boehm, A. B. Respiratory Syncytial Virus (RSV) RNA in Wastewater Settled Solids Reflects RSV Clinical Positivity Rates. *Environ. Sci. Technol. Lett.* **2022**, *9* (2), 173–178. https://doi.org/10.1021/acs.estlett.1c00963.

(6) McCall, C.; Wu, H.; O'Brien, E.; Xagoraraki, I. Assessment of Enteric Viruses during a Hepatitis Outbreak in Detroit MI Using Wastewater Surveillance and Metagenomic Analysis. *J. Appl. Microbiol.* **2021**, *131* (3), 1539–1554. https://doi.org/10.1111/jam.15027.

(7) Diemert, S.; Yan, T. Municipal Wastewater Surveillance Revealed a High Community Disease Burden of a Rarely Reported and Possibly Subclinical *Salmonella Enterica* Serovar Derby Strain. *Appl. Environ. Microbiol.* **2020**, *86* (17), e00814-20. https://doi.org/10.1128/AEM.00814-20.

(8) Peccia, J.; Zulli, A.; Brackney, D. E.; Grubaugh, N. D.; Kaplan, E. H.; Casanovas-Massana, A.; Ko, A. I.; Malik, A. A.; Wang, D.; Wang, M.; Warren, J. L.; Weinberger, D. M.; Arnold, W.; Omer, S. B. Measurement of SARS-CoV-2 RNA in Wastewater Tracks Community Infection Dynamics. *Nat. Biotechnol.* **2020**, *38* (10), 1164–1167.

https://doi.org/10.1038/s41587-020-0684-z.
(9) Graham, K. E.; Loeb, S. K.; Wolfe, M. K.; Catoe, D.; Sinnott-Armstrong, N.; Kim, S.; Yamahara, K. M.; Sassoubre, L. M.; Mendoza Grijalva, L. M.; Roldan-Hernandez, L.; Langenfeld, K.; Wigginton, K. R.; Boehm, A. B. SARS-CoV-2 RNA in Wastewater Settled Solids Is Associated with COVID-19 Cases in a Large Urban Sewershed. *Environ. Sci. Technol.* **2021**, *55* (1), 488–498. https://doi.org/10.1021/acs.est.0c06191.
(10) Fernandez-Cassi, X.; Scheidegger, A.; Bänziger, C.; Cariti, F.; Tuñas Corzon, A.; Ganesanandamoorthy, P.; Lemaitre, J. C.; Ort, C.; Julian, T. R.; Kohn, T. Wastewater Monitoring Outperforms Case Numbers as a Tool to Track COVID-19 Incidence Dynamics When Test Positivity Rates Are High. *Water Res.* **2021**, *200*, 117252. https://doi.org/10.1016/j.watres.2021.117252.
(11) Kim, S.; Kennedy, L. C.; Wolfe, M. K.; Criddle, C. S.; Duong, D. H.; Topol, A.; White, B. J.; Kantor, R. S.; Nelson, K. L.; Steele, J. A.; Langlois, K.; Griffith, J. F.; Zimmer-Faust, A. G.; McLellan, S. L.; Schussman, M. K.; Ammerman, M.; Wigginton, K. R.; Bakker, K. M.; Boehm, A. B. SARS-CoV-2 RNA Is Enriched by Orders of Magnitude in Primary Settled Solids Relative to Liquid Wastewater at Publicly Owned Treatment Works. *Environ. Sci. Water Res. Technol.* **2022**, 10.1039.D1EW00826A. https://doi.org/10.1039/D1EW00826A.
(12) Feng, S.; Roguet, A.; McClary-Gutierrez, J. S.; Newton, R. J.; Kloczko, N.; Meiman, J. G.; McLellan, S. L. Evaluation of Sampling, Analysis, and Normalization Methods for SARS-CoV-2 Concentrations in Wastewater to Assess COVID-19 Burdens in Wisconsin Communities. *ACS EST Water* **2021**, *1* (8), 1955–1965. https://doi.org/10.1021/acsestwater.1c00160.
(13) Borchardt, M. A.; Boehm, A. B.; Salit, M.; Spencer, S. K.; Wigginton, K. R.; Noble, R. T. The Environmental Microbiology Minimum Information (EMMI) Guidelines: QPCR and DPCR Quality and Reporting for Environmental Microbiology. *Environ. Sci. Technol.* **2021**, *55* (15), 10210–10223. https://doi.org/10.1021/acs.est.1c01767.
(14) Bio-Rad Laboratories, Inc. Droplet Digital PCR Application Guide.
(15) Kantor, R. S.; Nelson, K. L.; Greenwald, H. D.; Kennedy, L. C. Challenges in Measuring the Recovery of SARS-CoV-2 from Wastewater. *Environ. Sci. Technol.* **2021**, *55* (6), 3514–3519. https://doi.org/10.1021/acs.est.0c08210.
(16) Rosario, K.; Symonds, E. M.; Sinigalliano, C.; Stewart, J.; Breitbart, M. Pepper Mild Mottle Virus as an Indicator of Fecal Pollution. *Appl. Environ. Microbiol.* **2009**, *75* (22), 7261–7267. https://doi.org/10.1128/AEM.00410-09.
(17) Wu, S. L.; Mertens, A. N.; Crider, Y. S.; Nguyen, A.; Pokpongkiat, N. N.; Djajadi, S.; Seth, A.; Hsiang, M. S.; Colford, J. M.; Reingold, A.; Arnold, B. F.; Hubbard, A.; Benjamin-Chung, J. Substantial Underestimation of SARS-CoV-2 Infection in the United States. *Nat. Commun.* **2020**, *11* (1), 4507. https://doi.org/10.1038/s41467-020-18272-4.
(18) Topol, A.; Wolfe, M. K.; White, B.; Wigginton, K.; B Boehm, A. High Throughput Pre-Analytical Processing of Wastewater Settled Solids for SARS-CoV-2 RNA Analyses V1, 2021. https://doi.org/10.17504/protocols.io.btyqnpvw.
(19) Topol, A.; Wolfe, M. K.; Wigginton, K.; White, B.; B Boehm, A. High Throughput RNA Extraction and PCR Inhibitor Removal of Settled Solids for Wastewater Surveillance of SARS-CoV-2 RNA V1, 2021. https://doi.org/10.17504/protocols.io.btyrnpv6.
(20) Topol, A.; Wolfe, M. K.; White, B.; Wigginton, K.; B Boehm, A. High Throughput SARS-COV-2, PMMOV, and BCoV Quantification in Settled Solids Using Digital RT-PCR V1, 2021. https://doi.org/10.17504/protocols.io.btywnpxe.

Figures

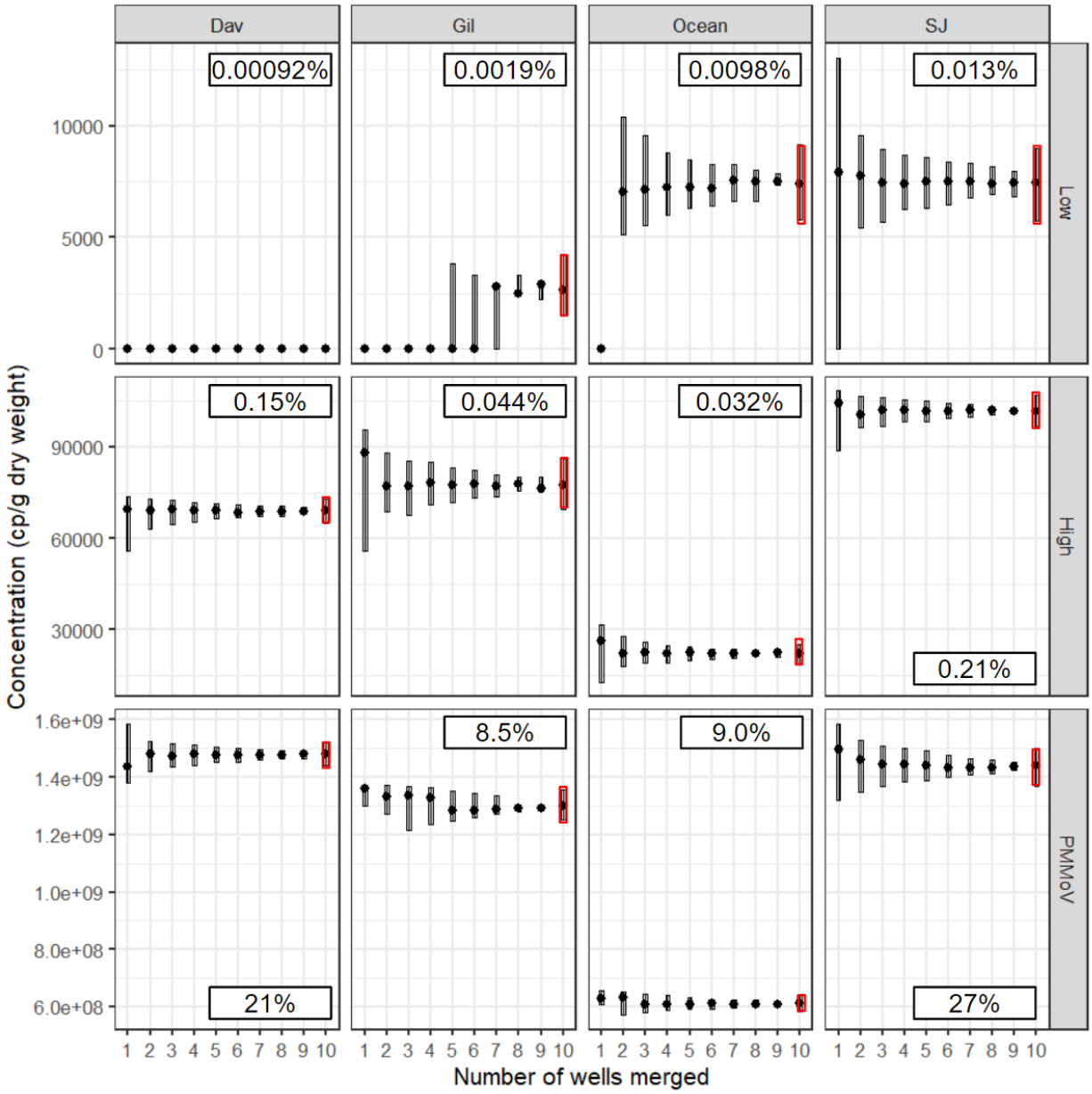


Figure 1. Example output of 1000 simulation results of selecting random sets of wells out of ten wells to calculate the final concentration in wastewater solids (cp/g dry weight) for SARS-CoV-2 N gene in June 1, 2021 sample during low COVID-19 incidence (top), SARS-CoV-2 N gene in August 31, 2021 sample during high COVID-19 incidence (middle), and for PMMoV in June 6, 2021 sample (bottom). For X = 1 - 9, the circle in the box represents the median, and the top and bottom of the box represent 75th and 25th percentile, respectively. For X = 10, the circle in the red box represents the software reported concentration from merging all ten wells, and the top and bottom of the box represent upper and lower confidence intervals, respectively, from 68% total error as given by the instrument software, which includes errors associated with the Poisson distribution and variability among replicate wells. Percentage of positive droplets in 10 wells is shown in boxes within each plot.

490

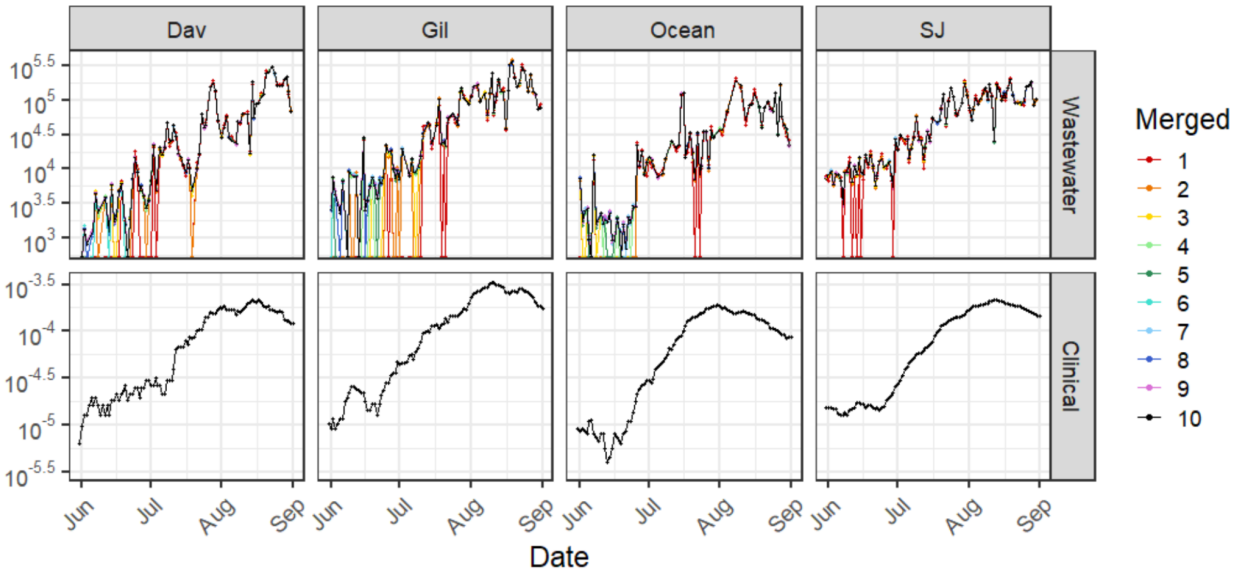


Figure 2. Time series of (top to bottom) SARS-CoV-2 N gene concentration in wastewater solids (cp/g dry weight) and 7 day centered smoothed average laboratory-confirmed SARS-CoV-2 incidence rate for each of the four POTWs from June 1, 2021 to August 31, 2021. Note that the SARS-CoV-2 N gene concentrations are displayed in $\log_{10}$-scale format for ease of visualization. Each wastewater data point for X = 1 - 9 represents median SARS-CoV-2 RNA concentration for a single sample obtained from 1000 simulations; for X = 10, each data point is the concentration obtained by merging 10 wells. Samples that resulted in ND were substituted with zero.

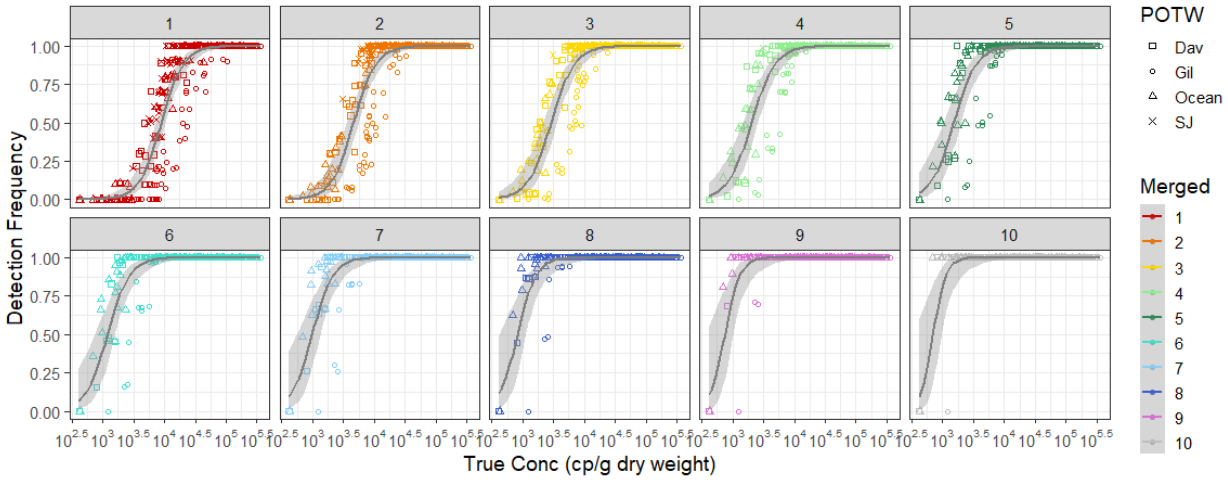


Figure 3. Detection frequency for all samples across four POTWs (y-axis) plotted against true concentration (x-axis), which is defined as concentration obtained by merging all ten wells. Each data point shows the fraction of 1000 simulations that did not result in ND in each well on the y-axis and its true concentration on the x-axis. ND for X = 10 was substituted with half of the theoretical lower measurement limit. 95% confidence intervals of the logistic regression are shown as the gray ribbon.

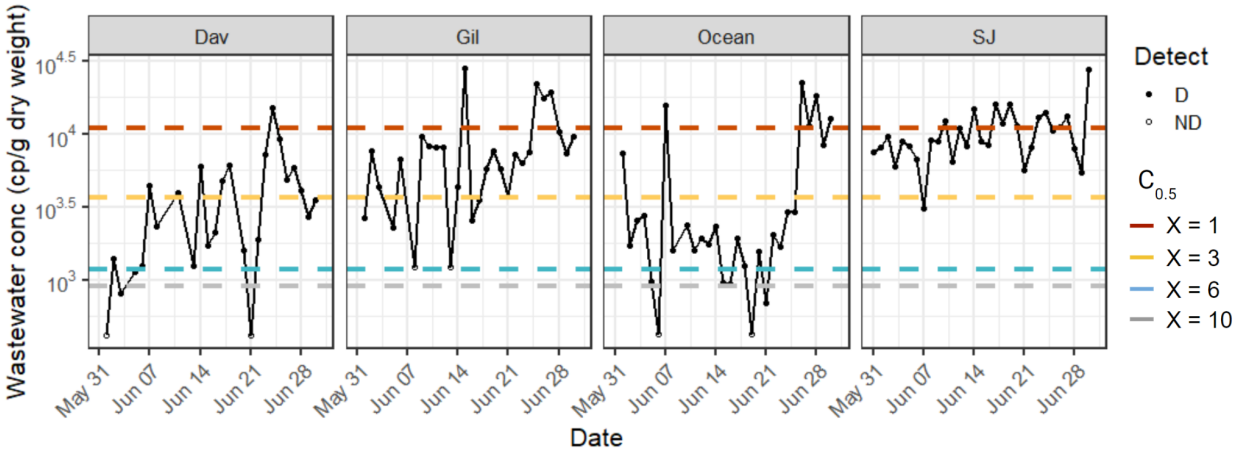


Figure 4. Time series of SARS-CoV-2 concentration in wastewater solids (cp/g dry weight) for each of the four POTWs during low COVID-19 incidence month of June 2021. Note that the SARS-CoV-2 RNA concentrations are displayed in $\log_{10}$-scale format for ease of visualization. Each wastewater data point represents SARS-CoV-2 concentration measured for a single sample. Samples above the lower measurement limit are shown as filled circles. Samples that resulted in ND, shown as empty circles, were substituted with half the lower measurement limit.

Tables

Table 1. $C_{0.5}$ shows the concentration that would yield a detect half the time in units of cp/g dry weight, calculated using the logistic regression (Figure 3). % Samples under $C_{0.5}$ shows percentage of samples in the data set that were under $C_{0.5}$ during the entire time series (All) and the month of June when COVID-19 incidence rate was low (June). Incidence rate shows the number of people out of 100,000 that corresponds to $C_{0.5}$, calculated using the empirical relationship between log-transformed SARS-CoV-2 RNA N gene concentrations and laboratory-confirmed COVID-19 incidence rates using 10 merged wells for all POTWs derived in Table S3.

Merged	$C_{0.5}$ (cp/g)	Incidence rate (#/100,000)	% Samples under $C_{0.5}$ (All, June)
1	10 900	3.7 ± 0.1	33%, 81%
2	5 620	2.4 ± 0.1	17%, 46%
3	3 650	1.9 ± 0.1	13%, 38%
4	2 430	1.4 ± 0.1	10%, 29%
5	1 650	1.1 ± 0.1	6.1%, 18%
6	1 180	0.9 ± 0.1	3.1%, 8.8%
7	895	0.8 ± 0.1	1.8%, 5.3%
8	783	0.7 ± 0.1	1.5%, 4.4%
9	772	0.7 ± 0.1	1.5%, 4.4%
10	746	0.7 ± 0.1	1.5%, 4.4%