

# **Title: Statistical Relationship Between Wastewater Data and Case Notifications for COVID-19 Surveillance in the United States, 2020-2023: A Bayesian Hierarchical Model.**

**Authors:** Masahiko Haraguchi [1], Fayette Klaassen [1], Ted Cohen [2], Joshua A. Salomon [3], Nicolas A. Menzies [1]

**Corresponding Author:** Masahiko Haraguchi ([mharaguchi@hsph.harvard.edu](mailto:mharaguchi@hsph.harvard.edu))

**Institutions:** [1] The Department of Global Health and Population, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston MA, USA; [2] Department of Epidemiology of Microbial Diseases and Public Health Modeling Unit, Yale School of Public Health, 60 College St, New Haven CT, USA; [3] Department of Health Policy, Stanford University School of Medicine, 1265 Welch Road, Stanford CA, USA.

**Email Address:** Fayette Klaassen [[fklaassen@hsph.harvard.edu](mailto:fklaassen@hsph.harvard.edu)], Ted Cohen [[theodore.cohen@yale.edu](mailto:theodore.cohen@yale.edu)], Joshua A. Salomon [[salomon1@stanford.edu](mailto:salomon1@stanford.edu)], Nicolas A. Menzies [[nmenzies@hsph.harvard.edu](mailto:nmenzies@hsph.harvard.edu)]

**Keywords:** wastewater epidemiology, COVID-19 surveillance; environmental surveillance; SARS-CoV-2; Bayesian modeling; the pandemic

## 1 Abstract

2 During the COVID-19 pandemic a number of jurisdictions in the United States began to regularly  
3 report levels of SARS-CoV-2 in wastewater for use as a proxy for SARS-CoV-2 incidence. Despite  
4 the promise of this approach for improving situational awareness, the degree to which viral levels  
5 in wastewater track with other outcome data has varied, and better evidence is needed to  
6 understand the situations in which wastewater surveillance tracks closely with traditional  
7 surveillance data. In this study, we quantified the relationship between wastewater data and  
8 traditional case-based surveillance data for multiple jurisdictions. To do so, we collated data on  
9 wastewater SARS-CoV-2 RNA levels and COVID-19 case reports from July 2020 to March 2023,  
10 and employed Bayesian hierarchical regression modeling to estimate the statistical relationship  
11 between wastewater data and reported cases, allowing for variation in this relationship across  
12 counties. We compared different model structural approaches and assessed how the strength of  
13 the estimated relationships varied between settings and over time. These analyses revealed a  
14 strong positive relationship between wastewater data and COVID-19 cases for the majority of  
15 locations, with a median correlation coefficient between observed and predicted cases of 0.904  
16 (interquartile range 0.823 – 0.943). Across locations, the COVID-19 case rate associated with a  
17 given level of wastewater SARS-CoV-2 RNA concentration declined over the study period.  
18 Counties with higher population size and of higher levels of urbanicity had stronger concordance  
19 between wastewater data and COVID-19 cases. Ideally, use of wastewater data for decision-  
20 making should be based on an understanding of their local historical performance.

21

22

## 1 Introduction

2 The practice of employing wastewater data to track pathogens has gained significant interest as  
3 an innovative method of infectious disease surveillance, with nearly 80% of the United States  
4 population connected to public wastewater systems (U.S. Government Accountability Office,  
5 2022; Yu et al., 2023). During the COVID-19 crisis, local health agencies started to gather and  
6 report data on COVID-19 concentrations in wastewater, using the trends in these data as an  
7 indirect indicator of SARS-CoV-2 transmission patterns.

8  
9 Tracking the presence and concentration of pathogens in wastewater, a passive method of  
10 environmental surveillance, has been used for several decades to track infectious diseases, such  
11 as polio (Matrajt et al., 2018), gastroenteritis (Kazama et al., 2017), hepatitis E (Alfonsi et al.,  
12 2018), and acute diarrhea (Prevost et al., 2015), among others. However, the COVID-19  
13 pandemic accelerated interest in this approach worldwide (Naughton et al., 2023), as it offers  
14 multiple benefits as a complement to traditional surveillance systems. First, it can serve as an  
15 early warning system for SARS-CoV-2 transmission, sensitive to asymptomatic and  
16 presymptomatic cases four to ten days prior to clinical testing signals (Wu et al., 2022). Second,  
17 it can monitor community-level transmission when implemented at downstream locations such  
18 as wastewater treatment plants (McClary-Gutierrez et al., 2021). The fact that it does not  
19 require individual testing circumvents the challenges created by variable supply of and demand  
20 for COVID-19 diagnostic testing (as has been observed over the pandemic), and the decline in  
21 reporting of test results (National Academies of Sciences & Medicine, 2023). Third, wastewater  
22 surveillance programs with specific methods can track virus variants in the early stage of their  
23 evolution, allowing for early identification of emerging variants (Jahn et al., 2022; Karthikeyan  
24 et al., 2022; McClary-Gutierrez et al., 2021). Lastly, it provides an early indicator of  
25 epidemiological changes (as compared to hospitalization and death data), so that mitigation  
26 and other response measures can be deployed more rapidly. With these advantages, tracking  
27 wastewater COVID-19 data is a potentially powerful tool for COVID-19 surveillance.

28  
29 Despite the potential advantages of SARS-CoV-2 wastewater surveillance, significant challenges  
30 remain. Existing studies have generally considered the relationship between reported cases and  
31 wastewater metrics at a limited number of locations, such as university dorms (Kotay et al.,  
32 2022), nursing homes (Davó et al., 2021), university campuses (Karthikeyan et al., 2021), and  
33 municipalities (e.g., Oklahoma City, Oklahoma (Kuhn et al., 2022) and Louisville, Kentucky  
34 (Klaassen et al., 2024)). While these studies provide valuable insights, they often do not fully  
35 account for heterogeneities across different geographical locations. Furthermore, the time  
36 periods covered by existing studies are limited. For example, Xiao et al. (2022) associated  
37 clinical case data with wastewater data within three Massachusetts counties from March 2020  
38 through May 2021. Similarly, Weidhaas et al. (2021) analyzed nine weeks of wastewater and  
39 COVID-19 case data related to ten wastewater treatment facilities in Utah. Notably, reported  
40 cases themselves are not a perfect indicator for true infection situations, given that they  
41 depend on various factors such as testing availability and access as well as the number of  
42 asymptomatic infections. As such, while quantifying the relationship between reported cases  
43 and wastewater metrics provides valuable insights, it should be noted as the next-best

1 alternative. Studies that compare the performance of COVID-19 wastewater surveillance data  
2 across sewersheds over extended time frames remain scarce (Dai et al., 2024). Such research is  
3 valuable for establishing the statistical basis for real-time trend analysis, and describing the  
4 conditions under which wastewater surveillance performs well.

5  
6 This study explored the quantitative relationship between SARS-CoV-2 wastewater surveillance  
7 data and reported COVID-19 diagnoses across multiple sewer-sheds over the initial years of the  
8 COVID-19 pandemic. Using weekly-aggregated wastewater and case report data for 107 U.S.  
9 counties over 2020-2023, we employed Bayesian hierarchical modeling to establish the  
10 statistical relationship between these two data sources, describe changes in these relationships  
11 over time and across locations, and identify how the strength of these relationships varied  
12 systematically by county characteristics.

## 14 Data and Methods

### 15 Data sources

16 For the study period July 1, 2020 to March 1, 2023, we collated county-level SARS-CoV-2  
17 wastewater surveillance data reported by Biobot Analytics, including 254 counties covering  
18 approximately 30% of the U.S. population. These data represent RNA copies per milliliter,  
19 normalized by the concentration of pepper mild mottle virus (PMMoV) to correct for variability  
20 in fecal content, which is influenced by environmental factors such as stormwater (Duvall et  
21 al., 2022). Surveillance data on county-level weekly COVID-19 reported case totals were  
22 extracted from the COVID-19 data repository in the Center for Systems Science and Engineering  
23 at Johns Hopkins University.

24 Due to varied implementation of wastewater surveillance operations, wastewater data were  
25 not available for all county-weeks. We restricted the analysis to counties with a minimum of 50  
26 weeks of available wastewater data during the study period, which resulted in 107 counties  
27 being included, covering a range of geographic areas within the United States, and with periods  
28 of data incompleteness for the majority of counties. We grouped counties into six ordinal  
29 urbanicity categories as defined by the National Center for Health Statistics (NCHS). Categories  
30 ranged from NCHS category 1 (large central metro) as the most urban to NCHS category 6 (non-  
31 core) as the most rural. Fig1 shows the geographic distribution of the 107 counties included in  
32 the analysis, coded by NCHS urbanicity category.

33

34 [Figure 1]

35

### 36 Hierarchical regression models

37 Using the wastewater and COVID-19 case data, we constructed hierarchical Bayesian regression  
38 models to capture key features of each data source, allowing for differences in the estimated  
39 relationship between these data over time and between modeled counties. Accounting for  
40 temporal variations is vital as the dynamics of the pandemic changed over time due to factors  
41 such as the emergence of new variants, public health interventions, vaccination rollouts, and

1 population immunity. Also, allowing for spatial variations accounts for local differences, such as  
2 vaccination coverage, and socioeconomic and demographic characteristics. As all of them  
3 influence the relationship between wastewater metrics and reported cases, a model that can  
4 account for the complex and evolving nature of the pandemic is critical. Namely, we modelled  
5 the weekly COVID-19 case reports for each county using a negative binomial likelihood,  
6 allowing for over-dispersion in these data:

$$7 \quad Y_{it} \sim \text{NegBin}(\mu_{it}, \varphi) \quad [1]$$

$$8 \quad \mu_{it} = \exp(\alpha_t + \beta_{0i} + \beta_{1i} * X_{it}) * n_i \quad [2]$$

9  
10 In Equation 1,  $Y_{it}$  represents reported COVID-19 cases for county  $i$  and week  $t$ . We used an  
11 alternative parameterization of the negative binomial in which  $\mu_{it}$  parameterizes the mean of  
12 the likelihood and  $\varphi$  parameterizes the extra-Poisson variation. The mean was specified as a  
13 function of  $\alpha_t$ , a time-varying coefficient given a random-walk prior;  $\beta_{0i}$ , a county-specific  
14 intercept; and  $\beta_{1i}$ , a county-specific slope term applied to  $X_{it}$ , the demeaned wastewater  
15 COVID-19 concentration value for each county and week (Equation 2). Both  $\beta_{0i}$  and  $\beta_{1i}$  were  
16 specified as random effects to pool information across counties.  $n_i$  represented the population  
17 of each country, as reported by the US Census Bureau. We fit this model to the COVID-19 case  
18 and wastewater data using the RStan package in R (Stan Development Team, 2022). Additional  
19 details on the specification of this regression model, including prior distributions, are provided  
20 in the Appendix.

## 21 [Analysis of fitted models](#)

22 We used several approaches to assess model fit. First, we visually compared the COVID-19 case  
23 time series for each county to the fitted values from the regression model. Second, we  
24 calculated a quantitative measure of model fit: the Median Absolute Deviation (MAD). We  
25 estimated this value for each county and used them to describe the overall level of model fit  
26 and how this varied across counties. We used the MAD to investigate whether the strength of  
27 estimated relationships differed systematically as a function of county characteristics  
28 (urbanicity, population size), examining these relationships visually and via univariable and  
29 multivariable regression models. Finally, we calculated the correlation coefficient between  
30 modelled values and raw case totals as a simple summary measure of model fit.

## 31 [Coefficient estimates](#)

32  
33 We used the fitted values of  $\alpha_t$  (the temporal trend in the regression model) to understand  
34 how the relationship between wastewater and COVID-19 case totals varied over the study  
35 period (July 2020 to March 2023). We used the fitted values of  $\beta_{0i}$  and  $\beta_{1i}$  to understand how  
36 the relationship between wastewater concentration and COVID-19 case totals varied within  
37 each county.  
38  
39

## 1 Results

### 2 Fitted relationship between COVID-19 cases and wastewater concentration

3 For the majority of modelled locations, we estimated a relationship between wastewater  
4 concentration and weekly COVID-19 cases, indicating that wastewater concentration serve as a  
5 useful predictor of case trends. Figure 2 shows the temporal trend in reported COVID-19 cases  
6 and fitted model estimates for each of six example counties, representing a range in terms of  
7 urbanicity and population size, which are often correlated but not identical. For each of these  
8 example counties the fitted model values (blue symbols) follow the empirical case data (black  
9 symbols) closely, with occasional deviations (e.g., 2021 estimates for Monterey County, CA).  
10 Figures for other counties are available in the Appendix. We also calculated correlation  
11 coefficients comparing observed and predicted values COVID-19 case counts. Across counties  
12 the median of these correlation coefficients was 0.904, with an inter-quartile range of 0.823 to  
13 0.943. Twenty-three counties (21.5% of the sample) had correlation coefficients below 0.8, and  
14 three had values below 0.6.

15

16 [Figure 2]

17

### 18 Systematic differences in model fit across counties

19 To further evaluate how the model performed in each county we calculated the MAD, for which  
20 smaller values indicate better model fit. For MAD, the median value was 0.259, with an inter-  
21 quartile range of 0.201 to 0.301. When we compared MAD to county population size (Figure 3-  
22 a), we found that model fit was better for counties with higher population numbers ( $p < 0.001$ ).

23

24 [Figure 3]

25

26 Furthermore, we found that MAD was associated with urbanicity (Figure 3-b), with more urban  
27 counties (Categories 1, 2, and 3) having lower MAD ( $p < 0.001$ ) and therefore better model fit.  
28 More rural counties had poorer model fits (higher values of MAD), with the exception of  
29 Chittenden VT.

30

31 When we fit a multivariable regression model including both logged population and urbanicity  
32 category as predictors we found both coefficients to have the same sign as in the univariate  
33 analyses but were no longer significant (population size:  $p = 0.09$ , urbanicity:  $p = 0.13$ ).

34

### 35 Time-trends

36 Figure 4 shows changes in the estimated relationship between wastewater concentration and  
37 COVID-19 case reports over the study period, quantified as the level of logged COVID-19 case  
38 totals consistent with a given wastewater concentration, shown in blue. Notably, the  
39 fluctuations in this relationship are closely associated with the significant US waves at the end  
40 of 2020 to the beginning of 2021 (Alpha wave), the summer of 2022 (Delta wave), and the  
41 beginning of 2022 (Omicron wave). During these periods, the ratio of COVID-19 cases to  
42 wastewater concentration is relatively high, as compared to the months before and afterwards.  
43 After early 2022 the estimated trend shows a progressive decline.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42

[Figure 4]

#### Differences across counties

Figure 5 shows fitted estimates for how changes in wastewater concentration values are associated with changes in COVID-19 cases for each county. As expected, the slopes shown in Figure 5 are positive (indicating that an increase in wastewater concentration was associated with an increase in reported cases), and relatively consistent across counties. In all cases the slope of these lines was estimated to be less than 1.0 (median: 0.551, IQR: 0.447 to 0.632), indicating that that fitted relationship between wastewater concentration and COVID-19 case totals is less than proportional (e.g., a 50% increase in wastewater concentration is associated with a <50% increase in case totals).

[Figure 5]

#### Discussion

Wastewater data have been used extensively during the SARS-CoV-2 pandemic to monitor disease trends and provide early evidence of rising community transmission. However, there is limited information on the statistical relationship between wastewater metrics and reported COVID-19 cases, and how this relationship varies over time and across jurisdictions (Dai et al., 2024). This knowledge is valuable for making decisions about how best to use wastewater data, and to understand the settings in which these data provide accurate information about COVID-19 case trends. In this study, we modeled the relationship between wastewater metrics and clinical cases at the county level in the United States from July 2020 to March 2023.

The results of our analysis show that models fit to wastewater data are better able to predict case counts in urban counties (based on NCHS categorization) as compared to more rural counties. This may be due to rural areas having lower levels of connection to centralized sewage systems, the source of wastewater surveillance data (National Academies of Sciences & Medicine, 2023; M. Varkila et al., 2023). We also noted a reduction in model performance among counties with smaller population sizes. This is consistent with other studies that have reported wastewater surveillance to have limited sensitivity as an early warning indicator in smaller geospatial scales (Gamage et al., 2024; Klaassen et al., 2024). We also estimated differences in the quality of model fit that were not explained by urbanicity and population size—these may relate to local differences in the coverage of wastewater surveillance, the processing of wastewater samples, or the quality of COVID-19 case reporting.

In addition to inter-county differences, the results also revealed fluctuations in the relationship between wastewater concentration and COVID-19 case totals over the course of the pandemic. Several factors could account for these temporal trends. First, viral shedding patterns among individuals developing COVID-19 in recent years likely differ from those developing disease in the early stage of the pandemic, as immunity in the population through previous infections and



1 vaccination increased significantly throughout the pandemic (Puhach et al., 2023). Second, the  
2 transition between different dominant variants may have also influenced the dynamics of  
3 discharged RNA copies in human waste. For example, the viral and antibody dynamics are  
4 distinct between omicron and delta variants (Yang et al., 2023). Also, as mutations may affect  
5 the quantification of SARS-CoV-2 concentration in wastewater, such viral changes may need to  
6 be accounted for in estimating the relationship between wastewater levels and case  
7 notifications (Endo et al., 2023). Lastly, case reporting systems have changed over time,  
8 influencing the ratio of reported and unreported cases (Alvarez et al., 2023; Silk et al., 2023). In  
9 particular, the declining trend over the final year of the time series (i.e. a declining number of  
10 reported COVID-19 cases for a given wastewater level) likely relates to changes in COVID-19  
11 testing and reporting practices, with a progressively smaller fraction of COVID-19 cases  
12 diagnosed and reported to public health authorities.

13  
14 Designing effective wastewater surveillance systems require trade-offs among cost-  
15 effectiveness, speed, and local feasibility (Levy et al., 2023). Most current sequencing is  
16 implemented with hundreds to thousands of samples in parallel with expensive machinery and  
17 intensive investment in human resources (Levy et al., 2023). This implies that counties with  
18 fewer resources may find it difficult to finance and support the required laboratory  
19 infrastructure and human resources.

20  
21 While the quality of model fit was generally good, our analyses revealed substantial variation in  
22 the utility of wastewater surveillance across counties. It is also important to note that reported  
23 case counts, which we used as a proxy for infection trends, are an imperfect measure of true  
24 incidence (M. R. Varkila et al., 2023). Ideally, a population-based survey, such as the United  
25 Kingdom's Office for National Statistics' COVID-19 Infection Survey, would provide a more  
26 accurate information for assessing the predictive performance of wastewater surveillance.  
27 However, without such data in the US, we rely on case reports as the best available data.  
28 Further investigation, validation, and standardized data collection frameworks are required to  
29 better understand the relationship between wastewater and epidemiological data. The  
30 incorporation of next-generation sequencing and automation of wastewater data collection  
31 processes could enhance the effectiveness of wastewater surveillance (Iwamoto et al., 2023;  
32 Singer et al., 2023).

## 33 Conclusions

34 The SARS-CoV-2 pandemic made it clear that traditional event-based surveillance systems have  
35 critical deficiencies for providing prompt and valid information about the local epidemiological  
36 situation. Wastewater surveillance may provide health agencies with another early detection  
37 and effective surveillance tool, unaffected by several of the deficiencies of traditional  
38 surveillance data. As of March 2024, more than 1300 locations in the US and over 72 countries  
39 globally conducted wastewater surveillance (CDC, 2024; University of California Merced, 2024).  
40 When implemented effectively, these can data provide a comprehensive picture of SARS-CoV-2  
41 transmission, capturing asymptomatic and non-tested infections. Our study demonstrates that  
42 analyzing wastewater metrics across multiple jurisdictions can establish the relationship  
43 between wastewater and potential cases, and how these differ across locations and over time.

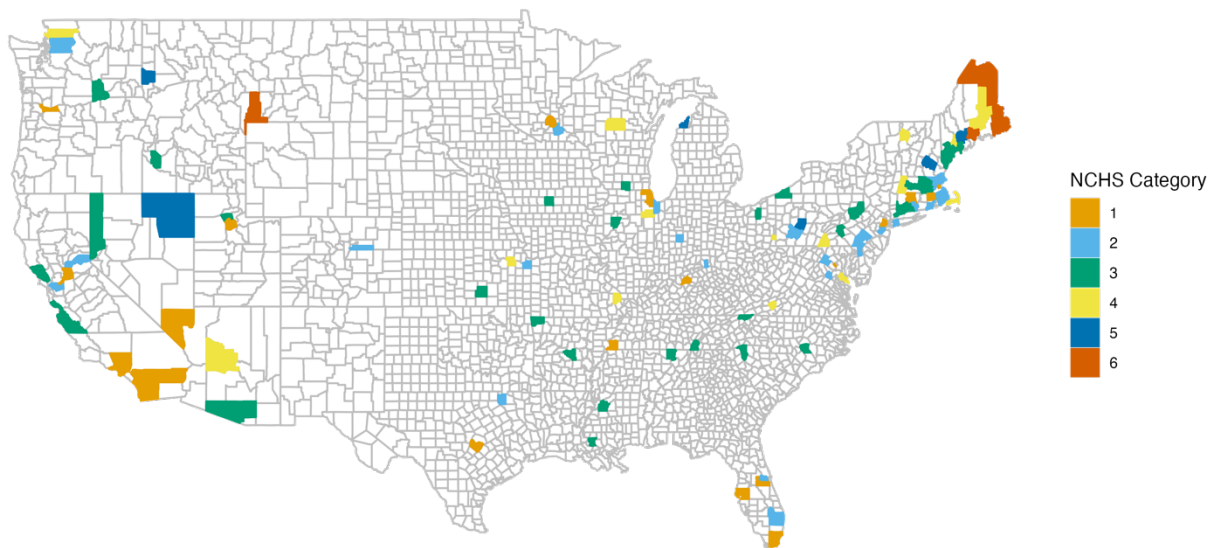


1 However, the missing data in wastewater and uncertainty in case data require future efforts to  
2 make the relationship between them more established. Efforts to collect wastewater data in a  
3 more standardized manner should be enhanced further to fully realize their potential.

4

1 Figures

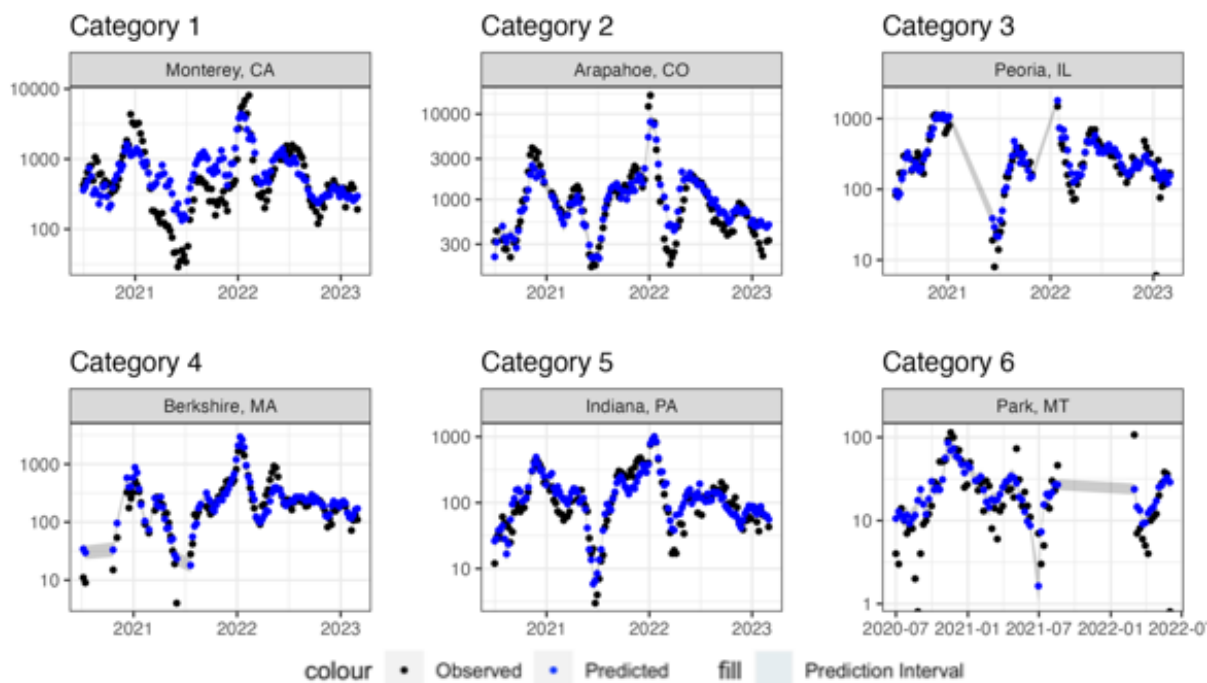
2  
3



4  
5 **Figure 1: This shows geographic location and NCHS urbanicity category for the 107 counties included in the**  
6 **analysis.**

7

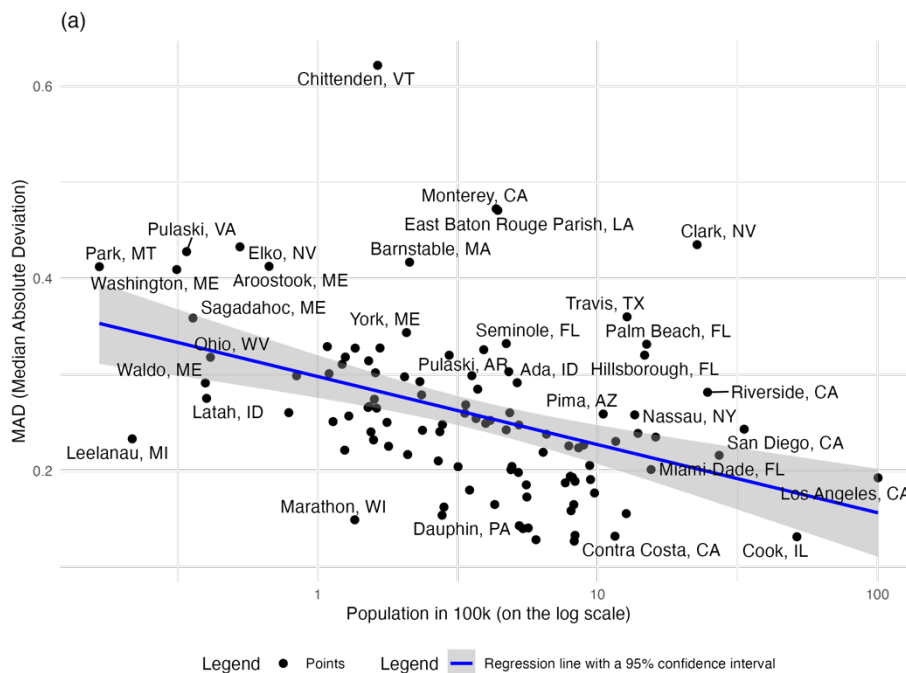
1  
2



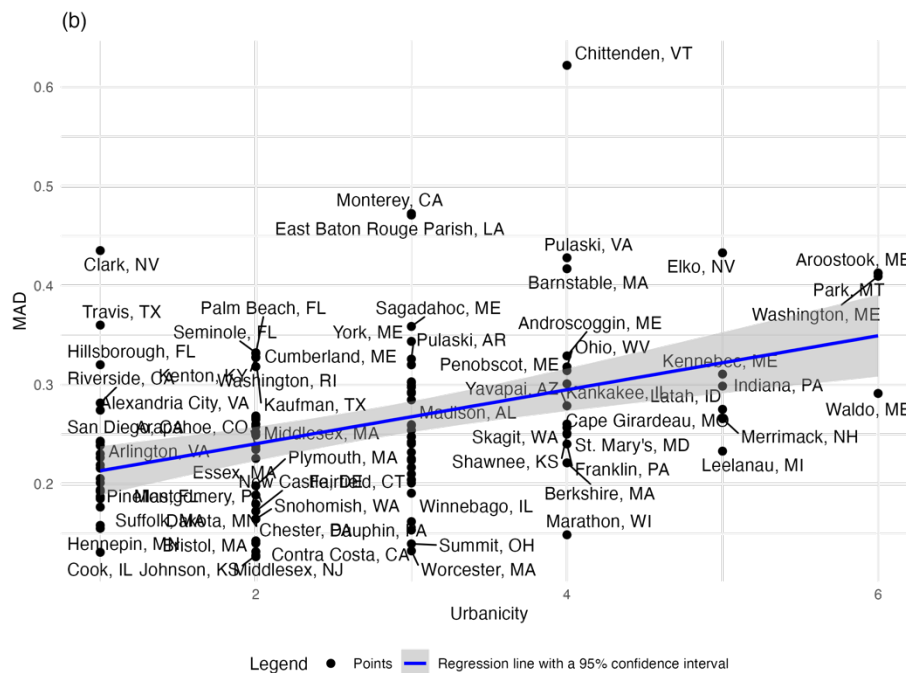
3  
4 **Figure 2: Comparison of observed and predicted COVID-19 case counts for a select group of counties within**  
5 **each urban-rural category as defined by the National Center for Health Statistics (NCHS).**

6  
7

1  
2



3



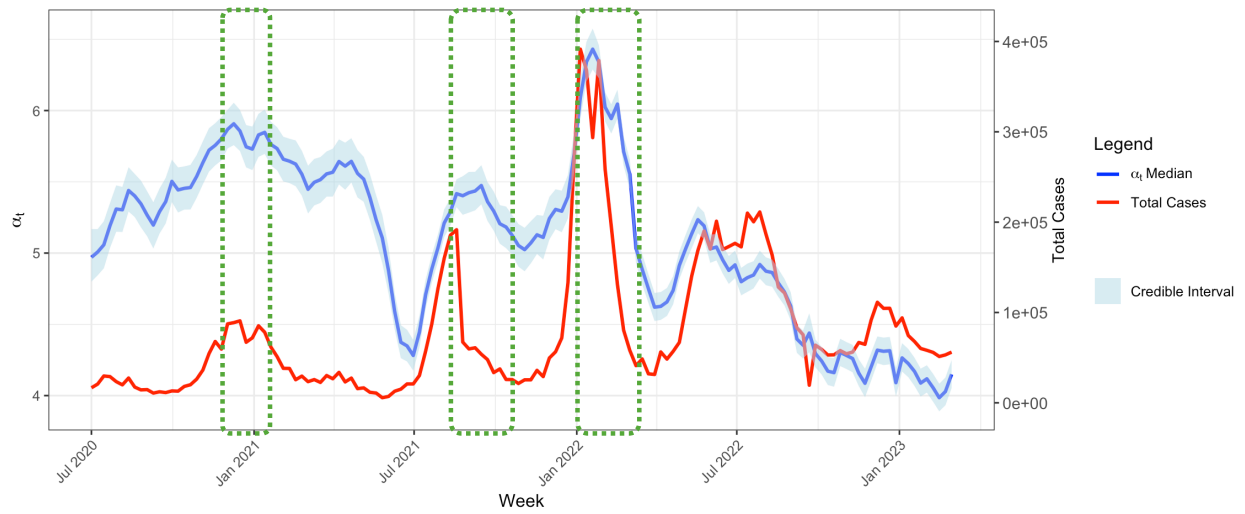
4

5

6

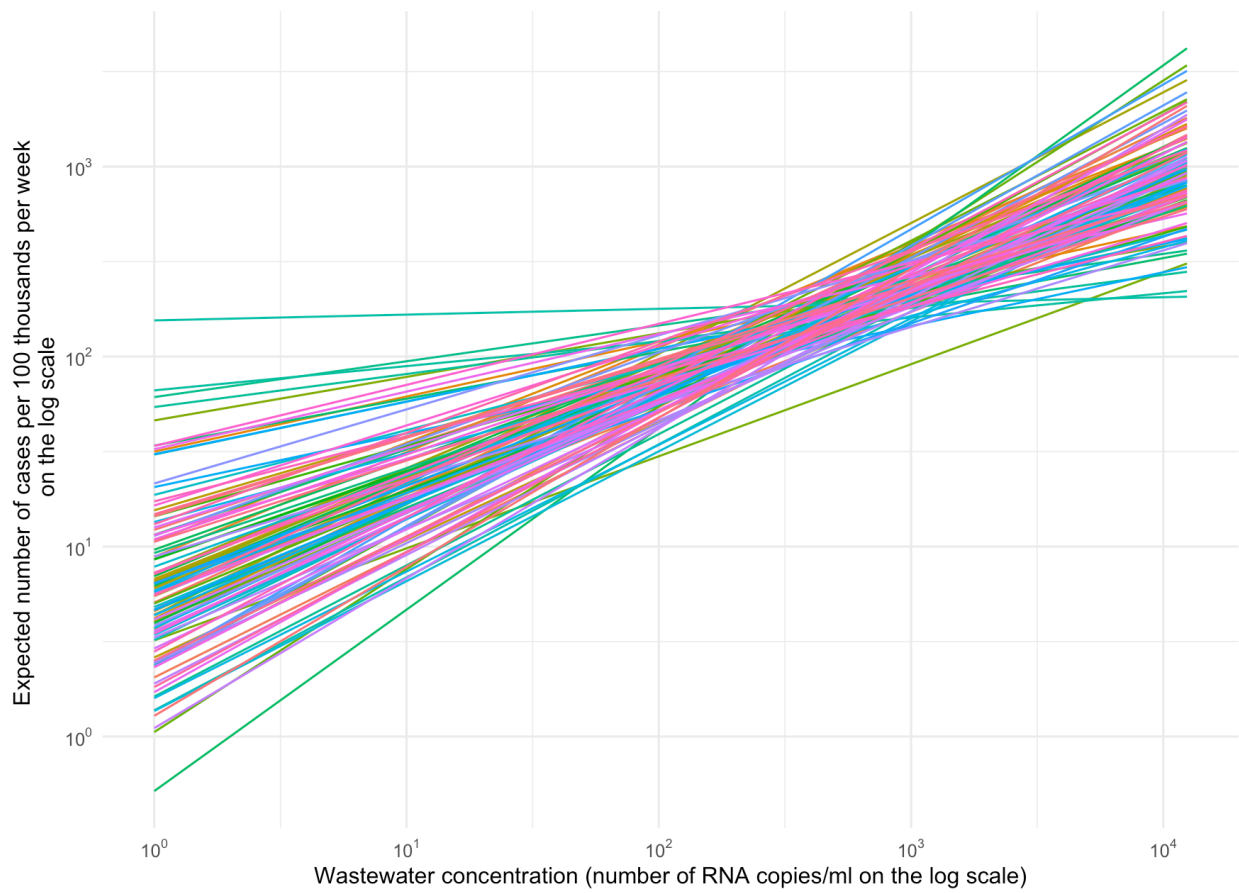
**Figure 3: Panel A shows how the quality of model fit (MAD) varies with county population size. Panel B shows how the quality of model fit (MAD) varies with country urbanicity category.**

7



1  
2  
3  
4

**Figure 4: Time series of the time-varying coefficient ( $\alpha_t$ ) alongside the aggregated case counts across counties from July 2020 to February 2023.**



1  
 2 **Figure 5: Estimated relationship between wastewater concentration and mean weekly cases per 100**  
 3 **thousand people for each county.**

4  
 5  
 6  
 7



## References

- Alfonsi, V., Romanò, L., Ciccaglione, A. R., La Rosa, G., Bruni, R., Zanetti, A., Della Libera, S., Iaconelli, M., Bagnarelli, P., & Capobianchi, M. R. (2018). Hepatitis E in Italy: 5 years of national epidemiological, virological and environmental surveillance, 2012 to 2016. *Eurosurveillance*, 23(41), 1700517.
- Alvarez, E., Bielska, I. A., Hopkins, S., Belal, A. A., Goldstein, D. M., Slick, J., Pavalagantharajah, S., Wynfield, A., Dakey, S., & Gedeon, M.-C. (2023). Limitations of COVID-19 testing and case data for evidence-informed health policy and practice. *Health Research Policy and Systems*, 21(1), 11.
- CDC. (2024). *COVID-19: Wastewater Surveillance*. Retrieved March 21 from <https://covid.cdc.gov/covid-data-tracker/#wastewater-surveillance>
- Dai, X., Acosta, N., Lu, X., Hubert, C. R., Lee, J., Frankowski, K., Bautista, M. A., Waddell, B. J., Du, K., & McCalder, J. (2024). A Bayesian framework for modeling COVID - 19 case numbers through longitudinal monitoring of SARS - CoV - 2 RNA in wastewater. *Statistics in Medicine*.
- Davó, L., Seguí, R., Botija, P., Beltrán, M. J., Albert, E., Torres, I., López-Fernández, P. Á., Ortí, R., Maestre, J. F., & Sánchez, G. (2021). Early detection of SARS-CoV-2 infection cases or outbreaks at nursing homes by targeted wastewater tracking. *Clinical Microbiology and Infection*, 27(7), 1061-1063.
- Duvallet, C., Wu, F., McElroy, K. A., Imakaev, M., Endo, N., Xiao, A., Zhang, J., Floyd-O'Sullivan, R., Powell, M. M., & Mendola, S. (2022). Nationwide trends in COVID-19 cases and SARS-CoV-2 RNA wastewater concentrations in the United States. *ACS Es&t Water*, 2(11), 1899-1909.
- Endo, N., Nihei, Y., Fujita, T., Yasojima, M., Daigo, F., Takemori, H., Nakamura, M., Matsuda, R., Sovannrlaksmey, S., & Ihara, M. (2023). Mutations affect the quantification of SARS-CoV-2 in wastewater: observations, mechanisms, and recommendations.
- Gamage, S. D., Jinadatha, C., Rizzo Jr, V., Chatterjee, P., Choi, H., Mayo, L., Brackens, E., Hwang, M., Xu, J., & Bennett, M. (2024). Nursing home wastewater surveillance for early warning of SARS-CoV-2-positive occupants—Insights from a pilot project at eight facilities. *American Journal of Infection Control*.
- Iwamoto, R., Yamaguchi, K., Katayama, K., Ando, H., Setsukinai, K.-i., Kobayashi, H., Okabe, S., Imoto, S., & Kitajima, M. (2023). Identification of SARS-CoV-2 variants in wastewater using targeted amplicon sequencing during a low COVID-19 prevalence period in Japan. *Science of the Total Environment*, 887, 163706.
- Jahn, K., Dreifuss, D., Topolsky, I., Kull, A., Ganesanandamoorthy, P., Fernandez-Cassi, X., Bänziger, C., Devaux, A. J., Stachler, E., & Caduff, L. (2022). Early detection and surveillance of SARS-CoV-2 genomic variants in wastewater using COJAC. *Nature Microbiology*, 7(8), 1151-1160.
- Karthikeyan, S., Levy, J. I., De Hoff, P., Humphrey, G., Birmingham, A., Jepsen, K., Farmer, S., Tubb, H. M., Valles, T., & Tribelhorn, C. E. (2022). Wastewater sequencing reveals early cryptic SARS-CoV-2 variant transmission. *Nature*, 609(7925), 101-108.

- 1 Karthikeyan, S., Nguyen, A., McDonald, D., Zong, Y., Ronquillo, N., Ren, J., Zou, J., Farmer, S.,  
2 Humphrey, G., & Henderson, D. (2021). Rapid, large-scale wastewater surveillance and  
3 automated reporting system enable early detection of nearly 85% of COVID-19 cases on  
4 a university campus. *Msystems*, 6(4), e00793-00721.
- 5 Kazama, S., Miura, T., Masago, Y., Kōta, Y., Tohma, K., Manaka, T., Liu, X., Nakayama, D.,  
6 Tanno, T., & Saito, M. (2017). Environmental surveillance of norovirus genogroups I and  
7 II for sensitive detection of epidemic variants. *Applied and environmental microbiology*,  
8 83(9), e03406-03416.
- 9 Klaassen, F., Holm, R. H., Smith, T., Cohen, T., Bhatnagar, A., & Menzies, N. A. (2024). Predictive  
10 power of wastewater for nowcasting infectious disease transmission: a retrospective  
11 case study of five sewershed areas in Louisville, Kentucky. *Environmental Research*, 240,  
12 117395.
- 13 Kotay, S. M., Tanabe, K. O., Colosi, L. M., Poulter, M. D., Barry, K. E., Holstege, C. P., Mathers, A.  
14 J., & Porter, M. D. (2022). Building-level wastewater surveillance for SARS-CoV-2 in  
15 occupied university dormitories as an outbreak forecasting tool: one year case study.  
16 *ACS Es&t Water*, 2(11), 2094-2104.
- 17 Kuhn, K. G., Jarshaw, J., Jeffries, E., Adesigbin, K., Maytubby, P., Dundas, N., Miller, A. C.,  
18 Rhodes, E., Stevenson, B., & Vogel, J. (2022). Predicting COVID-19 cases in diverse  
19 population groups using SARS-CoV-2 wastewater monitoring across Oklahoma City.  
20 *Science of the Total Environment*, 812, 151431.
- 21 Levy, J. I., Andersen, K. G., Knight, R., & Karthikeyan, S. (2023). Wastewater surveillance for  
22 public health. *Science*, 379(6627), 26-27.
- 23 Matrajt, G., Naughton, B., Bandyopadhyay, A. S., & Meschke, J. S. (2018). A review of the most  
24 commonly used methods for sample collection in environmental surveillance of  
25 poliovirus. *Clinical Infectious Diseases*, 67(suppl\_1), S90-S97.
- 26 McClary-Gutierrez, J. S., Mattioli, M. C., Marcenac, P., Silverman, A. I., Boehm, A. B., Bibby, K.,  
27 Balliet, M., Gerrity, D., Griffith, J. F., & Holden, P. A. (2021). SARS-CoV-2 wastewater  
28 surveillance for public health action. *Emerging infectious diseases*, 27(9).
- 29 National Academies of Sciences, E., & Medicine. (2023). Wastewater-based disease surveillance  
30 for public health action.
- 31 Naughton, C. C., Roman, F. A., Alvarado, A. G. F., Tariqi, A. Q., Deeming, M. A., Kadonsky, K. F.,  
32 Bibby, K., Bivins, A., Medema, G., & Ahmed, W. (2023). Show us the data: global COVID-  
33 19 wastewater monitoring efforts, equity, and gaps. *FEMS microbes*, 4.
- 34 Prevost, B., Lucas, F. S., Goncalves, A., Richard, F., Moulin, L., & Wurtzer, S. (2015). Large scale  
35 survey of enteric viruses in river and waste water underlines the health status of the  
36 local population. *Environment International*, 79, 42-50.
- 37 Puhach, O., Meyer, B., & Eckerle, I. (2023). SARS-CoV-2 viral load and shedding kinetics. *Nature  
38 Reviews Microbiology*, 21(3), 147-161.
- 39 Silk, B. J., Scobie, H. M., Duck, W. M., Palmer, T., Ahmad, F. B., Binder, A. M., Cisewski, J. A.,  
40 Kroop, S., Soetebier, K., & Park, M. (2023). COVID-19 Surveillance After Expiration of the  
41 Public Health Emergency Declaration—United States, May 11, 2023. *Morbidity and  
42 Mortality Weekly Report*, 72(19), 523.
- 43 Singer, A. C., Thompson, J. R., Filho, C. R. M., Street, R., Li, X., Castiglioni, S., & Thomas, K. V.  
44 (2023). A world of wastewater-based epidemiology. *Nature Water*, 1(5), 408-415.

1 Stan Development Team. (2022). *RStan: the R interface to Stan. R package version 2.21.7*. In  
2 <https://mc-stan.org/>

3 U.S. Government Accountability Office. (2022). *Science & Tech Spotlight: Wastewater*  
4 *Surveillance*. (GAO-22-105841). U.S. Government Accountability Office, Retrieved from  
5 <https://www.gao.gov/products/gao-22-105841>

6 University of California Merced. (2024). *COVIDPoops19*.  
7 <https://www.arcgis.com/apps/dashboards/c778145ea5bb4daeb58d31afee389082>

8 Varkila, M., Montez-Rath, M., Salomon, J., Yu, X., Block, G., Owens, D. K., Chertow, G. M.,  
9 Parsonnet, J., & Anand, S. (2023). Use of wastewater metrics to track COVID-19 in the  
10 US: a national time-series analysis over the first three quarters of 2022. *medRxiv*.

11 Varkila, M. R., Montez-Rath, M. E., Salomon, J. A., Yu, X., Block, G. A., Owens, D. K., Chertow, G.  
12 M., Parsonnet, J., & Anand, S. (2023). Use of wastewater metrics to track COVID-19 in  
13 the US. *JAMA Network Open*, 6(7), e2325591-e2325591.

14 Weidhaas, J., Aanderud, Z. T., Roper, D. K., VanDerslice, J., Gaddis, E. B., Ostermiller, J.,  
15 Hoffman, K., Jamal, R., Heck, P., & Zhang, Y. (2021). Correlation of SARS-CoV-2 RNA in  
16 wastewater with COVID-19 disease burden in sewersheds. *Science of the Total*  
17 *Environment*, 775, 145790.

18 Wu, F., Xiao, A., Zhang, J., Moniz, K., Endo, N., Armas, F., Bonneau, R., Brown, M. A., Bushman,  
19 M., & Chai, P. R. (2022). SARS-CoV-2 RNA concentrations in wastewater foreshadow  
20 dynamics and clinical presentation of new COVID-19 cases. *Science of the Total*  
21 *Environment*, 805, 150121.

22 Xiao, A., Wu, F., Bushman, M., Zhang, J., Imakaev, M., Chai, P. R., Duvallet, C., Endo, N.,  
23 Erickson, T. B., & Armas, F. (2022). Metrics to relate COVID-19 wastewater data to  
24 clinical testing dynamics. *Water research*, 212, 118070.

25 Yang, Y., Guo, L., Yuan, J., Xu, Z., Gu, Y., Zhang, J., Guan, Y., Liang, J., Lu, H., & Liu, Y. (2023). Viral  
26 and antibody dynamics of acute infection with SARS-CoV-2 omicron variant (B. 1.1. 529):  
27 a prospective cohort study from Shenzhen, China. *The Lancet Microbe*, 4(8), e632-e641.

28 Yu, Q., Olesen, S. W., Duvallet, C., & Grad, Y. H. (2023). Assessment of sewer connectivity in the  
29 United States and its implications for equity in wastewater-based epidemiology.  
30 *medRxiv*, 2023.2005. 2024.23290486.

31

**Conflict of Interest Statement:** We declare no competing interests

**Acknowledgement:** This project is supported by cooperative agreement NU38OT000297 from the Centers for Disease Control and Prevention (CDC) and the Council of State and Territorial Epidemiologists (CSTE), and SHEPherD contract 200-2016-91779 from the CDC. This work does not necessarily represent the views of the CDC or CSTE.