The dynamic relationship between COVID-19 cases and SARS-CoV-2 wastewater concentrations across time and space: Considerations for model training data sets

Rebecca Schill [a], Kara L. Nelson [b], Sasha Harris-Lovett [c], Rose S. Kantor [b],*

[a] TUM School of Engineering and Design, Technical University of Munich, Germany
[b] Civil and Environmental Engineering, University of California, Berkeley, CA, USA
[c] Berkeley Water Center, University of California, Berkeley, CA, USA

HIGHLIGHTS

- The ratio of wastewater SARS-CoV-2 to COVID-19 cases varied over time and space.
- Simple linear regression models fail to reflect this dynamic relationship.
- Neither normalization of wastewater data nor of case data improved correlations.
- Less SARS-CoV-2 RNA was measured in wastewater per COVID-19 case during Omicron BA.1.
- Viral shedding dynamics, diagnostic testing, and temporal offset are key factors.

GRAPHICAL ABSTRACT
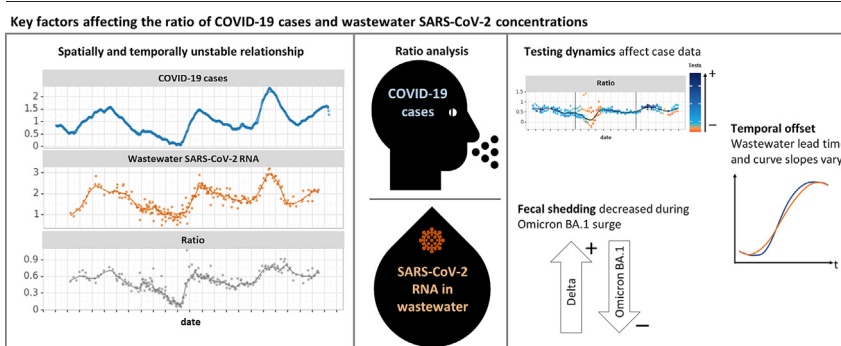
ARTICLE INFO

ABSTRACT

During the COVID-19 pandemic, wastewater-based surveillance has been used alongside diagnostic testing to monitor infection rates. With the decline in cases reported to public health departments due to at-home testing, wastewater data may serve as the primary input for epidemiological models, but training these models is not straightforward. We explored factors affecting noise and bias in the ratio between wastewater and case data collected in 26 sewersheds in California from October 2020 to March 2022. The strength of the relationship between wastewater and case data appeared dependent on sampling frequency and population size, but was not increased by wastewater normalization to flow rate or case count normalization to testing rates. Additionally, the lead and lag times between wastewater and case data varied over time and space, and the ratio of log-transformed individual cases to wastewater concentrations changed over time. This ratio decreased between the Epsilon/Alpha and Delta variant surges of COVID-19 and increased during the Omicron BA.1 variant surge, and was also related to the diagnostic testing rate. Based on this analysis, we present a framework of scenarios describing the dynamics of the case to wastewater ratio to aid in data handling decisions for ongoing modeling efforts.

1. Introduction

The COVID-19 pandemic stimulated worldwide research on how wastewater-based surveillance of SARS-CoV-2 RNA can be used to monitor infections at the population level. Many studies have found strong correlations between SARS-CoV-2 wastewater RNA samples and COVID-19 cases

via diagnostic testing (Bonanno Ferraro et al., 2021; Cluzel et al., 2022; Greenwald et al., 2021; Ho et al., 2021), and routine wastewater surveillance has supported decision-makers in choosing appropriate public health responses (Diamond et al., 2022; Harris-Lovett et al., 2022; Hopkins et al., 2023). With the widespread availability of at-home tests and decreased severity of disease due to vaccination and/or prior infection, the reliability of reported case data has decreased substantially since December 2021 (Usher, 2022). To prepare for new surges due to emerging variants or waning immunity there is a need to build forecasting and nowcasting models that use wastewater data as a main input (Huisman et al., 2022; Soller et al., 2022). For training, these models require high-quality paired retrospective wastewater and diagnostic testing data. However, both the wastewater and case count data in these retrospective datasets are imperfect, necessitating careful consideration of factors contributing to noise and bias prior to modeling.

1.1. Causes of inaccuracies in wastewater data

Concentration of SARS-CoV-2 in wastewater is affected by the number of infected individuals and their respective viral shedding dynamics, but also by precipitation events, infiltration and inflow (McCall et al., 2022), industrial flow contributions, and many other factors (Soller et al., 2022; Wade et al., 2022; de Araújo et al., 2023; Mitranescu et al., 2022). Flow rates at wastewater sampling sites can be used to adjust for dilution, but flow data is not always available, especially for samples collected from manholes or small wastewater treatment facilities where no flow meter is present. Additionally, the heterogeneity of sewage samples and the degradation of SARS-CoV-2 in sewers (Bivins et al., 2020) cannot be accounted for by flow normalization. To address these sources of variability many studies measure cross-assembly phage (crAssphage) or Pepper Mild Mottle Virus (PMMoV) as biological human fecal indicators (Greenwald et al., 2021; Farkas et al., 2019; Langeveld et al., 2021). Physicochemical parameters such as total nitrogen, ammonia, conductivity, total suspended solids (TSS), and biological oxygen demand (Hoar et al., 2022; Yaniv et al., 2021) can also be used to account for variation in wastewater strength, but they may be substantially affected by industrial inputs (Carducci et al., 2020). Although the US CDC has published recommendations on the wastewater sampling process and established a reporting database (CDC, n.d.), there is currently no overall standard for wastewater SARS-CoV-2 sampling and analysis. Thus, the causes of noise need to be considered individually for each dataset.

1.2. Causes of inaccuracies in diagnostic testing data

Diagnostic testing data also includes uncertainty, which may stem from biased allocation of and access to tests across the population, variation in reporting date assigned to each case (e.g. symptom onset, testing date, or date of positive test result), underreporting of at-home test results, and fluctuations in testing rates across time and space (Noh and Danuser, 2021). In 2020, the WHO recommended a threshold of 5 % test positivity as a metric of sufficient testing. However, this threshold is only valid under certain conditions of contact tracing and sufficient testing of symptomatic individuals, and may only reflect the beginning stages of the pandemic (World Health Organization, 2020). Generally, case data may be less reliable when testing rates are low, and as of May 26, 2022, Noh and Danuser (2021) estimated a total rate of undetected cases of approximately 55 % for California (Noh and Danuser, 2021). Modeling testing bias was shown to improve case data accuracy when compared to seroprevalence (Chiu and Ndeffo-Mbah, 2021), but normalization in wastewater testing studies is typically focused only on accounting for wastewater strength. Although the importance of assessing testing rates prior to modeling was demonstrated in a recent study (Daza-Torres et al., 2022), to our knowledge, few wastewater studies have directly addressed bias in diagnostic testing data.

1.3. Correlation and the ratio between wastewater and case data

Prior research has used correlation between wastewater and case data as a readout for the effectiveness of normalization methods, for determination of lead/lag times between datasets, and as a means to state the value of wastewater monitoring in general (Greenwald et al., 2021; Ho et al., 2021; Mitranescu et al., 2022; Feng et al., 2021; Maal-Bared et al., 2023). However, statistical caveats of this analysis are often ignored – for example the fact that the correlation of two variables that measure the same phenomenon in a time series is inflated due to autocorrelation (Hamed, 2009; Wolfe et al., 2021). Critically, the correlation coefficient reflects the global relationship between the diagnostic testing and wastewater surveillance data and does not offer an insight into the development of this relationship over time. For this purpose, the ratio of log-scaled COVID-19 cases over log-scaled wastewater RNA concentrations may be more appropriate. This ratio should be representative of shedding per person assuming perfect diagnostic testing and accurate wastewater data (not accounting for SARS-CoV-2 RNA decay in the sewer). Log-scaling reduces extreme values in the datasets and mimics a linear relationship between the variables, as they are not normally distributed (Hopkins et al., 2023). Several studies have proposed using this ratio for analysis, and have reported values between 0.24 and 0.39, or up to 0.67 after flow normalization (Wolfe et al., 2021; Nourbakhsh et al., 2022; Rodríguez Rasero et al., 2022; Zdenkova et al., 2022). These studies either report an average measured value for this ratio (Wolfe et al., 2021; Zdenkova et al., 2022) or investigate temporal changes based on model data (Nourbakhsh et al., 2022; Rodríguez Rasero et al., 2022), however, time-series analysis of this ratio has not been performed on real-world data.

1.4. Study objectives

The goal of this study was to investigate the nature of the relationship between wastewater and case data over space and time to provide a basis for future modeling efforts. We present a large, curated dataset with wastewater and case data collected in California during the first two years of the COVID-19 pandemic, when case data quality was high. Our analyses reveal the instability of the relationship between wastewater and case counts and identify three main variables that could affect models for predicting cases from wastewater: dynamic lead/lag, changes in fecal shedding due to viral variants, and changes in reporting of individual cases to public health departments.

2. Materials & methods

2.1. Wastewater sample collection and analysis

Raw wastewater samples ($n = 2480$) were collected via 24-h flow- or time-weighted composite samplers from 26 sewersheds in California between 1 and 5 times per week (Tables S1, S2). All sewer systems had separate storm sewers, with the exception of system D, where wastewater and storm sewers were combined. Sampling dates ranged between October 2, 2020 and June 29, 2022, although not all sewersheds were sampled for the full time period. Sample collection points were at wastewater treatment plant influent ("sewersheds") and at pump stations and manholes ("sub-sewersheds"). Samples were aliquoted (40 mL) into tubes containing the 4S method lysis mixture, shipped overnight to UC Berkeley, and analyzed according to the laboratory procedure described by Kantor et al. (2022). Analysis used the 4S method for total RNA extraction (Whitney et al., 2021) followed by RT-qPCR for SARS-CoV-2 CDC N1, Pepper Mild Mottle Virus, and Bovine Coronavirus (Greenwald et al., 2021). Quality controls included extraction negative controls, duplicate extractions, extraction spike-in controls (Bovine Coronavirus), triplicate RT-qPCR reactions, no-template controls, and standard curves, as described in Kantor et al. (2022). Data not passing quality control were removed and were replaced with repeated analyses wherever possible.

2.2. Wastewater data preparation

Wastewater data were preprocessed as previously described (Kantor et al., 2022). Briefly, RT-qPCR outliers were removed, Cq values were

converted to gene copy numbers using an aggregated standard curve, RT-qPCR replicates were combined by taking the geometric mean, and sample weight was used to calculate the gene copies per milliliter of wastewater. Extraction replicates were combined by taking the geometric mean. Five outliers that could be directly attributed to changes in plant operations or autosampler failures were manually removed.

The wastewater concentration was normalized by flow to reduce the effects of dilution by precipitation, groundwater infiltration, and industrial wastewater. Precipitation data for the years 2020–2022 were downloaded from the NOAA (National Oceanic and Atmospheric Administration) website for each county (Menne et al., 2012). Using this dataset, we calculated the median dry flow for each sewershed by taking the median of daily flow rates for all days that were recorded as dry (precipitation <0.2 in.) within the county. We then used this median dry flow to recalculate the SARS-CoV-2 RNA concentration in the wastewater and removed the industrial proportion of flow estimated by the wastewater agencies from the daily flow, as in Eq. (1) (Method 1). A second variation on this method (Method 2) entailed normalizing values only for days on which precipitation occurred. For Method 2, values were normalized according to Eq. (1) to offset a potential dilution and remove the industrial flow proportion, and all other values were normalized according to Eq. (2) to remove only the industrial flow proportion. We tested different time frames of up to three days after rain events to account for potential delays in the effect of precipitation on the dilution of the signal, however, including only the day of the rain event resulted in the highest correlations (not shown).

$$c(normalized) = \frac{c(raw)}{100\% - f_{industrial}} * \frac{q}{q_{dry\ median}} \quad (1)$$

$$c(normalized) = \frac{c(raw)}{100\% - f_{industrial}} \quad (2)$$

where c(normalized) is the flow-normalized RNA concentration (gc/mL), c(raw) is the measured SARS-CoV-2 RNA concentration (gc/mL), q is the daily flow (MGD), $q_{dry\ median}$ is the median dry flow (MGD), $f_{industrial}$ is the percentage of total flow estimated to come from industrial sources.

Normalization with PMMoV, TSS and conductivity was performed according to the following Eq. (3).

$$c(normalized) = c(raw) * \frac{np(median)}{np(raw)} \quad (3)$$

where c(normalized) is the normalized SARS-CoV-2 RNA concentration (gc/mL), c(raw) is the measured SARS-CoV-2 RNA concentration (gc/mL), np(raw) is the concentration of the normalization parameter, and np(median) is the median concentration of the normalization parameter.

Previous studies recommend applying a 7-day or 10-day moving average to the wastewater data (Daza-Torres et al., 2022). However, as the sampling frequencies in our dataset varied over time and space, lowess smoothing and interpolation was chosen for analysis of lag and lead times and for data visualization (Greenwald et al., 2021). The smoothing coefficient alpha was defined as alpha = X/n, where n was the total number of data points at a given site. We note that because lowess smoothing depends on the total number and density of data points, it may have led to slightly different effects on data from different sites. Unless stated, other analyses were performed on the original wastewater dataset to maintain the integrity of the recorded data.

2.3. COVID-19 case data collection and preparation

Masked daily case counts per sewershed were provided by the California Department of Public Health, based on sewershed boundaries provided by wastewater agencies. Cases were attributed to the earlier of 1) the date of diagnostic testing or 2) the reported date of first symptoms, when both dates were available. Sewershed population estimates were based on reports by the wastewater agencies and, if unavailable, government census

data (Table S1). Daily case counts were masked below 3 cases for sewersheds representing populations of 200,000 or less, and below 5 cases for populations of 50,000 or less, but instances of zero cases were reported as zero. During data preparation, masked values were filled with the mean of the masked ranges (Table S3). Case counts and testing rates were normalized to a population of 100,000 and a centered 7-day moving average value was calculated to smooth weekly periodicity. For log-scaled analyses, days with zero average cases were dropped prior to analysis.

2.4. Normalization of case data to account for diagnostic testing rates

County-level diagnostic testing rate data were acquired from publicly available sources (COVID-19 Time-Series Metrics by County and State, 2022). In order to compensate for fluctuations in how accurately the case count data reflected the true incidence of infection, we adjusted the reported cases to the diagnostic testing rates according to the following equations. Eq. (4) linearly inflates the daily cases according to the fraction of utilized testing capacity on a given day. Eq. (5) compensates for a positivity rate bias as defined by Chiu and Ndeffo-Mbah (2021).

$$n'_c = n_c \frac{n_{t\ max}}{n_t} \quad (4)$$

$$n'_c = \frac{p}{n_t^{-0.5}} * 100, \text{where } p = \frac{n_c}{nt} \quad (5)$$

where n_c' is the adjusted number of cases, n_c is the original number of cases, n_t is the number of tests, $n_{t\ max}$ is the maximum number of tests ever reported on a single day within the study period, and p is the test positivity rate. As described in Section 2.3, all values are normalized by sewershed population size.

2.5. Data analysis

The data analysis pipeline was created in Python 3.7 using the Pandas v1.3.5 and Numpy v1.21.6 libraries. The rank correlation of smoothed daily cases with raw and normalized wastewater sampling data was quantified for each sewershed using the Kendall's Tau b coefficient (SciPy v1.7.3) (Greenwald et al., 2021). Autocorrelation and lowess smoothing were calculated using Statsmodels v0.10.2, and data visualization was performed using Plotnine v0.9.0.

For analyses of individual surges, the following timeframes were used: the first major surge we observed (including Epsilon, Alpha, and other minor variants) was defined from the start of the time series (October 2020) to April 15, 2021, the Delta surge from April 16 to November 26, 2021, and the Omicron BA.1 surge from November 27 to March 15, 2022 based on California Department of Public Health (CDPH, n.d.) and COVID-CG (COVID CG, n.d.) and our wastewater sequencing data (unpublished).

To assess the stability of the relationship of log-scaled COVID-19 cases and log-scaled wastewater SARS-CoV-2 concentrations, we implemented a linear regression model with Scikit-Learn v1.13 using these inputs during the Epsilon/Alpha variant surge for sewersheds D1, D2, K, L, and M. This model was then applied to the subsequent Delta and Omicron variant surges and evaluated using the R^2 goodness-of-fit parameter (Table S4). The data analysis pipeline and all necessary datasets are available at GitHub (github.com/RebeccaSchill/WBE).

3. Results and discussion

We analyzed the SARS-CoV-2 RNA concentration in 2480 wastewater samples from 26 sewersheds sampled between 1 and 5 times per week from approximately October 2020–April 2022. This data was paired with sewershed-specific COVID-19 daily case counts and county-level diagnostic testing rates and positivity rates. Populations of the sewersheds ranged from 12,000 to 4 million, and flow rates ranged from 0.2 to 243 million

gallons per day (Table S1). Precipitation was infrequent (0 % - 26 % of days in the time series for each site), due to a combination of drought and mediterranean climate in California.

3.1. Denoising via normalization of wastewater and case data

We first compared methods for removing noise from the wastewater and case data. As previously described, denoising efficacy was evaluated based on changes to Kendall's tau calculated for the relationship between wastewater and case data (Table 1) (Greenwald et al., 2021; Zheng et al., 2022). Flow normalization marginally improved the correlation for 14 sewersheds, but the effect of normalization was minimal, likely because of infrequent precipitation (Table S5). Normalization of the wastewater data from two major sewersheds (D1 and D2) to PMMoV, TSS, or conductivity also did not increase correlations with case data (Table 1, Table S6). Other studies have shown that normalization of wastewater to PMMoV can decrease noise, but successes have been inconsistent and appear to be dependent on the laboratory method used for virus concentration and extraction, as well as sewershed size (Graham et al., 2021; Kim et al., 2021; Nagarkar et al., 2021), and possible dietary variation. Our laboratory method for RNA extraction (4S, Whitney et al., 2021) lacked bead-beating and therefore may not have achieved complete and consistent lysis of PMMoV, required for accurate quantification.

Normalization of the sewershed-level case counts to the county-level diagnostic testing rate (Eq. (4)) reduced the strength of the correlation to wastewater data. Accounting for the test positivity rate in addition to testing rate in a bias function (Eq. (5)) resulted in a more modest decrease in correlation (Table 1). This suggests that additional calibration of the testing bias model (e.g. with regional seroprevalence data) is likely required.

3.2. The correlations between wastewater and case data differed by sewershed

Flow-normalized Kendall's tau for wastewater and case data from different sewersheds exhibited a wide range, from 0.27 to 0.74 (Fig. 1). In general, larger treatment plants with more frequent sampling and less masking of individual case data showed the highest tau values (Table S3). Meanwhile, smaller sewersheds appeared subject to higher noise, for several possible reasons. First, when the total absolute number of infected individuals are low, as is often typical in small sewersheds, each individual contributes a higher fraction of the total wastewater SARS-CoV-2 concentration, and sampling effects (e.g. missing a flush) can create more noise (Wade et al., 2022). Additionally, the effect of mobility (e.g. one infected person entering or leaving the sewershed) (Gudra et al., 2022), variability in personal water use, and dynamic viral shedding profiles (de Araújo et al., 2023) are stronger.

Second, consistent with recommendations from the US CDC (Centers for Disease Control and Prevention, 2022), we found that sewersheds with fewer than 2 samples per week tended to produce weaker correlations, and these were often smaller treatment facilities. This is in line with reports of lower sampling capacities at smaller wastewater treatment plants (Hill et al., 2022). Many of these small sewersheds also had fewer than 50 total sampling events (Table S1). Additionally, we note that the time series of wastewater and case counts were autocorrelated (Durbin-Watson statistic d < 1.5 in all sewersheds), and autocorrelation may have increased with increasing sampling frequency, affecting tau values differently in each sewershed.

Third, a larger proportion of daily case data was masked in the smallest sewersheds, with a median value of 45 % of data masked (Table S3). Two sewersheds with a masking proportion > 95 % were removed from further analyses. Lastly, within-sewershed fluctuations in diagnostic testing rates may also have led to differences in wastewater-case correlation between sewersheds (Lieberman-Cribbin et al., 2020) (see Section 3.6). We were unable to assess disparities in testing rates given that testing rate data were available at the county level only, which may not be representative of individual sewersheds. Overall, the collection of high-resolution datasets improves the reliability of the relationship between case counts and wastewater data and the accuracy of forecasting models (Vaughan et al., 2023). These findings motivate policy to report detailed diagnostic testing and COVID-19 case data and to provide support for smaller communities to increase wastewater sampling frequency in locations where case data may be the least accurate (Medina et al., 2022).

3.3. Lag between case data and wastewater data was dynamic over time and space

Modeling work may need to take into consideration the lead/lag between case counts and wastewater data. As previous studies have reported wastewater lead times over case data of between 0 and 14 days (Olesen et al., 2021), we hypothesized that lead time could vary substantially by sewershed and over time due to factors such as evolving virus variants, sewage travel distance, and access to and frequency of diagnostic testing (Bibby et al., 2021). To assess lead/lag times, we first smoothed the wastewater data to remove noise (see Methods; Section 2.2), then calculated the cross-correlation Kendall's tau-b between the flow-normalized wastewater data and case data shifted in each direction by 1 to 14 days (see Methods; Fig. S1A). A wastewater lag/lead time between −3 days and + 4 days was detected in four of the seven sewersheds that were sampled throughout the entire time series, but the corresponding increases in Kendall's tau-b were very low with a maximum increase of 3 %.

We next examined whether the wastewater lead/lag changed during periods when different variants predominated. Overall, varying wastewater lead times from + 1 to + 13 days were observed in 12 of 17 sewersheds during the Epsilon/Alpha variant-dominated surge. This lead time was also observed in 14 out of 20 sewersheds during the Delta variant surge but faded during the Omicron variant surge, where wastewater data lagged and led case data in an equal number of sewersheds (Fig. S1B, S1C, S1D). The dynamic behavior of the time shift between wastewater and case data across variants is demonstrated in detail at two sewersheds (D1 and K) in Fig. 2. Notably, during the first surge we observed, the peaks in wastewater and case data are not aligned, resulting in very long lead times. This is likely due to a combination of testing fluctuations over the winter holidays and the multiple overlapping surges of different variants (Epsilon, Alpha, Gamma, and others). The wastewater lead time lessened significantly during the Delta surge in both sewersheds and disappeared during the Omicron surge. This aligns with previous reports of reduced wastewater lead times after the Alpha surge (Hopkins et al., 2023; Xiao et al., 2022; de Araújo et al., 2022).

The use of cross-correlation to determine lag/lead times between case data and wastewater data is based on the assumption that there is a static lag between the two datasets. Static lag could reasonably stem from near-constant factors such as sewer transit time (constant within a sewershed) or delay between infection and symptom onset that would trigger
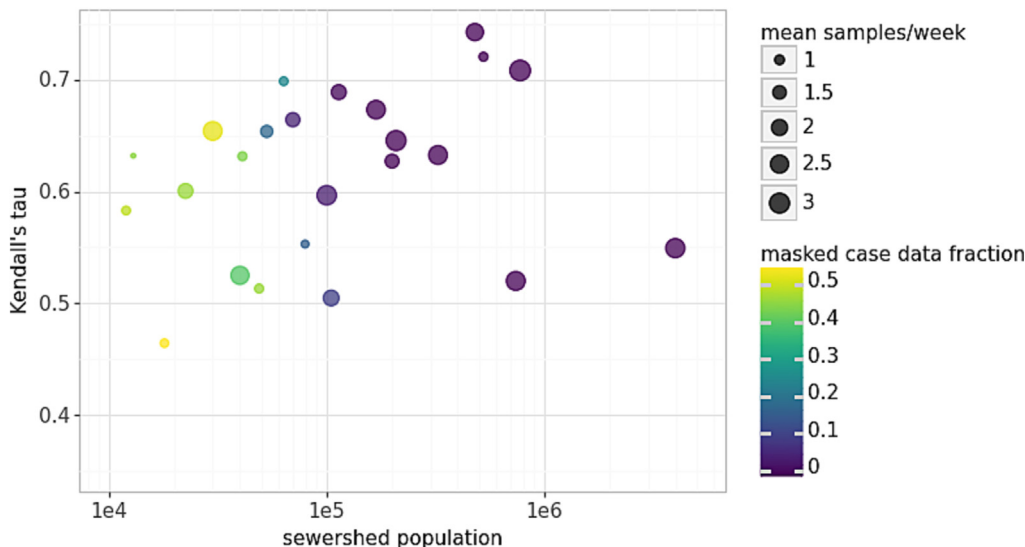
Table 1

Average Kendall's correlation of log-scaled values (for all chosen sewersheds) before and after applying different normalization methods to 7-day moving average case data and unsmoothed wastewater data. Averages represent all 26 sewersheds.

Case data normalization methods	Wastewater data normalization methods			
	No normalization	Flow (Method 1, Eq. (1))	Flow (Method 2, Eqs. (1) & (2))	PMMoV
No normalization	0.57	0.58	0.57	0.47
Normalized by testing capacity (Eq. (4))	0.50	0.50	0.49	0.12
Normalized by testing bias (Eq. (5))	0.54	0.54	0.53	0.14

Fig. 1. The Kendall's correlation coefficient between log-scaled cases and unsmoothed log-scaled wastewater SARS-CoV-2 RNA concentrations (y-axis) increased with increasing sewershed population (x-axis) and was affected by weekly sampling frequency (point size) and by the fraction of case data that was masked (point color).
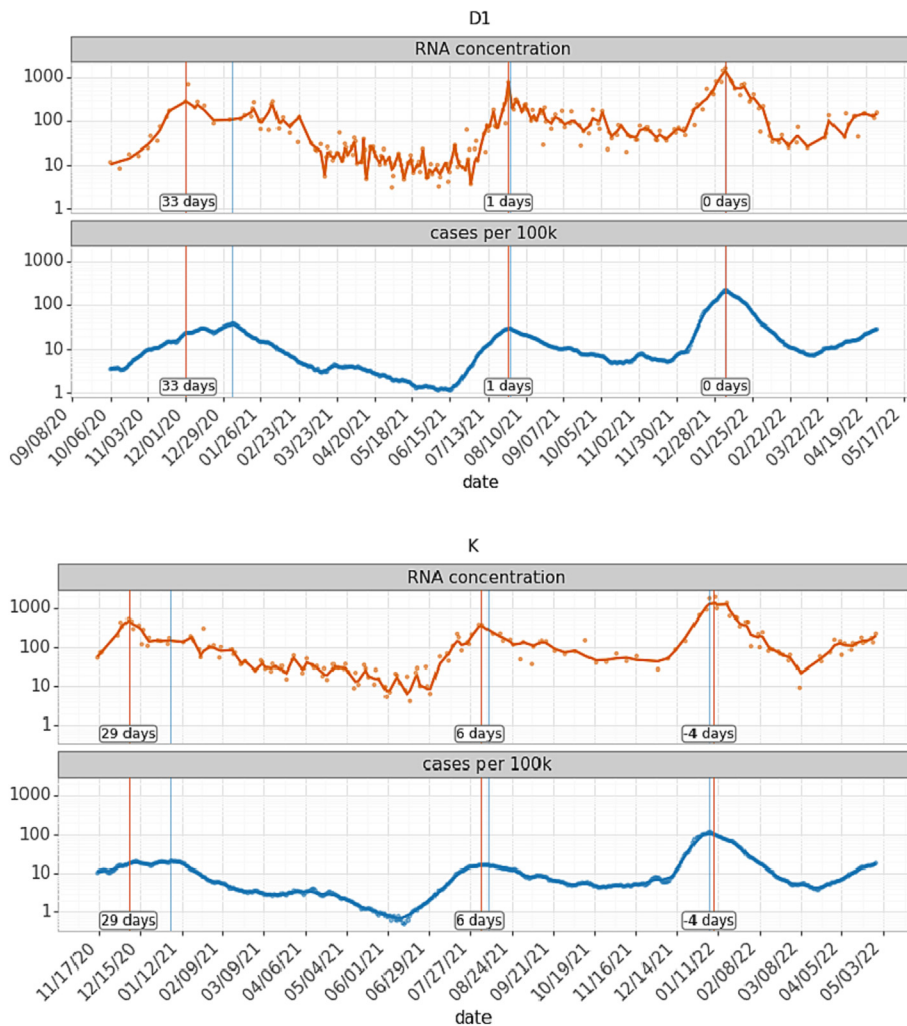

Fig. 2. Lag between wastewater and individual case data varied over time and between sewersheds. Flow-adjusted wastewater SARS-CoV-2 concentrations in gene copies per milliliter (orange), COVID-19 cases per 100,000 people (blue), including a lowess smoother (alpha = 0.05) are shown for two sewersheds. Sewershed D1 (top) represents a population of 750,000, with high sampling frequency, while sewershed K (bottom) represents 480,000 people with intermittently reduced sampling frequency. Vertical lines indicate minima and maxima of the timeseries of cases (blue) and wastewater (orange). Labels indicate the wastewater lead time at the surge peak in days.

diagnostic testing (assumed constant for each variant). However, our findings of dynamic lag over time and across sewersheds suggest that other factors are at play. Wastewater sampling frequency, population-level immunity, or changes in diagnostic testing strategy/availability differed between surges and locations and likely affected lead times. Previous studies have highlighted that lead time calculations need to be adapted to different purposes, for example real-time decision-making versus retrospective data analysis (Olesen et al., 2021). In this study, due to the applied smoothing methods, lag calculations do not represent real-time data availability, but instead reveal a potential delay in measurable signal between wastewater and diagnostic testing. Our findings of dynamic lead times suggest that cross-correlation, and by extension, simple linear regression models (Table S4), are therefore insufficient for describing the relationship between case and wastewater data for retrospective data analysis, and dynamic lead times will affect input data for modeling.

3.4. The ratio of cases per wastewater RNA was not constant over time and space

To explore the dynamic nature of the relationship between wastewater and case data, we calculated the ratio of log(cases) per log(wastewater concentration) (see Fig. S2 for example). For five large sewersheds sampled continuously throughout the analyzed time frame (Fig. 3), we found that the magnitude of the ratio was different at each sewershed, likely affected by the accuracy of the population estimates. The ratio also changed over time: during the Epsilon/Alpha surge, the ratio remained stable overall before decreasing to a minimum just before the peak of the Delta surge. The ratio then recovered and increased to a maximum during the first Omicron surge. Towards the end of this surge, the ratio decreased once more. These developments were similar at sewersheds D1, D2, and K, but less

pronounced or more stochastic in sewersheds L and M and others where sampling was less frequent (Fig. 3; see Fig. S3 for all sewersheds).

3.5. Ratio of cases per wastewater RNA differed by variant

To test for the effect of evolving virus variants on wastewater surveillance data, we calculated point estimates for the ratios in each sewershed as follows: for each variant surge, we identified the maximum number of cases per 100,000 people (centered 7-day average) and the maximum lowess-smoothed wastewater concentration reported. Then, we calculated the ratio by dividing the log-scaled maximum cases by the log-scaled maximum wastewater concentration. We could not isolate the Epsilon and Alpha variants, as the surges partially coincided. Although our analysis was limited to 5 sewersheds, we observed a significant increasing trend in the ratio from the Delta variant to the Omicron variant (Fig. 4). The decrease between the Epsilon/Alpha and Delta variants could be observed as well but the difference was not statistically significant. The ratio appears consistently lower for sewershed K but sewershed-specific differences were not significant (Kruskall-Wallis, $p = 0.429$).

Additionally, we found that a linear regression model trained to predict case data from flow-normalized unsmoothed wastewater data for Alpha/Epsilon surge deteriorated in fit during the Delta and Omicron BA.1 variants (see Table S4). Overall, these findings agree with the reports of increased fecal and oro-nasopharyngeal viral loads during the Delta surge (Li et al., 2022; Prasek et al., 2022) and with reduced fecal shedding observed with the Omicron variant (Bloemen et al., 2022; Yuan et al., 2022). While SARS-CoV-2 oro-nasopharyngeal viral load was reportedly reduced after vaccination (Bramante et al., 2022; Levine-Tiefenbrun et al., 2021; McEllistrem et al., 2021), more research is needed to determine potential effects of vaccination and prior infection on fecal shedding rates.
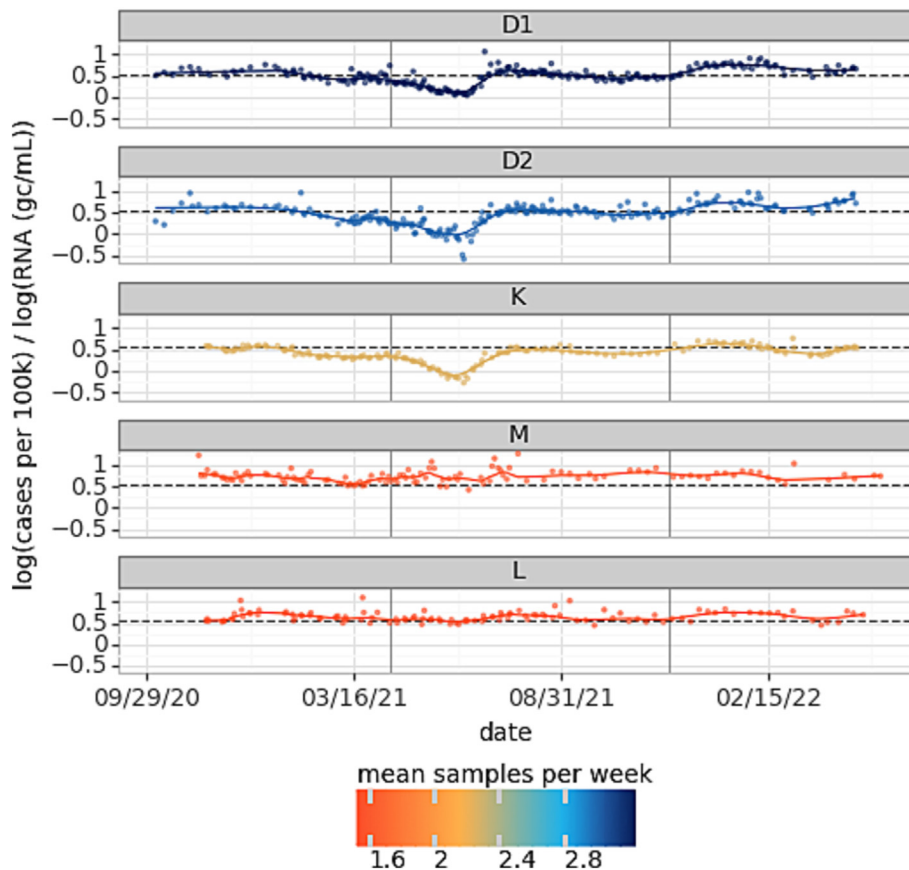

Fig. 3. Ratio of log(cases per 100,000) over log(wastewater concentration) at sewersheds D1 ($n = 262$), D2 ($n = 244$), K ($n = 172$), L ($n = 121$), and M ($n = 131$), for three surges (separated by vertical gray lines). Smoothed lines were generated with lowess (alpha = 0.05), and the dashed line represents the median ratio across all 5 sewersheds (0.55). For each sewershed shown, case data masking was below 5 % of all data points. Two outliers were removed at sewershed M for visualization.
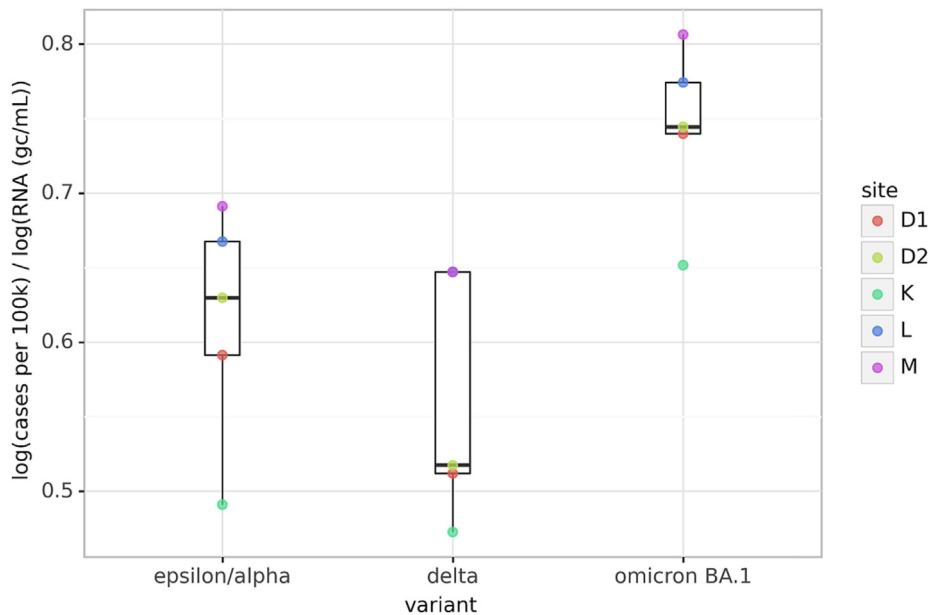
Fig. 4. The ratio of log(cases) over log(wastewater RNA) changes for each variant at sewersheds D1, D2, K, L, and M. The ratio was calculated from the peaks of the three surges after lowess smoothing (alpha = 5 / total samples). The difference between the Omicron BA.1 variant and the other variants was statistically significant (Mann-Whitney, * $p = 0.032$, ** $p = 0.008$), while the difference between the Epsilon/Alpha and other variants to the Delta variant was not ($p = 0.421$).

3.6. Diagnostic testing rates influenced the cases-to-wastewater RNA ratio

Given the drop in the cases-to-wastewater RNA ratio during the pre-Delta period, we hypothesized that changes in diagnostic testing dynamics might influence this ratio. Indeed, we observed that low testing rates corresponded with low ratios throughout the time series, including during the pre-Delta trough (Fig. 5). Correlations between cases-to-wastewater ratios and diagnostic testing rates over time were significant in several sewersheds (Table S7). Notably, the strengths of these relationships

differed for sewersheds in the same county (Table S7), suggesting that sewershed-level testing rates differed from those at the county level or the quality of wastewater data differed for sewersheds in the same county. Additionally, the measurement uncertainty was likely higher during periods of low case counts and low wastewater concentrations, which could also have contributed to the change in the ratio observed during these periods.

Although low diagnostic testing rates may partially explain low ratios between surges, we note that the cases-to-wastewater RNA ratio recovers more quickly than the testing rates, suggesting that undertesting cannot
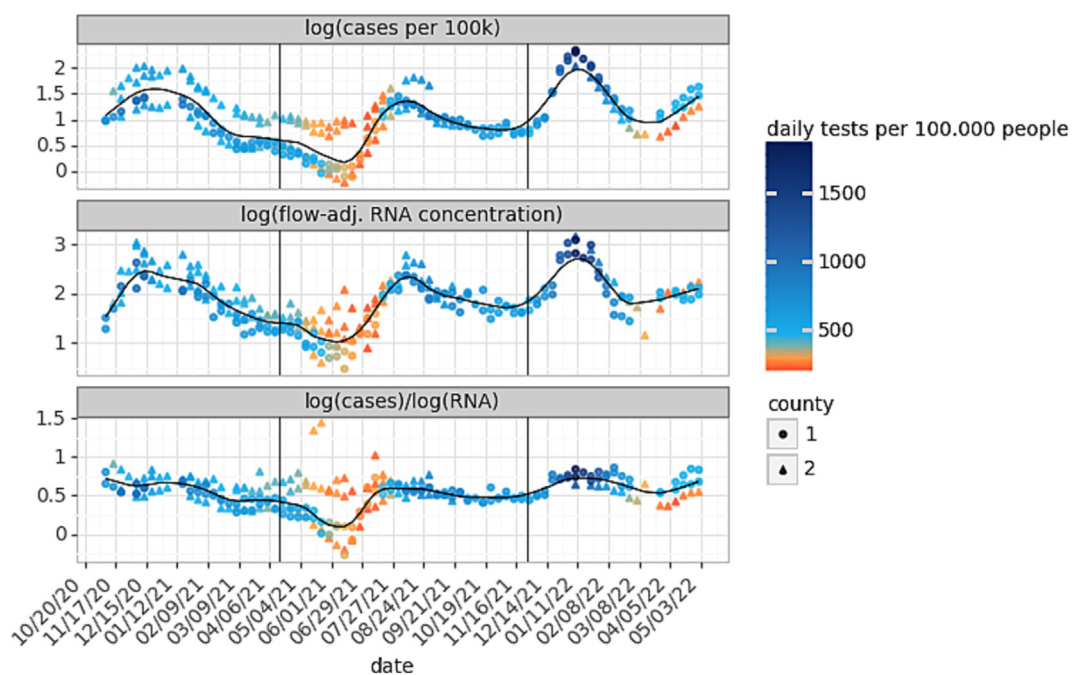

Fig. 5. Time series from 5 sewersheds (D1, D2, K, L, M) of the median weekly COVID-19 cases per day (top), median weekly flow-adjusted wastewater SARS-CoV-2 RNA (gc/mL) (middle) and the ratio between them (bottom). For each sewershed, for a given week, a minimum of two data points was required for a weekly median to be shown. The color of the data points represents the daily diagnostic testing rate per 100,000 people, and the shape indicates the county in which the sewershed is located.

Table 2
Scenarios that can lead to changes in the ratio of log transformed case to wastewater data.

Ratio	Changes to factors contributing to case data	Changes to factors contributing to wastewater signal	Both
Increase	Increased diagnostic testing rates without proportional increase in incidence	Increase in incidence that is captured in the case data but is not reflected in wastewater data due to lower shedding rate or duration	Tested cases increase before this is reflected in the wastewater data
Decrease	Increase in incidence reflected in wastewater data is not reported in case data due to undertesting	Higher wastewater values due to increased shedding rate or duration	Tested cases increase more steeply than is reflected in the wastewater data. Wastewater values increase before cases reflect the change (undertesting)

be the only cause of lower ratios (Fig. 5). This is underlined by the fact that normalizing by testing rates (via Eqs. (4) and (5)) did not completely flatten the cases-to-wastewater ratio over time (Fig. S4). Critically, differences in the slopes of the case and wastewater curves may also have affected the ratio between them. As has been shown in previous studies (Daza-Torres et al., 2022), we suggest that proportionally more cases remained undetected at the very beginning of a surge until the diagnostic testing rates adapted, as the case curves increased more steeply than the wastewater curves before the peak of each surge (Fig. S5). This affected the ratio as well (Table 2). After the peak of each surge, the decline in wastewater RNA concentrations was more gradual than the decline in cases, perhaps due to prolonged fecal shedding (Zhang et al., 2021). We echo the suggestion by Daza-Torres et al. (2022), that input data for modeling should be drawn from time periods with adequate testing. Future work could assess testing behavior and the distribution of tests across the population (e.g. symptomatic vs. asymptomatic, retesting, etc.), to further adjust case data.

4. Conclusions

Based on our observations, we define a framework of key factors that may affect the cases-to-wastewater SARS-CoV-2 RNA ratio over time, encompassing variation in diagnostic testing rates, changes to fecal shedding, and fluctuations in the temporal off-set between wastewater surveillance and case count data (Table 2). Future efforts could model these factors to come to a more accurate understanding of the ground truth case counts. Additional work could also incorporate hospitalization (Hopkins et al., 2023), vaccination, mobility and other data types that were not considered here. Importantly, modeling work should ensure that the case data used to train a predictive model are drawn from a period (s) when testing was adequate (Daza-Torres et al., 2022). Additionally, we found that within our dataset, wastewater data varied in quality, and our analysis was limited by the changing frequency of wastewater sample collection throughout each time series. Thus, wastewater data to be used for modeling requires careful curation and potentially smoothing. However, we observed that smoothing led to the loss of extreme values, many of which were important maxima and minima. Smoothing can hide or delay rapid changes in the time series, affecting the lead/lag between wastewater data and case counts. Future work should compare raw and smoothed model inputs to ensure that smoothing maintains the integrity of the data.

Looking forward, once sewershed-specific models have been established, subsequent modeling will benefit from the fact that the sewersheds themselves will remain relatively consistent: the structure of the sewer system itself, transit time of sewage, noise from precipitation, and industrial discharge can be taken into account with ongoing data and existing normalization methods. These models will be independent from case counts and testing, and thus independent from lead/lag times relative to cases. The key factor subject to change will be fecal shedding rate and duration (Table 2). Models will require updated in vivo studies of fecal shedding profiles to adjust for new SARS-CoV-2 variants and evolving immunity in the population. Overall, this approach will reduce uncertainty to an extent. However, wastewater surveillance is limited in that uncertainties cannot be completely eliminated and they may still prevent a quantitative prediction of COVID-19 cases from SARS-CoV-2 concentrations in wastewater.

Thus, the largest potential of WBE may lie in trend analysis and detection of variants of concern, as is currently implemented by the CDC National Wastewater Surveillance System (Centers for Disease Control and Prevention, 2023).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2023.162069.

References

de Araújo, J.C., Mota, V.T., Teodoro, A., Leal, C., Leroy, D., Madeira, C., Machado, E.C., Dias, M.F., Souza, C.C., Coelho, G., et al., 2022. Long-term monitoring of SARS-CoV-2 RNA in sewage samples from specific public places and STPs to track COVID-19 spread and identify potential hotspots. Sci. Total Environ. 838, 155959. https://doi.org/10.1016/j.scitotenv.2022.155959.

de Araújo, J.C., Madeira, C.L., Bressani, T., Leal, C., Leroy, D., Machado, E.C., Fernandes, L.A., Espinosa, M.F., Freitas, G.T.O., Leão, T., et al., 2023. Quantification of SARS-CoV-2 in wastewater samples from hospitals treating COVID-19 patients during the first wave of the pandemic in Brazil. Sci. Total Environ. 860, 160498. https://doi.org/10.1016/j.scitotenv.2022.160498.

Bibby, K., Bivins, A., Wu, Z., North, D., 2021. Making waves: plausible lead time for wastewater based epidemiology as an early warning system for COVID-19. Water Res. 202, 117438. https://doi.org/10.1016/j.watres.2021.117438.

Bivins, A., Greaves, J., Fischer, R., Yinda, K.C., Ahmed, W., Kitajima, M., Munster, V.J., Bibby, K., 2020. Persistence of SARS-CoV-2 in water and wastewater. Environ. Sci. Technol. Lett. 7, 937–942. https://doi.org/10.1021/acs.estlett.0c00730.

Bloemen, M., Delang, L., Rector, A., Raymenants, J., Thibaut, J., Pussig, B., Fondu, L., Aertgeerts, B., Van Ranst, M., Van Geet, C., et al., 2022. Detection of SARS-COV-2 variants of concern in wastewater of Leuven, Belgium. medRxiv https://doi.org/10.1101/2022.05.12.22274823.

Bonanno Ferraro, G., Veneri, C., Mancini, P., Iaconelli, M., Suffredini, E., Bonadonna, L., Lucentini, L., Bowo-Ngandji, A., Kengne-Nde, C., Mbaga, D.S., et al., 2021. A state-of-the-art scoping review on SARS-CoV-2 in sewage focusing on the potential of wastewater surveillance for the monitoring of the COVID-19 pandemic. Food Environ. Virol. https://doi.org/10.1007/s12560-021-09498-6.

Bramante, C.T., Proper, J.L., Boulware, D.R., Karger, A.B., Murray, T., Rao, V., Hagen, A., Tignanelli, C.J., Puskarich, M., Cohen, K., et al., 2022. Vaccination against SARS-CoV-2 is associated with a lower viral load and likelihood of systemic symptoms. Open ForumInfect. Dis. 9, ofac066. https://doi.org/10.1093/ofid/ofac066.

Carducci, A., Federigi, I., Liu, D., Thompson, J.R., Verani, M., 2020. Making waves: coronavirus detection, presence and persistence in the water environment: state of the art and knowledge needs for public health. Water Res. 179, 115907. https://doi.org/10.1016/j.watres.2020.115907.

CDC, .. Wastewater Surveillance Testing Methods.Available online https://www.cdc.gov/healthywater/surveillance/wastewater-surveillance/testing-methods.html (accessed on 14 November 2022).

CDPH, .. COVID-19 variant data.Available online https://data.chhs.ca.gov/dataset/covid-19-variant-data (accessed on 17 November 2022).

Centers for Disease Control and Prevention, 2022. Developing a Wastewater Surveillance Sampling Strategy.

Centers for Disease Control and Prevention, 2023. COVID Data Tracker: Wastewater Surveillance.Available onlineUS Department of Health and Human Services, CDC, Atlanta, GA (accessed on 17 November 2022).

Chiu, W.A., Ndeffo-Mbah, M.L., 2021. Using test positivity and reported case rates to estimate state-level COVID-19 prevalence and seroprevalence in the United States. PLoS Comput. Biol. 17, e1009374. https://doi.org/10.1371/journal.pcbi.1009374.

Cluzel, N., Courbariaux, M., Wang, S., Moulin, L., Wurtzer, S., Bertrand, I., Laurent, K., Monfort, P., Gantzer, C., Guyader, S.L., et al., 2022. A nationwide indicator to smooth and normalize heterogeneous SARS-CoV-2 RNA data in wastewater. Environ. Int. 158, 106998. https://doi.org/10.1016/j.envint.2021.106998.

COVID CG, .. Available online https://covidcg.org/ (accessed on 22 November 2022).

COVID-19 Time-Series Metrics by County and State, 2022. COVID-19 Time-series Metrics by County And State.

Daza-Torres, M.L., Montesinos-López, J.C., Kim, M., Olson, R., Bess, C.W., Rueda, L., Susa, M., Tucker, L., García, Y.E., Schmidt, A.J., et al., 2022. Model training periods impact estimation of COVID-19 incidence from wastewater viral loads. Sci. Total Environ., 159680 https://doi.org/10.1016/j.scitotenv.2022.159680.

Diamond, M.B., Keshaviah, A., Bento, A.I., Conroy-Ben, O., Driver, E.M., Ensor, K.B., Halden, R.U., Hopkins, L.P., Kuhn, K.G., Moe, C.L., et al., 2022. Wastewater surveillance of pathogens can inform public health responses. Nat. Med. 28, 1992–1995. https://doi.org/10.1038/s41591-022-01940-x.

Farkas, K., Adriaenssens, E.M., Walker, D.I., McDonald, J.E., Malham, S.K., Jones, D.L., 2019. Critical evaluation of CrAssphage as a molecular marker for human-derived wastewater contamination in the aquatic environment. Food Environ. Virol. 11, 113–119. https://doi.org/10.1007/s12560-019-09369-1.

Feng, S., Roguet, A., McClary-Gutierrez, J.S., Newton, R.J., Kloczko, N., Meiman, J.G., McLellan, S.L., 2021. 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 1, 1955–1965. https://doi.org/10.1021/acsestwater.1c00160.

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., et al., 2021. SARS-CoV-2 RNA in wastewater settled solids is associated with COVID-19 cases in a large urban sewershed. Environ. Sci. Technol. 55, 488–498. https://doi.org/10.1021/acs.est.0c06191.

Greenwald, H.D., Kennedy, L.C., Hinkle, A., Whitney, O.N., Fan, V.B., Crits-Christoph, A., Harris-Lovett, S., Flamholz, A.I., Al-Shayeb, B., Liao, L.D., et al., 2021. Tools for interpretation of wastewater SARS-CoV-2 temporal and spatial trends demonstrated with data collected in the San Francisco Bay Area. Water Res. X 12, 100111. https://doi.org/10.1016/j.wroa.2021.100111.

Gudra, D., Dejus, S., Bartkevics, V., Roga, A., Kalnina, I., Strods, M., Rayan, A., Kokina, K., Zajakina, A., Dumpis, U., et al., 2022. Detection of SARS-CoV-2 RNA in wastewater and importance of population size assessment in smaller cities: an exploratory case study from two municipalities in Latvia. Sci. Total Environ. 823, 153775. https://doi.org/10.1016/j.scitotenv.2022.153775.

Hamed, K.H., 2009. Effect of persistence on the significance of Kendall's tau as a measure of correlation between natural time series. Eur. Phys. J. Spec. Top. 174, 65–79. https://doi.org/10.1140/epjst/e2009-01090-x.

Harris-Lovett, S., Nelson, K.L., Kantor, R., Korfmacher, K.S., 2022. Wastewater surveillance to inform public health decision making in residential institutions. J. Public Health Manag. Pract. https://doi.org/10.1097/PHH.0000000000001636.

Hill, D.T., Cousins, H., Dandaraw, B., Faruolo, C., Godinez, A., Run, S., Smith, S., Willkens, M., Zirath, S., Larsen, D.A., 2022. Wastewater treatment plant operators report high capacity to support wastewater surveillance for COVID-19 across New York state, USA. Sci. Total Environ. 837, 155664. https://doi.org/10.1016/j.scitotenv.2022.155664.

Ho, J., Stange, C., Suhrborg, R., Wurzbacher, C., Drewes, J.E., Tiehm, A., 2021. SARS-CoV-2 wastewater surveillance in Germany: long-term PCR monitoring, suitability of primer/probe combinations and biomarker stability. Water Res. 210, 117977. https://doi.org/10.1016/j.watres.2021.117977.

Hoar, C., Li, Y., Silverman, A.I., 2022. Assessment of commonly measured wastewater parameters to estimate sewershed populations for use in wastewater-based epidemiology: insights into population dynamics in New York City during the COVID-19 pandemic. ACS EST Water, acsestwater.2c00052 https://doi.org/10.1021/acsestwater.2c00052.

Hopkins, L., Persse, D., Caton, K., Ensor, K., Schneider, R., McCall, C., Stadler, L.B., 2023. City-wide wastewater SARS-CoV-2 levels strongly correlated with multiple disease surveillance indicators and outcomes over three COVID-19 waves. Sci. Total Environ. 855, 158967. https://doi.org/10.1016/j.scitotenv.2022.158967.

Huisman, J.S., Scire, J., Caduff, L., Xavier, Fernandez-Cassi, Ganesanandamoorthy, P., Kull, A., Scheidegger, A., Stachler, E., Boehm, A.B., Hughes, B., et al., 2022. Wastewater-based estimation of the effective reproductive number of SARS-CoV-2. Environ. Health Perspect. 130, 057011. https://doi.org/10.1289/EHP10050.

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., 2022. Operationalizing a routine wastewater monitoring laboratory for SARS-CoV-2. PLOS Water 1, e0000007. https://doi.org/10.1371/journal.pwat.0000007.

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., et al., 2021. SARS-CoV-2 RNA is enriched by orders of magnitude in solid relative to liquid wastewater at publicly owned treatment works. Environ. Sci.: Water Res. Technol. https://doi.org/10.1039/d1ew00826a.

Langeveld, J., Schilperoort, R., Heijnen, L., Elsinga, G., Schapendonk, C.E.M., Fanoy, E., de Schepper, E.I.T., Koopmans, M.P.G., de Graaf, M., Medema, G., 2021. Normalisation of SARS-CoV-2 concentrations in wastewater: the use of flow, conductivity and CrAssphage. Sci. Total Environ. 865, 161196. https://doi.org/10.1016/j.scitotenv.2022.161196.

Levine-Tiefenbrun, M., Yelin, I., Katz, R., Herzel, E., Golan, Z., Schreiber, L., Wolf, T., Nadler, V., Ben-Tov, A., Kuint, J., et al., 2021. Initial report of decreased SARS-CoV-2 viral load after inoculation with the BNT162b2 vaccine. Nat. Med. 27, 790–792. https://doi.org/10.1038/s41591-021-01316-7.

Li, B., Deng, A., Li, K., Hu, Y., Li, Z., Shi, Y., Xiong, Q., Liu, Z., Guo, Q., Zou, L., et al., 2022. Viral infection and transmission in a large, well-traced outbreak caused by the SARS-CoV-2 Delta variant. Nat. Commun. 13, 460. https://doi.org/10.1038/s41467-022-28089-y.

Lieberman-Cribbin, W., Tuminello, S., Flores, R.M., Taioli, E., 2020. Disparities in COVID-19 testing and positivity in New York City. Am. J. Prev. Med. 59, 326–332. https://doi.org/10.1016/j.amepre.2020.06.005.

Maal-Bared, R., Qiu, Y., Li, Q., Gao, T., Hrudey, S.E., Bhavanam, S., Ruecker, N.J., Ellehoj, E., Lee, B.E., Pang, X., 2023. Does normalization of SARS-CoV-2 concentrations by pepper mild mottle virus improve correlations and Lead time between wastewater surveillance and clinical data in Alberta (Canada): comparing twelve SARS-CoV-2 normalization approaches. Sci. Total Environ. 856, 158964. https://doi.org/10.1016/j.scitotenv.2022.158964.

McCall, C., Fang, Z.N., Li, D., Czubai, A.J., Juan, A., LaTurner, Z.W., Ensor, K., Hopkins, L., Bedient, P.B., Stadler, L.B., 2022. Modeling SARS-CoV-2 RNA degradation in small and large sewersheds. Environ. Sci. Water Res. Technol. 8, 290–300. https://doi.org/10.1039/D1EW00717C.

McEllistrem, M.C., Clancy, C.J., Buehrle, D.J., Lucas, A., Decker, B.K., 2021. Single dose of an MRNA severe acute respiratory syndrome coronavirus 2 (SARS-Cov-2) vaccine is associated with lower nasopharyngeal viral load among nursing home residents with asymptomatic coronavirus disease 2019 (COVID-19). Clin. Infect. Dis. 73, e1365–e1367. https://doi.org/10.1093/cid/ciab263.

Medina, C.Y., Kadonsky, K.F., Roman, F.A., Tariqi, A.Q., Sinclair, R.G., D'Aoust, P.M., Delatolla, R., Bischel, H.N., Naughton, C.C., 2022. The need of an environmental justice approach for wastewater based epidemiology for rural and disadvantaged communities: a review in California. Curr. Opin. Environ. Sci. Health 27, 100348. https://doi.org/10.1016/j.coesh.2022.100348.

Menne, M.J., Durre, I., Korzeniewski, B., McNeill, S., Thomas, K., Yin, X., Anthony, S., Ray, R., Vose, R.S., Gleason, B.E., et al., 2012. Global Historical Climatology Network - Daily (GHCN-Daily), Version 3.

Mitranescu, A., Uchaikina, A., Kau, A.-S., Stange, C., Ho, J., Tiehm, A., Wurzbacher, C., Drewes, J.E., 2022. Wastewater-based epidemiology for SARS-CoV-2 biomarkers: evaluation of normalization methods in small and large communities in southern Germany. ACS EST Water https://doi.org/10.1021/acsestwater.2c00306.

Nagarkar, M., Keely, S.P., Jahne, M., Wheaton, E., Hart, C., Smith, B., Garland, J., Varughese, E.A., Braam, A., Wiechman, B., et al., 2021. SARS-CoV-2 monitoring at three sewersheds of different scales and complexity demonstrates distinctive relationships between waste-water measurements and COVID-19 case data. Sci. Total Environ., 151534 https://doi. org/10.1016/j.scitotenv.2021.151534.

Noh, J., Danuser, G., 2021. Estimation of the fraction of COVID-19 infected people in U.S. states and countries worldwide. PLOS ONE 16, e0246772. https://doi.org/10.1371/jour-nal.pone.0246772.

Nourbakhsh, S., Fazil, A., Li, M., Mangat, C.S., Peterson, S.W., Daigle, J., Langner, S., Shurgold, J., D'Aoust, P., Delatolla, R., et al., 2022. A wastewater-based epidemic model for SARS-CoV-2 with application to three Canadian cities. Epidemics 39, 100560. https://doi.org/10.1016/j.epidem.2022.100560.

Olesen, S.W., Imakaev, M., Duvallet, C., 2021. Making waves: defining the lead time of wastewater-based epidemiology for COVID-19. Water Res. 202, 117433. https://doi. org/10.1016/j.watres.2021.117433.

Prasek, S.M., Pepper, I.L., Innes, G.K., Slinski, S., Ruedas, M., Sanchez, A., Brierley, P., Betancourt, W.Q., Stark, E.R., Foster, A.R., et al., 2022. Population level SARS-CoV-2 fecal shedding rates determined via wastewater-based epidemiology. Sci. Total Environ. 838, 156535. https://doi.org/10.1016/j.scitotenv.2022.156535.

Rodríguez Rasero, F.J., Moya Ruano, L.A., Rasero Del Real, P., Cuberos Gómez, L., Lorusso, N., 2022. Associations between SARS-CoV-2 RNA concentrations in wastewater and COVID-19 rates in days after sampling in small urban areas of Seville: a time series study. Sci. Total Environ. 806, 150573. https://doi.org/10.1016/j.scitotenv.2021.150573.

Soller, J., Jennings, W., Schoen, M., Boehm, A., Wigginton, K., Gonzalez, R., Graham, K.E., McBride, G., Kirby, A., Mattioli, M., 2022. Modeling infection from SARS-CoV-2 wastewa-ter concentrations: promise, limitations, and future directions. J. Water Health 20, 1197–1211. https://doi.org/10.2166/wh.2022.094.

Usher, A.D., 2022. FIND documents dramatic reduction in COVID-19 testing. Lancet Infect. Dis. 22, 949. https://doi.org/10.1016/S1473-3099(22)00376-0.

Vaughan, L., Zhang, M., Gu, H., Rose, J.B., Naughton, C.C., Medema, G., Allan, V., Roiko, A., Blackall, L., Zamyadi, A., 2023. An exploration of challenges associated with machine learning for time series forecasting of COVID-19 community spread using wastewater-based epidemiological data. Sci. Total Environ. 858, 159748. https://doi.org/10.1016/ j.scitotenv.2022.159748.

Wade, M.J., Lo Jacomo, A., Armenise, E., Brown, M.R., Bunce, J.T., Cameron, G.J., Fang, Z., Farkas, K., Gilpin, D.F., Graham, D.W., et al., 2022. Understanding and managing uncer-tainty and variability for wastewater monitoring beyond the pandemic: lessons learned from the United Kingdom national COVID-19 surveillance programmes. J. Hazard. Mater. 424, 127456. https://doi.org/10.1016/j.jhazmat.2021.127456.

Whitney, O.N., Kennedy, L.C., Fan, V.B., Hinkle, A., Kantor, R., Greenwald, H., Crits-Christoph, A., Al-Shayeb, B., Chaplin, M., Maurer, A.C., et al., 2021. Sewage, salt, silica, and SARS-CoV-2 (4S): an economical kit-free method for direct capture of SARS-CoV-2 RNA from wastewater. Environ. Sci. Technol. 55, 4880–4888. https://doi.org/10.1021/ acs.est.0c08129.

Wolfe, M.K., Archana, A., Catoe, D., Coffman, M.M., Dorevich, S., Graham, K.E., Kim, S., Grijalva, L.M., Roldan-Hernandez, L., Silverman, A.I., et al., 2021. Scaling of SARS-CoV-2 RNA in settled solids from multiple wastewater treatment plants to compare incidence rates of laboratory-confirmed COVID-19 in their sewersheds. Environ. Sci. Technol. Lett. 8, 398–404. https://doi.org/10.1021/acs.estlett.1c00184.

World Health Organization, 2020. Public Health Criteria to Adjust Public Health And Social Measures in the Context of COVID-19: Annex to Considerations in Adjusting Public Health And Social Measures in the Context of COVID-19. World Health Organization.

Xiao, A., Wu, F., Bushman, M., Zhang, J., Imakaev, M., Chai, P.R., Duvallet, C., Endo, N., Erickson, T.B., Armas, F., et al., 2022. Metrics to relate COVID-19 wastewater data to clin-ical testing dynamics. Water Res., 118070 https://doi.org/10.1016/j.watres.2022. 118070.

Yaniv, K., Shagan, M., Lewis, Y.E., Kramarsky-Winter, E., Weil, M., Indenbaum, V., Elul, M., Erster, O., Brown, A.S., Mendelson, E., et al., 2021. City-level SARS-CoV-2 sewage surveil-lance. Chemosphere 283, 131194. https://doi.org/10.1016/j.chemosphere.2021. 131194.

Yuan, S., Ye, Z.-W., Liang, R., Tang, K., Zhang, A.J., Lu, G., Ong, C.P., Poon, V.K.-M., Chan, C.C.-S., Mok, B.W.-Y., et al., 2022. The SARS-CoV-2 Omicron (B.1.1.529) variant exhibits altered pathogenicity, transmissibility, and fitness in the golden Syrian hamster model. Science 377, 428–433. https://doi.org/10.1126/science.abn8939.

Zdenkova, K., Bartackova, J., Cermakova, E., Demnerova, K., Dostalkova, A., Janda, V., Jarkovsky, J., Lopez Marin, M.A., Novakova, Z., Rumlova, M., et al., 2022. Monitoring COVID-19 spread in Prague local neighborhoods based on the presence of SARS-CoV-2 RNA in wastewater collected throughout the sewer network. Water Res. 216, 118343. https://doi.org/10.1016/j.watres.2022.118343.

Zhang, N., Gong, Y., Meng, F., Shi, Y., Wang, J., Mao, P., Chuai, X., Bi, Y., Yang, P., Wang, F., 2021. Comparative study on virus shedding patterns in nasopharyngeal and fecal speci-mens of COVID-19 patients. Sci. China Life Sci. 64, 486–488. https://doi.org/10.1007/ s11427-020-1783-9.

Zheng, X., Li, S., Deng, Y., Xu, X., Ding, J., Lau, F.T.K., In Yau, C., Poon, L.L.M., Tun, H.M., Zhang, T., 2022. Quantification of SARS-CoV-2 RNA in wastewater treatment plants mir-rors the pandemic trend in Hong Kong. Sci. Total Environ. 844, 157121. https://doi.org/ 10.1016/j.scitotenv.2022.157121.