Development of a COVID-19 warning system for neighborhood-scale wastewater-based epidemiology in low incidence situations

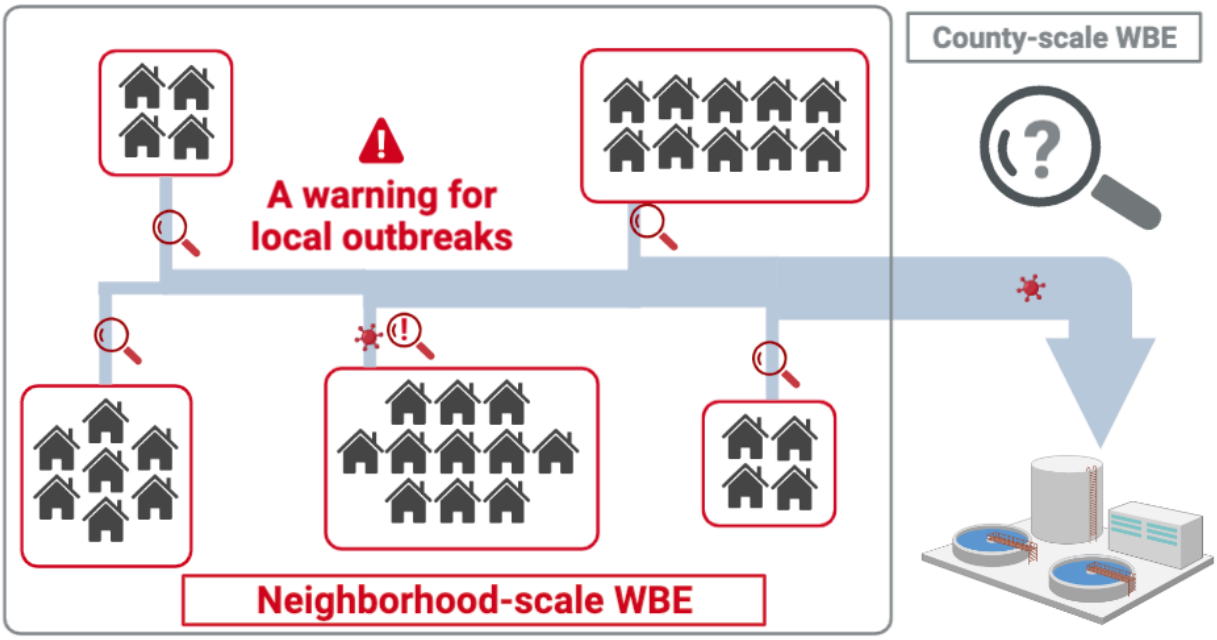
Authors

Chamteut Oh[1,*], Aijia Zhou[1], Kate O'Brien[2], Yusuf Jamal[3], Joanna L. Shisler[4], Antarpreet Jutla[3], Arthur Schmidt[1], Laura Keefer[5], William M. Brown[6], Thanh H. Nguyen[1,7]

*Corresponding authors: Chamteut Oh (co14@illinois.edu)

1) Department of Civil and Environmental Engineering, University of Illinois Urbana-Champaign, United States
2) School of Integrative Biology, University of Illinois Urbana-Champaign, United States
3) Department of Environmental Engineering Sciences, University of Florida, Gainesville, United States
4) Department of Microbiology, University of Illinois Urbana-Champaign, United States
5) Illinois State Water Survey, Prairie Research Institute, University of Illinois Urbana-Champaign, United States
6) Department of Pathobiology, College of Veterinary Medicine, University of Illinois Urbana-Champaign, United States
7) Institute of Genomic Biology, University of Illinois Urbana-Champaign, United States

Graphical abstract



Abstract

Wastewater-based epidemiology (WBE) is an emerging approach for community-wide COVID-19 surveillance. WBE was characterized mostly at large sewersheds such as wastewater treatment plants serving a large population. Although informed public health measures can be implemented in a small population, WBE for neighborhood-scale sewersheds is less studied and not fully understood. This study applied WBE to seven neighborhood-scale sewersheds (average population 1,471) from January to November 2021. Community testing data showed an average of 0.004% incidence rate in these sewersheds (97% of monitoring periods reported two or fewer daily infections). In 92% of sewage samples, SARS-CoV-2 N gene fragments were below the limit of quantification. COVID-19 cases poorly correlated with SARS-CoV-2 N gene concentrations except for one location where frequent COVID-19 testing was required. Thus, we developed a COVID-19 warning system for high COVID-19 risk when the SARS-CoV-2 N gene concentration normalized to pepper mild mottle virus (PMMOV) in wastewater is higher than 10^{-2}. This COVID-19 warning system now identified neighborhood-scale outbreaks (COVID-19 incidence rate higher than 0.5%) with 73% sensitivity and 71% specificity. Importantly, it can discern neighborhood-level outbreaks that would not otherwise be identified by county-scale WBE. We also demonstrated that SARS-CoV-2 variant-specific RT-qPCR assays could be combined with the COVID-19 warning system to identify the introduction of SARS-CoV-2 variants into sewersheds with a low COVID-19 incidence. This new COVID-19 warning system will contribute to effective community-wide disease surveillance when COVID-19 incidence may be maintained at a low level.

Keywords: Wastewater-based epidemiology; neighborhood-scale sewersheds; SARS-CoV-2 variant-specific RT-qPCR assays; low COVID-19 incidence; COVID-19 endemic.

Introduction

Wastewater-based epidemiology (WBE) has been used as a cost effective epidemiological tool for community-wide COVID-19 surveillance worldwide (Hart and Halden, 2020)(Chen et al., 2020; Sherchan et al., 2021; Tiwari et al., 2021). For example, WBE can detect SARS-CoV-2 at an 0.01% incidence (1infection in 10,000 people), when the influent of wastewater treatment plants serving a community of more than 100,000 residents was analyzed (Hata et al., 2021). The viral genome concentrations in sewersheds may be a better proxy for actual disease incidence than the reported COVID-19 infection cases because reported cases are significantly affected by temporal clinical testing capacity (Xiao et al., 2022). Furthermore, WBE affords the use of longitudinal wastewater surveillance to predict disease outbreaks in advance of clinical epidemiology because SARS-CoV-2 genes are shed before symptom onset (Wu et al., 2022b).

Large-scale sewersheds (such as wastewater treatment plants serving hundreds of thousands to millions of people) routinely have been surveilled for COVID-19 since early 2020 (Hata et al., 2021; Xiao et al., 2022). However, the large-sized sewersheds may not provide detailed information for region-specific public health measures. For example, social determinants of COVID-19 risk, such as demographics and socioeconomic status, which are regional characteristics (Abrams and Szefler, 2020; Jamal et al., 2022; Upshaw et al., 2021), may not be reflected in the influent wastewater from large sewersheds. The heterogeneity of these social determinants across the sewersheds is a possible explanation for discrepancies between state-wide and county-specific COVID-19 data (Messner and Payson, 2020). WBE also has been applied to small-scale sewersheds, such as buildings on university campuses (Bivins and Bibby, 2021; Gibas et al., 2021; Karthikeyan et al., 2021). At this smaller scale, viruses are detected in wastewater

when there are COVID-19 infection cases in the buildings. However, building-scale monitoring is not practical for residential areas that have high numbers of single family homes.

Neighborhood scale monitoring for COVID-19 has been conducted (Barrios et al., 2021; Spurbeck et al., 2021). While this neighboordhood scale monitoring could allow a more targeted public health intervention for a specific community, SARS-CoV-2 N gene concentrations presented lower correlation coefficients with COVID-19 infection cases as the catchment populations decreased (Bitter et al., 2022; Rusiñol et al., 2021; Sangsanont et al., 2022). Additionally, the probability of SARS-CoV-2 detection in wastewater samples at a given COVID-19 incidence decreases with a decrease in the population size (Fitzgerald et al., 2021; Wu et al., 2021). Thus, the neighborhood scale monitoring strategies are ripe for improvement.

This study aims to develop a methodology for the neighborhood-scale WBE for community-wide COVID-19 surveillance. We collected sewage samples from seven neighborhood-scale sewersheds (average population 1,471 people) across Champaign County, IL, USA for eleven months. First, we analyzed the SARS-CoV-2 N gene, including the genomes for the Alpha and Delta variants. Then, we plotted these data onto the human COVID-19 testing data obtained from the neighborhood-scale sewersheds, allowing us to develop a COVID-19 warning system. This COVID-19 warning system provides valuable information for county-wide disease surveillance. For example, our data uncovered local outbreaks that county-scale WBE would not detect, and also the introduction of SARS-CoV-2 variants of concerns to the sewersheds even at low COVID-19 incidence. Our findings suggest that the COVID-19 warning system will contribute to COVID-19 surveillance, especially when COVID-19 incidence is maintained at a low level, such as at a COVID-19 endemic phase.

Materials and Methods

Collection of raw sewage composite samples

We selected seven sampling points of manholes across Champaign County, IL, USA. These seven sampling points receive sewage discharged from sewersheds whose areas vary from 0.09 to 1.70 km^2. Population in each sewershed ranged from 853 to 2402 in 2020 (**Fig. 1**). We chose an additional sampling point: collecting wastewater generated from a food processing plant. The amount of human fecal matter released in this wastewater is much smaller than the processing water from slaughtering and meatpacking. For this reason, we used this wastewater for process control, evaluating cross-contamination among the sewage samples during the sample processing.

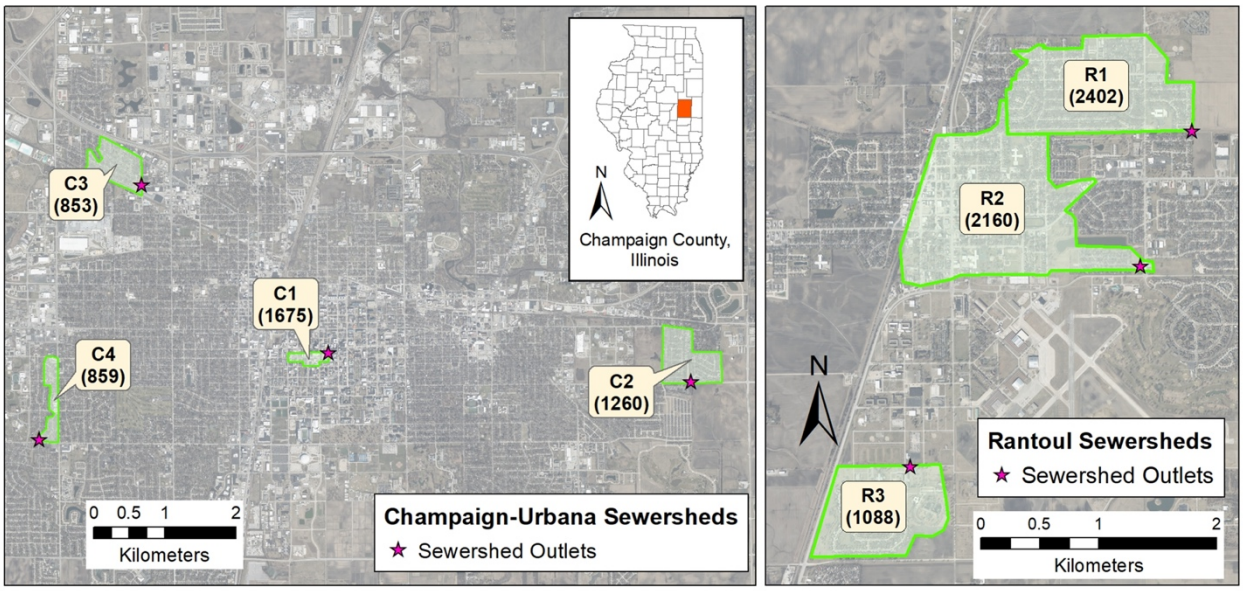


Fig. 1. Map of sewersheds in (A) Champaign and Urbana cities and (B) Rantoul city in Illinois. Values in parenthesis are population estimates using block-level US Census 2020 data.

We installed autosamplers (Teledyne ISCO, USA) programmed to collect a 1 to 2 L composite sample comprised of samples pumped at 4-hour intervals for four days. First, the composite samples were transferred to sterile sampling bags (14-955-001, Fisher Scientific, USA) and 20 mL

113 of 2.5 M $MgCl_2$ was added to the samples (i.e., final $MgCl_2$ concentrations from 25 to 50 mM) to
114 coagulate solids including virus particles (Ahmed et al., 2020b; Oh et al., 2022a). Samples were
115 transported on ice to a laboratory at the University of Illinois Urbana-Champaign within three
116 hours. All contact material was sanitized with 10% bleach and 70% ethanol to minimize cross-
117 contamination whenever handling different samples. We monitored the sewersheds from January
118 2021 to November 2021, during which we collected 254 sewage samples in total. The minimum
119 recommended meta-information on sewage samples is summarized in **Table S1**.
120

121 **Sample processing and viral nucleic acid extraction**

122 Upon arrival at the laboratory, supernatants from each composite sample were discarded. The
123 remaining 35 mL of sewage, in which solid particles were concentrated, were transferred to a 50
124 mL tube (12-565-271, Fisher Scientific, USA). Next, we added 350 μL of 10^5 gene copies (gc)/μL
125 of bovine coronavirus (BCoV; 16445-1, Merck Animal Health, USA) to the 35 mL of concentrated
126 sewage. After a 10 minutes incubation at room temperature, samples were centrifuged at 10,000 g
127 for 30 minutes (Sorvall™ RC 6 Plus, Thermo Scientific, USA). Supernatants were discarded and
128 a portion of the concentrated sludge (100 μL) was transferred to a sterile 1.5 mL tube (1415-2600,
129 USA Scientific, USA). Total RNA and DNA, including SARS-CoV-2 RNA, were extracted from
130 the sludge with QIAamp Viral RNA mini kit (Qiagen, German) following the manufacturer's
131 procedure with a minor modification. Specifically, the final volume of extracts was 100 μL.
132 Sewage collection and processing were conducted on the same day, and the RNA samples were
133 stored at -80 °C until downstream analysis.
134

135 **Analysis of viral genomes**

136 Total RNA was diluted serially 10-fold, and six reverse transcriptase quantitative polymerase chain
137 reaction (RT-qPCR) assays (**Table S2**) were used to analyze sampels for the presence of SARS-
138 CoV-2 RNA. We quantified the N gene (i.e., N1 assay with 2019-nCoV RUO kit, Integrated DNA
139 Technologies, USA) to identify a total SARS-CoV-2 RNA concentration. We also analyzed
140 samples for the presence of the S:A570D and S:P681R mutations present in the Alpha and Delta
141 variants because these were dominant variants in our study area during the time of our sample
142 collection (Oh et al., 2022b). For the data quality assurance, we quantified pepper mild mottle
143 virus (PMMOV), Tulane virus (TV), and BCoV, which were used to normalize SARS-CoV-2 to
144 human fecal matter, determine the impact of PCR inhibitors, and calculate SARS-CoV-2 recovery
145 efficiency, respectively.

146 The details for RT-qPCR assays are summarized in **Table S3,** following MIQE guidelines.
147 We analyzed five targets (i.e., N gene, Alpha, Delta, PMMOV, and TV) by Taqman-based RT-
148 qPCR assays. We designed two duplex RT-qPCR assays and one singleplex RT-qPCR to analyze
149 these five targets. Specifically, we used the first duplex RT-qPCR assay to detect total SARS-
150 CoV-2 RNA (i.e., N gene) and the Alpha variant. The second one was to analyze the Delta variant
151 and PMMoV. To validate the two duplex RT-qPCR assays, we compared Cq values of serial
152 dilutions of synthetic RNA/DNA standard controls determined by duplex RT-qPCR and its
153 corresponding singleplex RT-qPCR. We used singleplex RT-qPCR to measure TV RNA. The
154 Taqman-based RT-qPCR started with mixing 5 μL of RNA sample, 5 μL of Taqman Fast Virus 1-
155 step Master Mix (4444432, Applied Biosystems, USA), 1 μL of primers/probe mixture for each
156 assay (i.e., final concentrations of 400 nM for primers and 200 nM for probes), and various
157 volumes of nuclease-free water that fill up the mixture to 20 μL. For example, we added 1 μL of

158 primer/probe for the N gene, 1 μL of primer/probe for the Alpha variant, and 8 μL of nuclease-
159 free water to the RT-qPCR cocktail for the first duplex RT-qPCR assay while 1 μL of primer/probe
160 for TV and 9 μL of nuclease-free water to the RT-qPCR cocktail for the singleplex RT-qPCR assay.
161 The PCR cocktail was placed in 384-well plates (4309849, Thermofisher Scientific, USA) and
162 analyzed by QuantStudio 7 Flex (Thermofisher, USA) with a thermal cycle of 5 minutes at 50°C,
163 20 seconds at 95°C followed by 45 cycles of 15 seconds at 95°C and 60 seconds at 60°C. The PCR
164 standard curves were obtained for every RT-qPCR analysis with 10-fold serial dilutions of
165 synthetic RNA or DNA controls, and average PCR efficiencies for RT-qPCR were 92% (**Fig. S1**).
166 Four replicates of nuclease-free water were run for every RT-qPCR analysis to identify the
167 contamination of reagents or cross-contamination among reactions. These non-template controls
168 test negative at all RT-qPCR analyses. We also determined the limit of detection (LOD) and limit
169 of quantification (LOQ) for each assay with 20 replicates of serial dilutions of synthetic controls
170 (**Table S4**) following a previous study (Oh et al., 2022b).

171 We analyzed BCoV RNA by a SYBR-based RT-qPCR assay. The RT-qPCR mixture for
172 SYBR-based RT-qPCR assay included 3 μL of RNA sample, 0.3 μL of 10 μM forward and reverse
173 primer, 1.275 μL of molecular biology grade water (Corning, NY, USA), 5 μL of 2×iTaq universal
174 SYBR green reaction mix, 0.125 μL of iScript reverse transcriptase from the iTaq™ Universal
175 SYBR® Green One-Step Kit (1725151, Bio-Rad Laboratories, USA). The PCR cocktail was
176 placed in 96-well plates (4306737, Applied Biosystems, USA) and analyzed by an RT-qPCR
177 system (QuantStudio 3, Thermo Fisher Scientific, USA). The RT-qPCR reaction was performed
178 with a thermocycle of 50°C for 10 minutes and 95°C for 1 minute, followed by 40 cycles of 95°C
179 for 10 seconds and 60°C for 30 seconds. Melting curves were analyzed while the temperature
180 increased from 60°C to 95°C. Based on the melting curves, the primers were specifically bound to

181 the target genome. The SYBR signal was normalized to the ROX reference dye. The cycles of
182 quantification (Cq) were determined by QuantStudio Design & Analysis Software (v1.5.1). The
183 numbers of technical replicates were 4 for synthetic RNA controls and from 3 to 5 for sewage
184 samples. The PCR efficiencies for RT-qPCR were higher than 85% (R^2>0.99).

185 Once RNA concentrations in the final extracts were determined, we calculated RNA
186 concentrations in sewage samples using Eqs. 1-4.

187

188 $$C_{sludge} = C_{RNA\ extract} \times \frac{V_{RNA\ extract}}{V_{sludge}} \quad \text{(Eq. 1)}$$

189 $$RE = \frac{C_{BCoV\ in\ sludge} \times V_{sludge}}{C_{BCoV\ in\ stock\ solution} \times V_{stock\ solution}} \quad \text{(Eq. 2)}$$

190 $$CF = \frac{V_{sludge}}{V_{initial\ sewage}} \quad \text{(Eq. 3)}$$

191 $$C_{initial\ sewage} = C_{sludge} \times \frac{CF}{RE} \quad \text{(Eq. 4)}$$

192

193

194 Where $C_{RNA\ extract}$, C_{sludge}, and $C_{initial\ sewage}$ indicate RNA concentration in RNA extract, sludge, and
195 initial sewage, respectively. $V_{RNA\ extract}$, V_{sludge}, and $V_{initial\ solution}$ mean volume of RNA extract,
196 sludge, and initial sewage, respectively. RE and CF represent recovery efficiency and
197 concentration factor, respectively.

198

199 **Testing for PCR inhibitors**

200 We evaluated the impacts of potential remaining PCR inhibitors in the RNA extracts by spiking
201 RNA extracts with TV RNA. TV is a calicivirus that infects rhesus monkies, so TV should not be
202 present in our collected sewage samples. We added 1 μL of 10^6 gc/μL TV RNA to 10 μL RNA

203 extracts and 10 μL molecular biology-grade water (Millipore Sigma, Burlington, MA, USA). We
204 assumed that if an RNA extract includes PCR inhibitors, a Cq value for TV in the RNA extract
205 will be higher by 1 Cq value or more than in inhibition-free water.
206
207 **Presentation of viral genome concentrations**
208 If one of the technical replicates for RT-qPCR is undetermined until 45 PCR cycles, then viral
209 genome concentrations are not quantitatively reliable (Safford et al., 2022; **Fig. S2A**). Hewitt et al.
210 (2022) and **Fig. S2B** showed that positivity/negativity (Eq. 5) is a better indicator for the actual
211 genome concentrations than the Cq values when samples have undetermined RT-qPCR results.
212 We used standard curves to convert Cq values to concentrations when none of the technical RT-
213 qPCR replicates showed undetermined. On the other hand, when undetermined data was reported,
214 we used positivity to estimate viral RNA concentrations (Eq. 6).
215
216 $$P_S = \frac{The\ number\ of\ positive\ replicates}{Total\ number\ of\ replicates} \quad \text{(Eq. 5)}$$
217 $$C_S = C_L \times P_S \quad \text{(Eq. 6)}$$
218
219 Where C_S represents the N gene concentration of samples containing undetermined qPCR results
220 in technical replicates. C_L indicates the lowest N gene concentration from samples that do not have
221 undetermined qPCR results. In addition, as suggested in previous studies (Simpson et al., 2021;
222 Wolfe et al., 2021; Zhan et al., 2022), we presented SARS-CoV-2 RNA concentrations normalized
223 to PMMOV for comparisons with clinical epidemiology.
224

225 **Using BCoV spiked experiments to determine sample composite period**

226 We put 35 mL of sewage sample and 350 μL of 10^5 gc/μL BCoV particles in a 50 mL sterile tube.
227 The five tubes of the sewage samples containing BCoV were incubated in a water bath where the
228 temperature was maintained at 25°C. We collected samples every 24 hours for five days. The
229 number of BCoV RNA was determined by the extraction method and RT-qPCR assays mentioned
230 above. We assumed that the RNA decay follows the first-order decay model (Eq. 7).

231

232 $C_t = C_0 e^{-kt}$ (Eq. 7)

233 $lnC_t = lnC_0 - kt$ (Eq. 8)

234

235 Where C_t and C_0 are the concentration of BCoV RNA at time t and zero, k is the first-order decay
236 rate constant. The rate constant is equal to the slope of ln C and the time graph (Eq. 8), which was
237 determined by the linear least-squares regression model.

238

239 **Clinical epidemiology**

240 The Champaign-Urbana Public Health District (CUPHD) provided daily COVID-19 cases with
241 spatiotemporal information, with which we determined the 7-day average and relative infection
242 case. The relative infection case refers to the number of patients who actively contribute to sewage
243 viral load. The relative infection case is determined by multiplying temporal daily COVID-19
244 cases by a virus shedding model (He et al., 2020; Hewitt et al., 2022). We determined the COVID-
245 19 incidence rate for each sewershed by normalizing the COVID-19 occurrence (cases/day) to the
246 conresponding populations.

247

248 $$Incidence\ rate\ (\%) = \frac{COVID\text{-}19\ occurrence\ (cases/day)}{population\ in\ a\ sewershed\ (people)} \times 100 \qquad \text{(Eq. 9)}$$

249

250 We analyzed 138 sequences from the Global Initiative on Sharing All Influenza Data (GISAID),
251 which were reported from Champaign County in 2021 using an algorithm, PRIMES developed by
252 Oh et al. (2022b) to determine the temporal prevalence of SARS-CoV-2 variants in Champaign
253 County.

254

255 **Statistical analysis**

256 We conducted various statistical analyses to support the credibility of data comparisons.
257 Specifically, BCoV RNA concentrations in sewage with different incubation times were analyzed
258 by Mann-Whitney U Test and a linear least-squares regression model (**Fig. S3**). PMMOV and
259 BCoV RNA concentrations were not normally distributed, so Kruskal Wallis ANOVA was used
260 to compare these data across the seven sewersheds (**Fig. S4** and **Fig. S5**). The Cq values of serial
261 dilutions determined by multiplex and singleplex RT-qPCR assays were compared by F-tests (**Fig.**
262 **S7**). In addition, N/PMMOV was not normally distributed, so we used Spearman's rank correlation
263 to compare it to clinical epidemiology (**Table S5**). These statistical analyses were performed using
264 OriginPro (version 2019b). Furthermore, we analyzed the possibility of detecting the SARS-CoV-
265 2 N gene at a given COVID-19 incidence by a binary logistic regression model. This statistical
266 analysis was performed using the *statsmodels* library in Python (version 3.8.8) (Seabold and Josef
267 Perktold, 2010) and the algorithm is deposited at
268 https://github.com/Nguyen205/Logistic_regression.

269

Results and Discussion

Quality assurance and quality control

We first identified what an appropriate sample composite period was to establish an efficient WBE procedure by spiking samples with known amounts of bovine coronavirus (BCoV) RNA and examining its stability over time. **Fig. S3** used controlled laboratory conditions and showed the concentrations of BCoV RNA as the incubation time increases. The reduction in BCoV RNA concentrations followed the first-order decay kinetics (R^2=1.00) with a first-order decay rate constant of 0.22 day^{-1}, which is within the range of previously reported rate constants for SARS-CoV-2 and PMMOV (Roldan-Hernandez et al., 2022). Under these conditions, BCoV RNA concentrations were not significantly different from initial concentrations to four days post-incubation (Mann-Whitney U Test; $p>0.05$). Ahmed et al. (2020a) found similar results when using SARS-CoV-2.The temperatures measured at our sampling sites were lower than 25°C throughout the monitoring period. Therefore, we were assured that coronavirus RNA genomes would remain stable when we collected 4-day 4-day composite sewage samples.

PMMOV RNA concentrations were quantified for each composite sample as an indicator that fecal material was present in a composite sample (D'Aoust et al., 2021). The PMMOV RNA concentrations for all collected samples across the seven sewersheds are different (Kruskal Wallis ANOVA; $p<0.05$, **Fig. S4)**. This was expected because of the heterogeneity in the sewage composition. Median values of PMMOV RNA concentrations from the seven sewersheds ranged from 6.41 (C2) to 7.05 gc/L (R1), which were similar to those previously reported (Feng et al., 2021). Since PMMOV was detected from all collected sewage samples, our sampling method was indeed successfully collecting human fecal matter from sewer distribution systems.

292 We used the recovery efficiency for BCoV to evaluate the sample processing procedures
293 (e.g., sludge concentration and RNA extraction). The mean, median, minimum, and maximum
294 recovery efficiencies across the seven locations were 3.4%, 1.5%, 0.02%, and 64%, respectively.
295 The recovery efficiencies were significantly different (Kruskal Wallis ANOVA; $p<0.05$, **Fig. S5**).
296 A wide range of recovery efficiency was also reported previously (Feng et al., 2021), and the
297 differences were probably attributed to different sewage characteristics. Pecson et al. (2021)
298 investigated the reproducibility and sensitivity of 36 virus concentration methods performed by 32
299 laboratories across the U.S. and suggested a threshold of 0.01% of recovery efficiency for
300 exclusion of poorly processed samples. Based on this criterion, none of the samples were screened
301 out due to the low quality of sample processing.

302 Sewage samples routinely contain PCR inhibitors (Rački et al., 2014). Thus, it is essential
303 to determine the impact of remaining PCR inhibitors in the final RNA extracts. We conducted
304 experiments to determine the presence of RNA inhibitors by spiking samples with Tulane Virus
305 (TV) RNA. Results are shown in **Fig. S6**. We found all ΔCq values were within 1.0, confirming
306 that PCR inhibitors' impacts in our RNA samples were negligible after 10-fold dilution.

307 We used multiplex RT-qPCR assays to quantify the SARS-CoV-2 N gene, the S gene for
308 the Alpha variant, the S gene for the Delta variant, and PMMOV replication gene in the final RNA
309 extracts to save samples, reagents, and time. **Figure S7** shows that that Cq values from singleplex
310 and multiplex RT-qPCR were not significantly different for the two sets of multiplex RT-qPCR
311 assays (F-test; $p>0.05$). Thus, the fluorescence from FAM dye did not affect the HEX channel and
312 vice versa. As a final negative control, we applied the same procedures, including sample
313 collection, processing, and analysis, to wastewater discharged from a food processing plant and
314 sewage samples. We did not detect any SARS-CoV-2 RNA signal (i.e., Cq values were

315 undetermined by 45 cycles of RT-qPCR) from the food processing wastewater samples throughout
316 the monitoring period. This finding suggests no cross-contamination among samples from
317 different sites, so the positive Cq values were true positives.
318
319 **Comparisons of clinical epidemiology and wastewater-based epidemiology data**
320 WBE has been evaluated based on the goodness of fit to clinical epidemiology determined by
321 statistical models such as a Pearson correlation coefficient or a Spearman rank correlation
322 coefficient (Barua et al., 2022; Li et al., 2021), depending on the data normality. In a similar
323 manner, we compared N/PMMOV ratios to COVID-19 incidence rates determined by three
324 epidemiological parameters: daily COVID-19 cases (Xiao et al., 2022), 7-day average COVID-19
325 cases (Hoar et al., 2022), and relative infection COVID-19 cases (Hewitt et al., 2022). Then, we
326 evaluated the goodness of fit by Spearman rank correlation (**Table S5**) between the N/PMMOV
327 ratios and these epidemiological parameters. We found that the relative infection cases yielded
328 higher correlation coefficients than daily cases and 7-day averages at six sewersheds (C1, C2, C3,
329 C4, R1, and R2). Therefore, we used relative infection cases to determine the COVID-19 incidence
330 rate.
331 The correlation coefficient between COVID-19 incidence rate and N/PMMOV ranged
332 from 0.08 (R3) to 0.72 (C1) across the seven sewersheds (**Table S5**). Overall, the correlation
333 coefficients showed a considerable variation among the sewersheds. There are several possible
334 explanations for the lack of uniform correlations. First, the average COVID-19 incidence rate was
335 0.004% throughout the monitoring period, which is equivalent to only a few daily COVID-19 cases
336 in neighborhood-scale sewersheds serving populations of around 1,400 people. **Fig. 2A** shows no
337 COVID-19 patients were reported at the seven sewersheds for 1,544 days out of 2,104 days of

338 monitoring periods (i.e., 73.4%). One, two, and three or more patients were reported at 18.7, 4.9,
339 and 3.0%, respectively. Individual patients have significant differences in clinical symptoms, viral
340 shedding, and access to COVID-19 tests. For example, Ke et al. (2022) found significant person-
341 to-person variations in viral load and virus shedding duration among COVID-19 patients. A time
342 lag between symptom onset and testing date also varies because people have different access to
343 COVID-19 testing. The fact that the different time lags significantly impact the correlation to the
344 WBE data is demonstrated by the highest correlation coefficient (0.72) at location C1. Most
345 residents of location C1 were undergraduate students enrolled at the University of Illinois Urbana-
346 Champaign, which required students to undergo routine COVID-19 testing. For example, graduate
347 students were required to get tested once a week until August 2021. Undergraduate students were
348 requird to get tested twice a week until May 2021, and then once a week from May to August 2021.
349 While the mandatory COVID-19 testing was lifted for fully vaccinated individuals in September
350 2021, many residents in campus town were still required weekly testing. Therefore, compared to
351 other sites, the reported COVID-19 cases at C1 were much closer to the actual infection cases. The
352 individual heterogeneity in clinical symptoms including virus shedding and access to COVID-19
353 tests will become more significant as the number of COVID-19 cases decreases. Second, SARS-
354 CoV-2 RNA was present at a low concentration in sewage samples. The SARS-CoV-2 N gene
355 concentrations are summarized in three groups (**Fig. 2B)**; the N gene was not detected; the N gene
356 was detected in at least one of the qPCR replicates, but the concentrations were below LOQ; the
357 N gene concentrations were above the LOQ. We found that 92% of sewage samples showed
358 concentrations below the LOQ. By definition, the concentrations below LOQ are not quantitatively
359 reliable, and the concentrations may differ from the actual virus concentrations. Therefore, we
360 concluded that the unclear relationship between the WBE and the clinical epidemiological data

was attributed to a low number of COVID-19 cases in sewersheds and low SARS-CoV-2 concentrations in sewage.

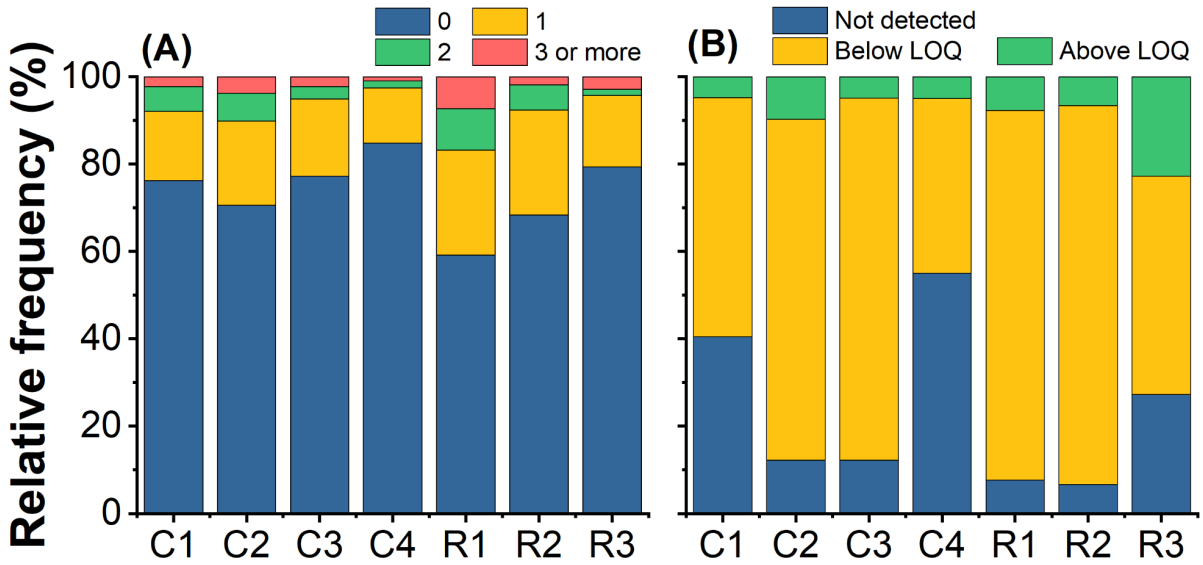


Fig 2. Stacked histograms for (A) COVID-19 daily cases in sewersheds and (B) SARS-CoV-2 concentrations in sewage.

A COVID-19 warning system for neighborhood-scale sewersheds with low COVID-19 incidences

We developed a COVID-19 warning system for a high COVID-19 incidence by setting a threshold of N/PMMOV to separate the monitoring periods into high- or low-risk. Specifically, the sewershed was classified as a high-risk period if N/PMMOV was higher than the threshold from two consecutive monitoring results. Conversely, this warning was dismissed when the N/PMMOV values were below the threshold for the two samples in a row. We decided to make decisions with two samples with the same results to exclude uncertainty in the measurement.

We applied the COVID-19 warning system with a N/PMMOV threshold of $10^{-2.0}$ to the seven sewersheds. Then, we presented the entire monitoring period classified into low (blue) and

378 high (pink) risk periods in **Fig. 3**. As a result, there were 26 local outbreaks with a COVID-19
379 incidence rate higher than 0.5% (**Table S6**), and the COVID-19 warning system alarmed the 19
380 outbreaks (73% of sensitivity). Also, the warning system assigned 17 high-risk periods for the
381 seven sewersheds, among which 12 high-risk periods were related to the local outbreaks of
382 COVID-19 with higher than 0.5% (71% of specificity). These findings suggest that the high-risk
383 periods determined by the COVID-19 warning system aligned with the local COVID-19 outbreaks
384 at the neighborhood-scale sewersheds.
385

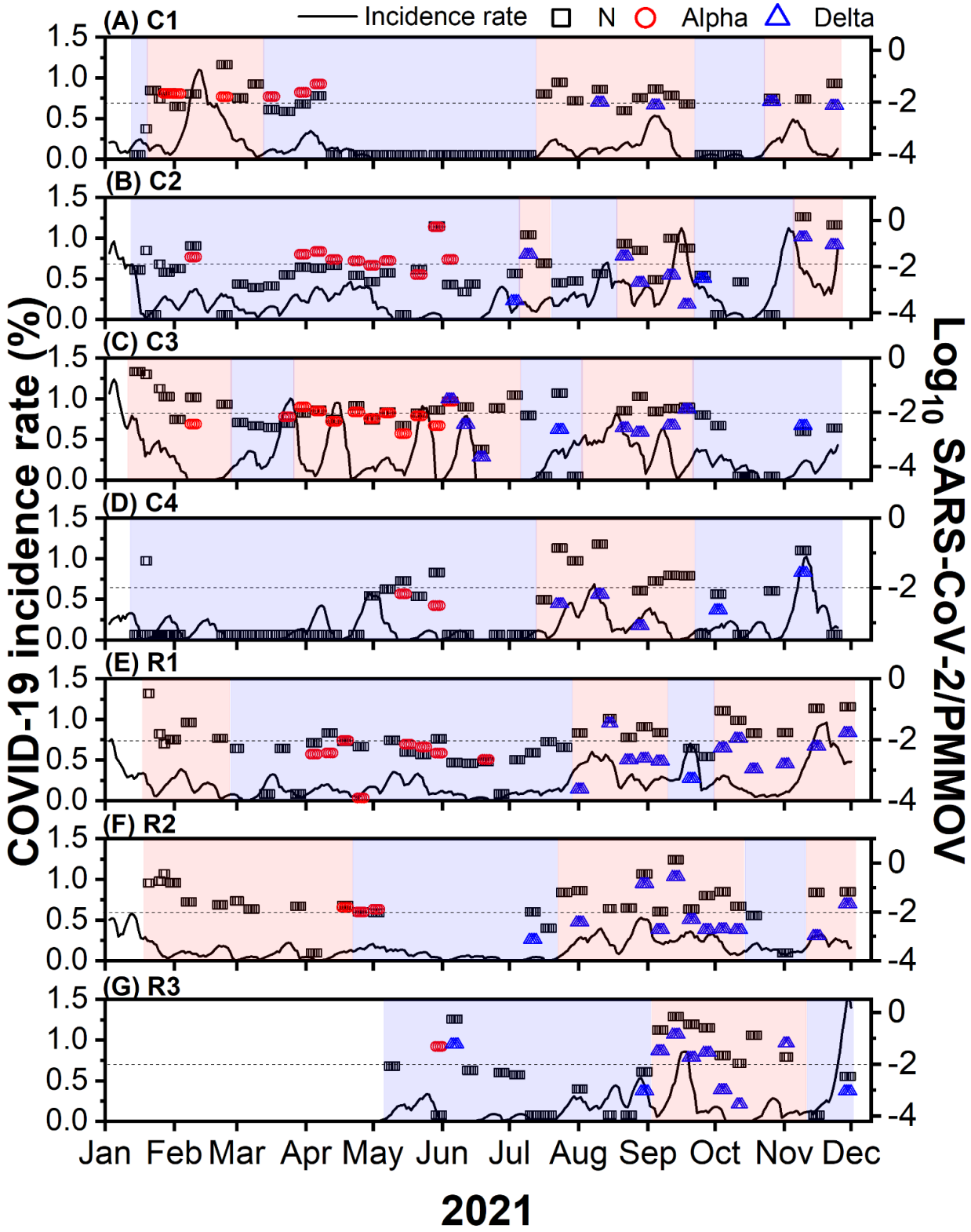


Fig. 3. Comparisons between clinical epidemiology (COVID-19 incidence rate) and wastewater-based epidemiology (N/PMMOV). Open symbols represent WBE results (i.e., rectangles for the

389 N gene, circles for the Alpha variant, and triangles for the Delta variant). We assigned about 10^{-4}
390 of N/PMMOV to samples tested negatives for the N gene to differentiate these samples from
391 samples not measured. Black solid curves show the COVID-19 incidence rates. The horizontal
392 dashed lines show the threshold N/PMMOV of $10^{-2.0}$. The monitoring period is separated either
393 into the high-risk (pink) or the low-risk (blue).
394

395 We also found that the sensitivity and specificity of the COVID-19 warning system to the
396 local outbreaks changed with different thresholds of N/PMMOV. **Fig. 4** showed the high-risk
397 periods at C1 with different thresholds. We found that several low-risk periods changed to the
398 high-risk as the threshold decreased. For example, November was a low-risk period with a
399 N/PMMOV threshold of $10^{-1.8}$ (**Fig. 4A**) but switched to a high-risk period with a threshold of 10^{-}
400 $^{2.0}$. Likewise, C1 is classified as a low-risk period in April with a threshold of $10^{-2.0}$, but it changes
401 to a high-risk period when the threshold decreases to $10^{-2.1}$. As a more extended period is classified
402 into the high-risk, the sensitivity increases while the specificity decreases. **Fig. S8 and S9** show
403 the COVID-19 warning systems applied to the seven sewersheds with the thresholds of $10^{-1.9}$ and
404 $10^{-2.1}$, respectively. We found that the sensitivity of the warning system increased from 54% to 81%
405 while the specificity decreased from 75% to 61% as the thresholds changed from $10^{-1.9}$ to $10^{-2.1}$.
406 This finding suggests that we can increase the possibility of detecting local outbreaks by
407 decreasing the thresholds, although the probability of not including the outbreak in high-risk
408 periods may increase (i.e., decreases in specificity).
409

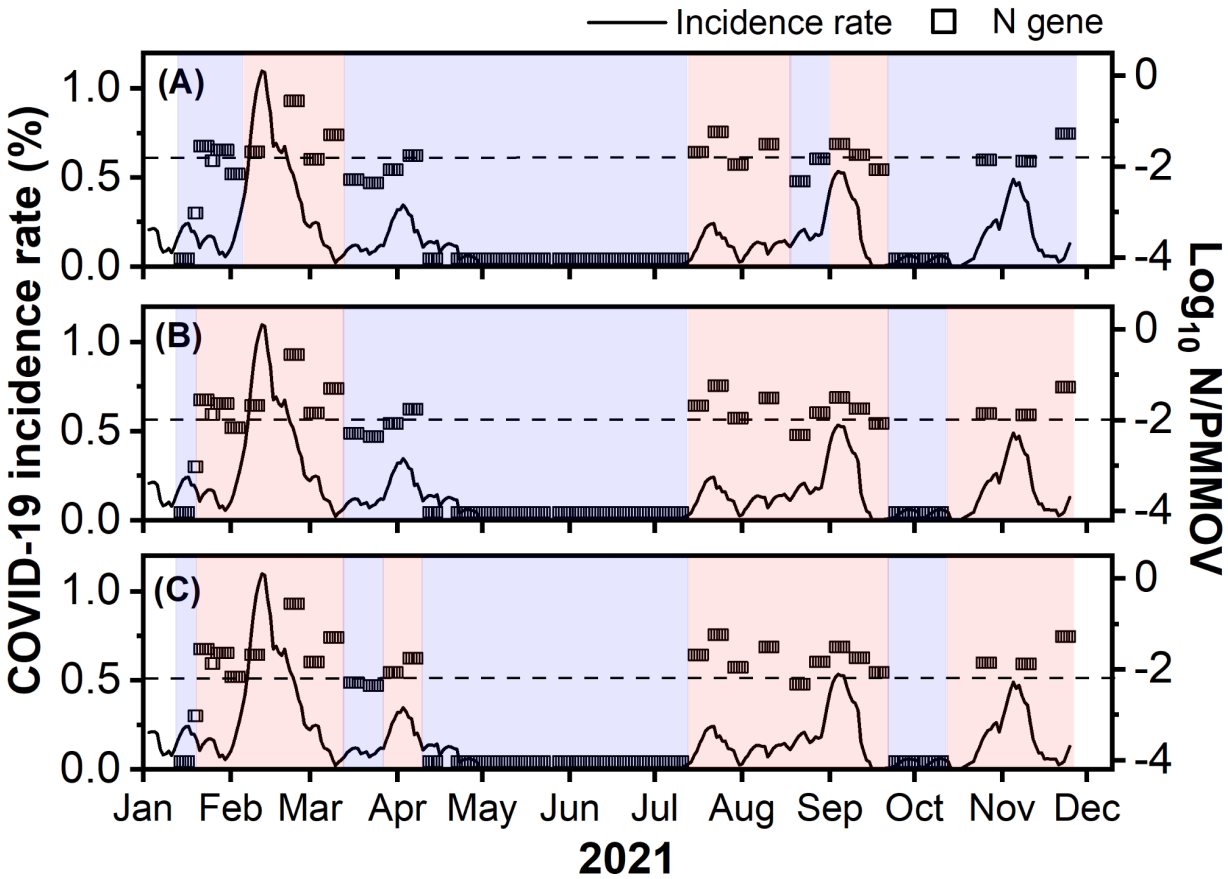


Fig. 4. Application of the COVID-19 warning system to location C1 with different thresholds; (A) $10^{-1.8}$, (B) $10^{-2.0}$, and (C) $10^{-2.1}$ of N/PMMOV. The high-risk and the low-risk periods are presented in red and blue shades, respectively.

We also compared WBE data to the county-wide COVID-19 incidence rate. **Fig. 5A** shows the COVID-19 incidence rate in Champaign County, which includes all seven sewersheds during the same monitoring period as our WBE program. There were four county-wide outbreaks in descending order of disease incidence rate in September, November, February, and April (Arrows in **Fig. 5A** point each peak). According to our COVID-19 warning system, all seven sewersheds were high-risk periods in September when the disease incidence was the highest during the monitoring period (the Delta variant outbreak) **(Fig. 3)**. Four sewersheds (C1, C2, R1, and R2) were high-risk periods in November (the beginning of the 2022 winter outbreak). Also, four sewersheds (C1, C3, R1, and R2) were high-risk periods in February (the end of the 2021 winter

outbreak). Finally, only one site (C3) was the high-risk period in April when disease incidence was the lowest among the four county-wide outbreaks. The location C3 showed clear peaks of the COVID-19 incidence rate from April to June (**Fig. 3C**), supporting that this high-risk period is a true positive. The finding that the number of sewersheds in high-risk periods is related to the scale of the county-level outbreaks confirms the credibility of the COVID-19 warning system.

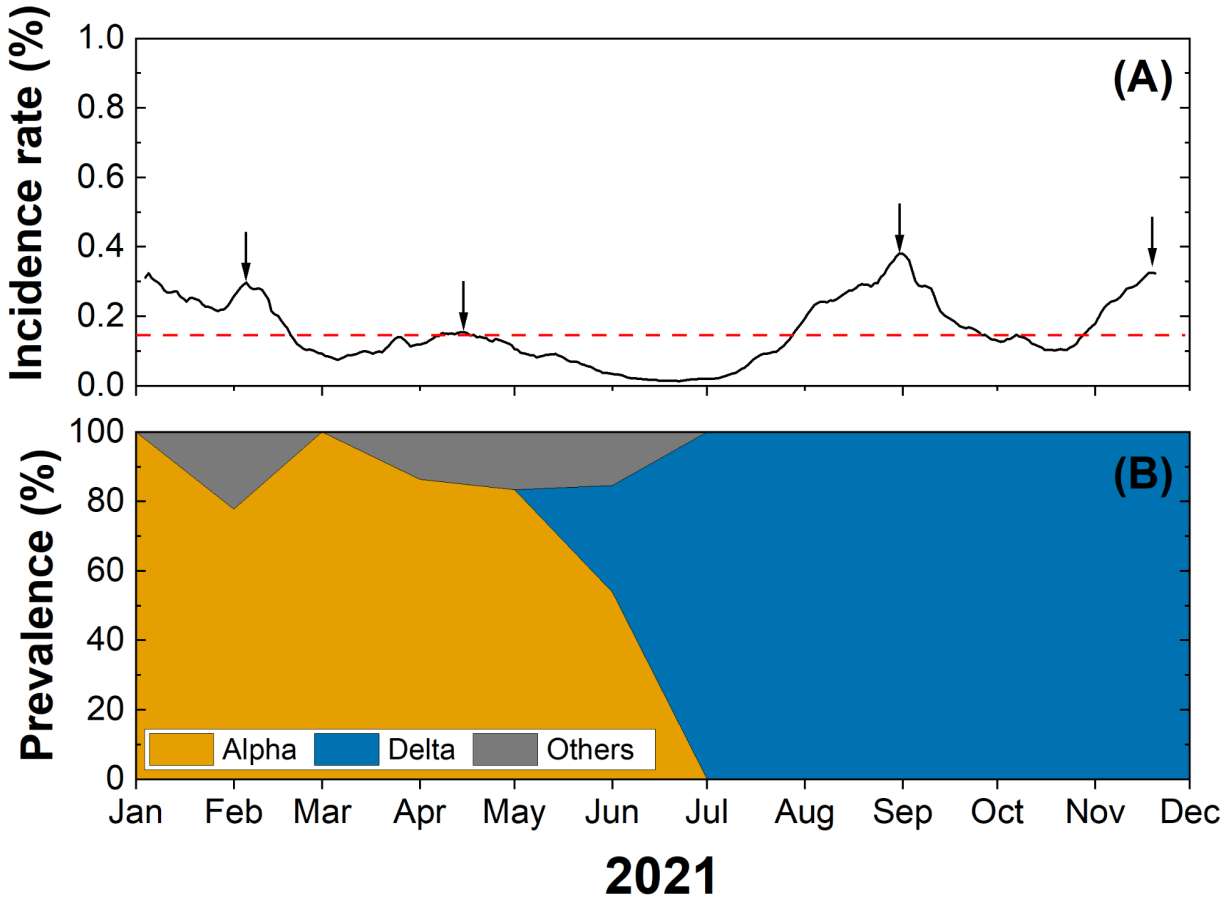


Fig. 5. COVID-19 clinical epidemiology data from Champaign County. (A) COVID-19 incidence rate and (B) the prevalence of SARS-CoV-2 variants. The red dashed line in (A) indicates a 0.14% of incidence rate with which viruses in sewage discharged from a county-scale sewershed are expected to be detected with 90% probability. The variant prevalence in (B) was determined based on genome sequences deposited at GISAID (n=138).

437 The high-risk period at C3 from April to June further highlighted the advantage of the
438 COVID-19 warning system. We determined a sensitivity of our WBE procedure, defined as the
439 COVID-19 incidence rate at which viruses in sewage are detected at 90% probability. The data
440 from C1 was used for a binary logistic regression analysis because the COVID-19 incidence rate
441 is most reliable in C1 due to the mandatory COVID-19 testing. We found that the COVID-19
442 incidence rate is a reliable predictor to explain the probability of SARS-CoV-2 N gene detection
443 (McFadden; $p<0.001$). Based on the logistic regression curve in **Fig. 6**, the WBE sensitivity was
444 0.14% (or one infection case out of 715 residents). Meanwhile, the COVID-19 incidence rate of
445 the county-wide outbreak in April peaked at 0.15%. In addition, the incidence rate was less than
446 the WBE sensitivity (0.14%) from February 20 to July 27. If the WBE had been applied to the
447 entire county, local outbreaks such as the high-risk period at C3 from April to June would have
448 been overlooked. This finding suggests that wastewater surveillance should be performed at
449 neighborhood-scale sewersheds to discern local outbreaks when the COVID-19 incidence rate is
450 low.
451

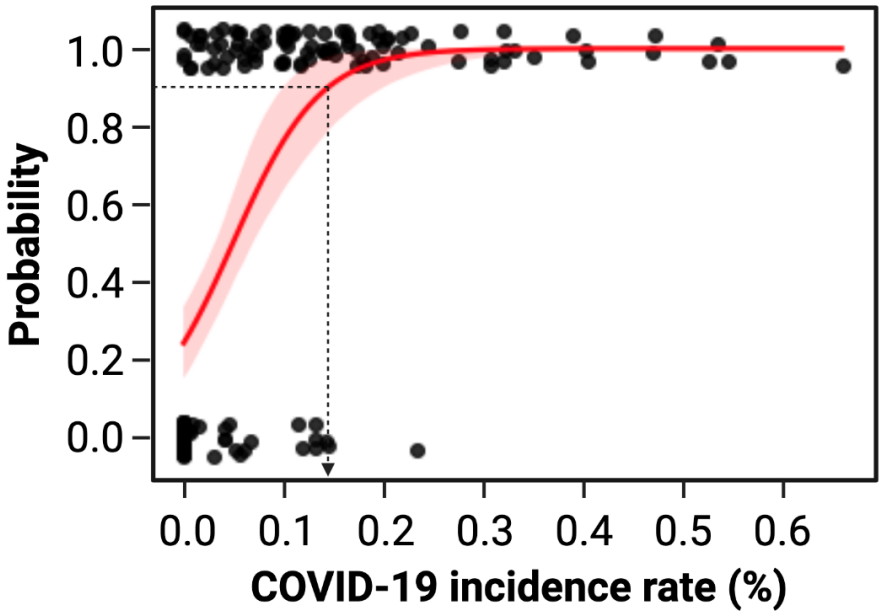


Fig. 6. Probability of detecting SARS-CoV-2 N gene in sewage samples from location C1. Black closed-circles present COVID-19 incidence rates on the x-axis. Either 1 or 0 was given to the y-value when sewage samples tested positive or negative for SARS-CoV-2, respectively. Logistic regression curves with a 95% confidence interval are presented in a red solid line and shade. McFadden's Pseudo R-squared of the regression analysis was 0.34 ($p<0.001$), which is considered a good fit (McFadden, 2021). The Black dotted line indicates that the 0.14% incidence rate results in a 90% probability.

We also applied RT-qPCR assay data that specifically detected the Alpha and Delta SARS-CoV-2 variants to the sewage samples (**Fig. 3**). In our sewersheds, the Alpha variant was first detected on January 27, and it continued to be reported until June 23. Also, the Delta variant was detected from June 3 to the end of the sampling (Dec. 2, 2021). In general, the periods when the Alpha and the Delta variant were detected aligned with the prevalence of SARS-CoV-2 variants determined by genome sequencing of clinical samples collected from COVID-19 patients in Champaign County, IL, USA (**Fig. 5B**). Furthermore, we found that Delta variant's shedding into the sewersheds might have increased the COVID-19 incidence rates. For example, five sites (C2, C3, C4, R1, and R2) were classified as the high-risk period as soon as the Delta variant was detected. In addition, the local outbreak in June at C3, the only high-risk period from the seven sewersheds,

471 was also associated with the Delta variant. This finding suggests that variant-specific RT-qPCR
472 assay can be combined with the COVID-19 warning system to determine the introduction of
473 emerging or dominant SARS-CoV-2 variants in neighborhood-scale sewersheds even at a low
474 COVID-19 incidence.
475
476 **Implication**
477 We developed a COVID-19 warning system that is activated when the ratio of N/PMMOV
478 concentrations are higher than $10^{-2.0}$ for two consecutive analyses. The COVID-19 warning system
479 showed 73% sensitivity and 71% specificity to neighborhood-scale outbreaks with an incidence
480 rate higher than 0.5%. Note that these specific sensitivity and specificity values should be
481 interpreted with caution. For example, the sensitivity varies depending on how we define a local
482 outbreak of interest. Although we chose an incidence rate of 0.5% for studies here, a higher
483 COVID-19 incidence rate such as 0.75% may be more appropriate for other regions. When we
484 examine local outbreaks with an incidence rate higher than 0.75% into account (12 outbreaks in
485 total), 10 outbreaks are now included in the high-risk period, and sensitivity increases from 71%
486 to 83%. In addition, it is likely that there were unreported COVID-19 cases in our sewersheds. For
487 example, the CDC estimates actual COVID-19 infections are from 6 to 24 times higher than
488 reported cases (Havers et al., 2020). These unreported COVID cases might have accounted for the
489 false-positive scenarios where the high-risk periods were made without the local outbreaks. If so,
490 the true specificity becomes higher than 71%.
491 Although the N/PMMOV threshold of $10^{-2.0}$ was used for our seven sewersheds, the
492 COVID-19 warning system can be applied with different thresholds to the other sites considering
493 regional characteristics such as local public health policies and the community's acceptance of the

494 COVID-19 risk (Dryhurst et al., 2020; Turska-Kawa and Pilch, 2022). Intensive care unit (ICU)
495 occupancy rate, influenced by vulnerable populations and ICU bed capacity, may be one of the
496 regional characteristics that should be considered to determine a threshold for the COVID-19
497 warning system. If some regions had experienced a high ICU occupancy rate during the COVID-
498 19 pandemic, COVID-19 put enormous stress on the local medical systems. Similar scenarios
499 could happen again in the future. According to CDC, the maximum ICU occupancy rate
500 significantly differs from site to site (http://covid.cdc.gov/covid-data-tracker). For example, in the
501 state of Illinois (USA), some counties, such as Cook County and Champaign County, have
502 managed ICU beds well during the COVID-19 pandemic, showing 35% of maximum ICU bed
503 occupancy. On the other hand, other counties, including Rock Island County (85%) and Effingham
504 County (70%), experienced a larger burden on ICU beds during the COVID-19 pandemic.
505 Therefore, these counties may need to be more vigilant about the COVID-19 outbreak by setting
506 a lower threshold.

507 Applying WBE to neighborhood-scale sewersheds may cost more than large size
508 sewersheds for monitoring the entire administrative area. To efficiently monitor multiple
509 neighborhood-scale sewersheds, the WBE can be deployed to some representative places of
510 diverse societal characteristics, such as socioeconomic status, demographics, or land use category,
511 so that we can minimize the number of monitoring sites while obtaining information about the
512 COVID-19 incidence representative of the entire area. The COVID-19 pandemic disproportionally
513 has impacted people depending on demographic and socioeconomic status (Bassett et al., 2020;
514 Krieger, 2020). Thus, those underserved regions could be chosen for wastewater surveillance. For
515 example, we chose the six sewersheds (C2, C3, C4, R1, R2, and R3) based on areas that reported
516 fewer cases of COVID-19 testing than the average testing in the Champaign County to attempt to

517 supplement clinical testing data, while we selected another sewershed to represent the highly-
518 tested campus area (location C1) (**Table S7**). In addition, sample collection frequency can be
519 optimized to operate WBE within a budget. Schoen et al. (2022) found that a sampling frequency
520 of once every 4 days is not significantly different from daily sampling in estimating clinical
521 epidemiology data. We collected 4-day composite samples to have a less frequent sampling
522 frequency, once every 8.2 days.

523 Our findings on the COVID-19 warning system have significant implications for public
524 health. COVID-19 might be moving from a pandemic to an endemic phase (Veldhoen and Simas,
525 2021), which involves a difference in disease surveillance strategy. In an endemic phase, we may
526 not be able to get as accurate statistics on COVID-19 by the clinical diagnosis as we did in the
527 pandemic phase because people will be less likely to visit hospitals to get tested or medical support.
528 This is because the illnesses may be self-limiting (Veldhoen and Simas, 2021). People also may
529 have less accessibility to testing sites and use rapid COVID-19 test kits instead at home (Kost,
530 2022). In this context, WBE will play a pivotal role in the COVID-19 surveillance because WBE
531 provides unbiased information about all residents in the sewershed. However, unlike during the
532 pandemic, the COVID-19 will be mostly managed at a much lower incidence level in the endemic
533 phase. At the same time, we may see intermittent COVID-19 outbreaks due to an emerging new
534 variant or waning immunity as we have seen measles or influenza outbreaks (Katzourakis, 2022).
535 Lak et al. (2021) and Shi et al. (2021) discovered that local COVID-19 outbreaks can spread
536 through the surrounding regions resulting in other outbreaks. Rader et al. (2020) explained the fast
537 virus transmission to the surrounding areas with frequent contact among residents. These studies
538 highlight the necessity of early detection of local COVID-19 outbreaks (Cariti et al., 2022; Wu et
539 al., 2022a). This study shows that local outbreaks would not be detected if WBE is applied to a

large-sized sewershed. Therefore, the COVID-19 warning system can be applied to neighborhood-scale sewersheds to identify covert local outbreaks in advance and the introduction of emerging SARS-CoV-2 variants, which provides a warning for large-scale outbreaks.

Conclusion

In this study, we applied WBE to seven neighborhood-scale sewersheds (an average catchment population was 1471) for eleven months. We found that WBE data from neighborhood-scale sewersheds may be poorly correlated to clinical testing data when COVID-19 incidence is low. Thus, we developed a COVID-19 warning system for high COVID-19 risk when the SARS-CoV-2 N gene concentration in wastewater is higher than a threshold, which can be assigned in collaboration with public health officials. We demonstrated that the COVID-19 warning system can discern neighborhood-scale COVID-19 outbreaks that a county-scale WBE would not detect due to the low disease incidence. We also identified that the shedding of a new COVID-19 variant (i.e., the delta variant) to neighborhood-scale sewersheds even at the low disease incidence, leading to local COVID-19 outbreaks. Our findings highlight that WBE should be applied to neighborhood-scale sewersheds when COVID-19 incidence is maintained at a low level.

Acknowledgment

We acknowledge the funding from the Grainger College of Engineering and the JUMP-ARCHES program of OSF Healthcare in conjunction with the University of Illinois. The authors also acknowledge Hayden Wennerdahl and Kip Stevenson for sampling deployment, and Yuqing Mao, Matthew Robert Loula, Aashna Patra, Kristin Joy Anderson, Mikayla Diedrick, Hubert Lyu,

Hamza Elmahi Mohamed, Jad R Karajeh, Runsen Ning, Rui Fu, and Kyukyoung Kim for sewage sampling and processing.

567 **References**

568 Abrams, E.M., Szefler, S.J., 2020. COVID-19 and the impact of social determinants of health.
569 Lancet Respir. Med. 8, 659–661. https://doi.org/10.1016/S2213-2600(20)30234-4

570 Ahmed, W., Bertsch, P.M., Bibby, K., Haramoto, E., Hewitt, J., Huygens, F., Gyawali, P., Korajkic,
571 A., Riddell, S., Sherchan, S.P., Simpson, S.L., Sirikanchana, K., Symonds, E.M., Verhagen,
572 R., Vasan, S.S., Kitajima, M., Bivins, A., 2020a. Decay of SARS-CoV-2 and surrogate
573 murine hepatitis virus RNA in untreated wastewater to inform application in wastewater-
574 based epidemiology. Environ. Res. 191, 110092.
575 https://doi.org/10.1016/J.ENVRES.2020.110092

576 Ahmed, W., Bertsch, P.M., Bivins, A., Bibby, K., Farkas, K., Gathercole, A., Haramoto, E.,
577 Gyawali, P., Korajkic, A., McMinn, B.R., Mueller, J.F., Simpson, S.L., Smith, W.J.M.,
578 Symonds, E.M., Thomas, K. V., Verhagen, R., Kitajima, M., 2020b. Comparison of virus
579 concentration methods for the RT-qPCR-based recovery of murine hepatitis virus, a surrogate
580 for SARS-CoV-2 from untreated wastewater. Sci. Total Environ. 739, 139960.
581 https://doi.org/10.1016/J.SCITOTENV.2020.139960

582 Barrios, R.E., Lim, C., Kelley, M.S., Li, X., 2021. SARS-CoV-2 concentrations in a wastewater
583 collection system indicated potential COVID-19 hotspots at the zip code level. Sci. Total
584 Environ. 800, 149480. https://doi.org/10.1016/J.SCITOTENV.2021.149480

585 Barua, V.B., Juel, M.A.I., Blackwood, A.D., Clerkin, T., Ciesielski, M., Sorinolu, A.J., Holcomb,
586 D.A., Young, I., Kimble, G., Sypolt, S., Engel, L.S., Noble, R.T., Munir, M., 2022. Tracking
587 the temporal variation of COVID-19 surges through wastewater-based epidemiology during
588 the peak of the pandemic: A six-month long study in Charlotte, North Carolina. Sci. Total
589 Environ. 814, 152503. https://doi.org/10.1016/J.SCITOTENV.2021.152503

590 Bassett, M.T., Chen, J.T., Krieger, N., 2020. Variation in racial/ethnic disparities in COVID-19
591 mortality by age in the United States: A cross-sectional study. PLOS Med. 17, e1003402.
592 https://doi.org/10.1371/JOURNAL.PMED.1003402
593 Bitter, L.C., Kibbee, R., Jiménez, G.C., Rmeci, B.O.¨, 2022. Wastewater Surveillance of SARS-
594 CoV-2 at a Canadian University Campus and the Impact of Wastewater Characteristics on
595 Viral RNA Detection. ACS ES&T Water.
596 https://doi.org/10.1021/ACSESTWATER.2C00060
597 Bivins, A., Bibby, K., 2021. Wastewater Surveillance during Mass COVID-19 Vaccination on a
598 College Campus. Environ. Sci. Technol. Lett. 8, 792–798.
599 https://doi.org/10.1021/ACS.ESTLETT.1C00519/ASSET/IMAGES/LARGE/EZ1C00519_
600 0001.JPEG
601 Cariti, F., Tuñ As Corzon, A., Fernandez-Cassi, X., Ganesanandamoorthy, P., Ort, C., Julian, T.R.,
602 Kohn, T., 2022. Wastewater Reveals the Spatiotemporal Spread of SARS-CoV-2 in the
603 Canton of Ticino (Switzerland) during the Onset of the COVID-19 Pandemic. ACS ES&T
604 Water. https://doi.org/10.1021/ACSESTWATER.2C00082
605 Catherine Hoar, Francoise Chauvin, Alexander Clare, Hope McGibbon, Esmeraldo Castro,
606 Samantha Patinella, Dimitrios Katehis, J. Dennehy, J., Monica Trujillo, S. Smyth, D.,
607 I. Silverman, A., 2022. Monitoring SARS-CoV-2 in wastewater during New York City's
608 second wave of COVID-19: sewershed-level trends and relationships to publicly available
609 clinical testing data. Environ. Sci. Water Res. Technol. 8, 1021–1035.
610 https://doi.org/10.1039/D1EW00747E
611 Chen, Y., Chen, L., Deng, Q., Zhang, G., Wu, K., Ni, L., Yang, Y., Liu, B., Wang, W., Wei, C.,
612 Yang, J., Ye, G., Cheng, Z., 2020. The presence of SARS-CoV-2 RNA in the feces of

613 COVID-19 patients. J. Med. Virol. 92, 833–840. https://doi.org/10.1002/JMV.25825
614 D'Aoust, P.M., Mercier, E., Montpetit, D., Jia, J.J., Alexandrov, I., Neault, N., Baig, A.T., Mayne,
615 J., Zhang, X., Alain, T., Langlois, M.A., Servos, M.R., MacKenzie, M., Figeys, D.,
616 MacKenzie, A.E., Graber, T.E., Delatolla, R., 2021. Quantitative analysis of SARS-CoV-2
617 RNA from wastewater solids in communities with low COVID-19 incidence and prevalence.
618 Water Res. 188, 116560. https://doi.org/10.1016/J.WATRES.2020.116560
619 Dryhurst, S., Schneider, C.R., Kerr, J., Freeman, A.L.J., Recchia, G., van der Bles, A.M.,
620 Spiegelhalter, D., van der Linden, S., 2020. Risk perceptions of COVID-19 around the world.
621 J. Risk Res. 23, 994–1006.
622 https://doi.org/10.1080/13669877.2020.1758193/SUPPL_FILE/RJRR_A_1758193_SM797
623 7.DOCX
624 Feng, S., Roguet, A., McClary-Gutierrez, J.S., Newton, R.J., Kloczko, N., Meiman, J.G., McLellan,
625 S.L., 2021. Evaluation of Sampling, Analysis, and Normalization Methods for SARS-CoV-2
626 Concentrations in Wastewater to Assess COVID-19 Burdens in Wisconsin Communities.
627 ACS ES&T Water 1, 1955–1965. https://doi.org/10.1021/ACSESTWATER.1C00160
628 Fitzgerald, S.F., Rossi, G., Low, A.S., McAteer, S.P., O'Keefe, B., Findlay, D., Cameron, G.J.,
629 Pollard, P., Singleton, P.T.R., Ponton, G., Singer, A.C., Farkas, K., Jones, D., Graham, D.W.,
630 Quintela-Baluja, M., Tait-Burkard, C., Gally, D.L., Kao, R., Corbishley, A., 2021. Site
631 Specific Relationships between COVID-19 Cases and SARS-CoV-2 Viral Load in
632 Wastewater Treatment Plant Influent. Environ. Sci. Technol. 55, 15276–15286.
633 https://doi.org/10.1021/ACS.EST.1C05029/ASSET/IMAGES/LARGE/ES1C05029_0004.J
634 PEG
635 Gibas, C., Lambirth, K., Mittal, N., Juel, M.A.I., Barua, V.B., Roppolo Brazell, L., Hinton, K.,

636 Lontai, J., Stark, N., Young, I., Quach, C., Russ, M., Kauer, J., Nicolosi, B., Chen, D., Akella,
637 S., Tang, W., Schlueter, J., Munir, M., 2021. Implementing building-level SARS-CoV-2
638 wastewater surveillance on a university campus. Sci. Total Environ. 782, 146749.
639 https://doi.org/10.1016/J.SCITOTENV.2021.146749
640 Hart, O.E., Halden, R.U., 2020. Computational analysis of SARS-CoV-2/COVID-19 surveillance
641 by wastewater-based epidemiology locally and globally: Feasibility, economy, opportunities
642 and challenges. Sci. Total Environ. 730, 138875.
643 https://doi.org/10.1016/J.SCITOTENV.2020.138875
644 Hata, A., Hara-Yamamura, H., Meuchi, Y., Imai, S., Honda, R., 2021. Detection of SARS-CoV-2
645 in wastewater in Japan during a COVID-19 outbreak. Sci. Total Environ. 758, 143578.
646 https://doi.org/10.1016/J.SCITOTENV.2020.143578
647 Havers, F.P., Reed, C., Lim, T., Montgomery, J.M., Klena, J.D., Hall, A.J., Fry, A.M., Cannon,
648 D.L., Chiang, C.F., Gibbons, A., Krapiunaya, I., Morales-Betoulle, M., Roguski, K., Rasheed,
649 M.A.U., Freeman, B., Lester, S., Mills, L., Carroll, D.S., Owen, S.M., Johnson, J.A.,
650 Semenova, V., Blackmore, C., Blog, D., Chai, S.J., Dunn, A., Hand, J., Jain, S., Lindquist, S.,
651 Lynfield, R., Pritchard, S., Sokol, T., Sosa, L., Turabelidze, G., Watkins, S.M., Wiesman, J.,
652 Williams, R.W., Yendell, S., Schiffer, J., Thornburg, N.J., 2020. Seroprevalence of
653 Antibodies to SARS-CoV-2 in 10 Sites in the United States, March 23-May 12, 2020. JAMA
654 Intern. Med. 180, 1576–1586. https://doi.org/10.1001/JAMAINTERNMED.2020.4130
655 He, X., Lau, E.H.Y., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y.C., Wong, J.Y., Guan, Y., Tan,
656 X., Mo, X., Chen, Y., Liao, B., Chen, W., Hu, F., Zhang, Q., Zhong, M., Wu, Y., Zhao, L.,
657 Zhang, F., Cowling, B.J., Li, F., Leung, G.M., 2020. Temporal dynamics in viral shedding
658 and transmissibility of COVID-19. Nat. Med. 2020 265 26, 672–675.

659 https://doi.org/10.1038/s41591-020-0869-5
660 Hewitt, J., Trowsdale, S., Armstrong, B.A., Chapman, J.R., Carter, K.M., Croucher, D.M., Trent,
661 C.R., Sim, R.E., Gilpin, B.J., 2022. Sensitivity of wastewater-based epidemiology for
662 detection of SARS-CoV-2 RNA in a low prevalence setting. Water Res. 211, 118032.
663 https://doi.org/10.1016/J.WATRES.2021.118032
664 Jamal, Y., Gangwar, M., Usmani, M., Adams, A., Wu, C.-Y., Nguyen, T.H., Colwell, R., Jutla, A.,
665 2022. Identification of thresholds on population density for understanding transmission of
666 COVID-19. GeoHealth e2021GH000449. https://doi.org/10.1029/2021GH000449
667 Karthikeyan, S., Ronquillo, N., Belda-Ferre, P., Alvarado, D., Javidi, T., Longhurst, C.A., Knight,
668 R., 2021. High-Throughput Wastewater SARS-CoV-2 Detection Enables Forecasting of
669 Community Infection Dynamics in San Diego County. mSystems 6.
670 https://doi.org/10.1128/MSYSTEMS.00045-21/SUPPL_FILE/MSYSTEMS.00045-21-
671 ST003.DOCX
672 Katzourakis, A., 2022. COVID-19: endemic doesn't mean harmless. Nature 601, 485.
673 https://doi.org/10.1038/D41586-022-00155-X
674 Ke, R., Martinez, P.P., Smith, R.L., Gibson, L.L., Mirza, A., Conte, M., Gallagher, N., Luo, C.H.,
675 Jarrett, J., Zhou, R., Conte, A., Liu, T., Farjo, M., Walden, K.K.O., Rendon, G., Fields, C.J.,
676 Wang, L., Fredrickson, R., Edmonson, D.C., Baughman, M.E., Chiu, K.K., Choi, H., Scardina,
677 K.R., Bradley, S., Gloss, S.L., Reinhart, C., Yedetore, J., Quicksall, J., Owens, A.N., Broach,
678 J., Barton, B., Lazar, P., Heetderks, W.J., Robinson, M.L., Mostafa, H.H., Manabe, Y.C.,
679 Pekosz, A., McManus, D.D., Brooke, C.B., 2022. Daily longitudinal sampling of SARS-
680 CoV-2 infection reveals substantial heterogeneity in infectiousness. Nat. Microbiol. 2022 75
681 7, 640–652. https://doi.org/10.1038/s41564-022-01105-z

682 Kost, G.J., 2022. Diagnostic Strategies for Endemic Coronavirus Disease 2019 (COVID-19)Rapid
683 Antigen Tests, Repeated Testing, and Prevalence Boundaries. Arch. Pathol. Lab. Med. 146,
684 16–25. https://doi.org/10.5858/ARPA.2021-0386-SA
685 Krieger, N., 2020. ENOUGH: COVID-19, structural racism, police brutality, plutocracy, climate
686 change-and time for health justice, democratic governance, and an equitable, sustainable
687 future. Am. J. Public Health 110, 1620–1623. https://doi.org/10.2105/AJPH.2020.305886
688 Lak, A., Sharifi, A., Badr, S., Zali, A., Maher, A., Mostafavi, E., Khalili, D., 2021. Spatio-temporal
689 patterns of the COVID-19 pandemic, and place-based influential factors at the neighborhood
690 scale in Tehran. Sustain. Cities Soc. 72, 103034. https://doi.org/10.1016/J.SCS.2021.103034
691 Li, X., Kulandaivelu, J., Zhang, S., Shi, J., Sivakumar, M., Mueller, J., Luby, S., Ahmed, W., Coin,
692 L., Jiang, G., 2021. Data-driven estimation of COVID-19 community prevalence through
693 wastewater-based epidemiology. Sci. Total Environ. 789, 147947.
694 https://doi.org/10.1016/J.SCITOTENV.2021.147947
695 M. Pecson, B., Emily Darby, N. Haas, C., M. Amha, Y., Mitchel Bartolo, Richard Danielson,
696 Yeggie Dearborn, Giovanni, G.D., Christobel Ferguson, Stephanie Fevig, Erica Gaddis,
697 Donald Gray, George Lukasik, Bonnie Mull, Liana Olivas, Adam Olivieri, Yan Qu,
698 Consortium, S.-C.-2 I., 2021. Reproducibility and sensitivity of 36 methods to quantify the
699 SARS-CoV-2 genetic signal in raw wastewater: findings from an interlaboratory methods
700 evaluation in the U.S. Environ. Sci. Water Res. Technol. 7, 504–520.
701 https://doi.org/10.1039/D0EW00946F
702 McFadden, D., 2021. Quantitative Methods for Analysing Travel Behaviour of Individuals : Some
703 Recent Developments. Behav. Travel Model. 279–318.
704 https://doi.org/10.4324/9781003156055-18

705 Messner, W., Payson, S.E., 2020. Variation in COVID-19 outbreaks at the US state and county
706 levels. Public Health 187, 15–18. https://doi.org/10.1016/J.PUHE.2020.07.035
707 Oh, C., Kim, K., Araud, E., Wang, L., Shisler, J.L., Nguyen, T.H., 2022a. A novel approach to
708 concentrate human and animal viruses from wastewater using receptors-conjugated magnetic
709 beads. Water Res. 212, 118112. https://doi.org/10.1016/J.WATRES.2022.118112
710 Oh, C., Sashittal, P., Zhou, A., Wang, L., El-Kebir, M., Nguyen, T.H., 2022b. Design of SARS-
711 CoV-2 Variant-Specific PCR Assays Considering Regional and Temporal Characteristics.
712 Appl. Environ. Microbiol. 88. https://doi.org/10.1128/AEM.02289-
713 21/SUPPL_FILE/AEM.02289-21-S0001.PDF
714 Rački, N., Dreo, T., Gutierrez-Aguirre, I., Blejec, A., Ravnikar, M., 2014. Reverse transcriptase
715 droplet digital PCR shows high resilience to PCR inhibitors from plant, soil and water
716 samples. Plant Methods 2014 101 10, 1–10. https://doi.org/10.1186/S13007-014-0042-6
717 Rader, B., Scarpino, S. V., Nande, A., Hill, A.L., Adlam, B., Reiner, R.C., Pigott, D.M., Gutierrez,
718 B., Zarebski, A.E., Shrestha, M., Brownstein, J.S., Castro, M.C., Dye, C., Tian, H., Pybus,
719 O.G., Kraemer, M.U.G., 2020. Crowding and the shape of COVID-19 epidemics. Nat. Med.
720 2020 2612 26, 1829–1834. https://doi.org/10.1038/s41591-020-1104-0
721 Roldan-Hernandez, L., Graham, K.E., Duong, D., Boehm, A.B., 2022. Persistence of Endogenous
722 SARS-CoV-2 and Pepper Mild Mottle Virus RNA in Wastewater-Settled Solids. ACS
723 Environ. Sci. Technol. Water.
724 https://doi.org/10.1021/ACSESTWATER.2C00003/ASSET/IMAGES/LARGE/EW2C0000
725 3_0002.JPEG
726 Rusiñol, M., Zammit, I., Itarte, M., Forés, E., Martínez-Puchol, S., Girones, R., Borrego, C.,
727 Corominas, L., Bofill-Mas, S., 2021. Monitoring waves of the COVID-19 pandemic:

728 Inferences from WWTPs of different sizes. Sci. Total Environ. 787, 147463.
729 https://doi.org/10.1016/J.SCITOTENV.2021.147463
730 Safford, H., Zuniga-Montanez, R.E., Kim, M., Wu, X., Wei, L., Sharpnack, J., Shapiro, K., Bischel,
731 H., 2022. Wastewater surveillance for COVID-19 response at multiple geographic scales:
732 Aligning wastewater and clinical results at the census-block level and addressing
733 pervasiveness of qPCR non-detects. medRxiv 2022.01.28.22269911.
734 https://doi.org/10.1101/2022.01.28.22269911
735 Sangsanont, J., Rattanakul, S., Kongprajug, A., Chyerochana, N., Sresung, M., Sriporatana, N.,
736 Wanlapakorn, N., Poovorawan, Y., Mongkolsuk, S., Sirikanchana, K., 2022. SARS-CoV-2
737 RNA surveillance in large to small centralized wastewater treatment plants preceding the
738 third COVID-19 resurgence in Bangkok, Thailand. Sci. Total Environ. 809, 151169.
739 https://doi.org/10.1016/J.SCITOTENV.2021.151169
740 Schoen, M.E., Wolfe, M.K., Li, L., Duong, D., White, B.J., Hughes, B., Boehm, A.B., 2022.
741 SARS-CoV-2 RNA Wastewater Settled Solids Surveillance Frequency and Impact on
742 Predicted COVID-19 Incidence Using a Distributed Lag Model. ACS ES&T Water.
743 https://doi.org/10.1021/ACSESTWATER.2C00074
744 Seabold, S., Josef Perktold, 2010. Statsmodels: Econometric and statistical modeling with python,
745 in: Proceedings of the 9th Python in Science Conference. pp. 92–96.
746 Sherchan, S.P., Shahin, S., Patel, J., Ward, L.M., Tandukar, S., Uprety, S., Schmitz, B.W., Ahmed,
747 W., Simpson, S., Gyawali, P., 2021. Occurrence of SARS-CoV-2 RNA in Six Municipal
748 Wastewater Treatment Plants at the Early Stage of COVID-19 Pandemic in The United States.
749 Pathog. 2021, Vol. 10, Page 798 10, 798. https://doi.org/10.3390/PATHOGENS10070798
750 Shi, J., Gao, X., Xue, S., Li, F., Nie, Q., Lv, Y., Wang, J., Xu, T., Du, G., Li, G., 2021. Spatio-

751 temporal evolution and influencing mechanism of the COVID-19 epidemic in Shandong
752 province, China. Sci. Reports 2021 111 11, 1–16. https://doi.org/10.1038/s41598-021-86188-
753 0
754 Simpson, A., Topol, A., White, B.J., Wolfe, M.K., Wigginton, K.R., Boehm, A.B., 2021. Effect
755 of storage conditions on SARS- CoV-2 RNA quantification in wastewater solids. PeerJ 9,
756 e11933. https://doi.org/10.7717/PEERJ.11933/SUPP-5
757 Spurbeck, R.R., Minard-Smith, A., Catlin, L., 2021. Feasibility of neighborhood and building scale
758 wastewater-based genomic epidemiology for pathogen surveillance. Sci. Total Environ. 789,
759 147829. https://doi.org/10.1016/J.SCITOTENV.2021.147829
760 Tiwari, A., Phan, N., Tandukar, S., Ashoori, R., Thakali, O., Mousazadesh, M., Dehghani, M.H.,
761 Sherchan, S.P., 2021. Persistence and occurrence of SARS-CoV-2 in water and wastewater
762 environments: a review of the current literature. Environ. Sci. Pollut. Res. 1, 1–11.
763 https://doi.org/10.1007/S11356-021-16919-3/FIGURES/2
764 Turska-Kawa, A., Pilch, I., 2022. Political beliefs and the acceptance of the SARS-CoV-2
765 pandemic restrictions. The case of Poland. PLoS One 17, e0264502.
766 https://doi.org/10.1371/JOURNAL.PONE.0264502
767 Upshaw, T.L., Brown, C., Smith, R., Perri, M., Ziegler, C., Pinto, A.D., 2021. Social determinants
768 of COVID-19 incidence and outcomes: A rapid review. PLoS One 16, e0248336.
769 https://doi.org/10.1371/JOURNAL.PONE.0248336
770 Veldhoen, M., Pedro Simas, J., 2021. Covid-19 will become endemic but with decreased potency
771 over time, scientists believe. BMJ 372, n494. https://doi.org/10.1136/BMJ.N494
772 Veldhoen, M., Simas, J.P., 2021. Endemic SARS-CoV-2 will maintain post-pandemic immunity.
773 Nat. Rev. Immunol. 2021 213 21, 131–132. https://doi.org/10.1038/s41577-020-00493-9

774 Wolfe, M.K., Archana, A., Catoe, D., Coffman, M.M., Dorevich, S., Graham, K.E., Kim, S.,
775 Grijalva, L.M., Roldan-Hernandez, L., Silverman, A.I., Sinnott-Armstrong, N., Vugia, D.J.,
776 Yu, A.T., Zambrana, W., Wigginton, K.R., Boehm, A.B., 2021. Scaling of SARS-CoV-2
777 RNA in Settled Solids from Multiple Wastewater Treatment Plants to Compare Incidence
778 Rates of Laboratory-Confirmed COVID-19 in Their Sewersheds. Environ. Sci. Technol. Lett.
779 8, 398–404. https://doi.org/10.1021/ACS.ESTLETT.1C00184
780 Wu, F., Lee, W.L., Chen, H., Gu, X., Chandra, F., Armas, F., Xiao, A., Leifels, M., Rhode, S.F.,
781 Wuertz, S., Thompson, J., Alm, E.J., 2022a. Making waves: Wastewater surveillance of
782 SARS-CoV-2 in an endemic future. Water Res. 219, 118535.
783 https://doi.org/10.1016/J.WATRES.2022.118535
784 Wu, F., Xiao, A., Zhang, J., Moniz, K., Endo, N., Armas, F., Bonneau, R., Brown, M.A., Bushman,
785 M., Chai, P.R., Duvallet, C., Erickson, T.B., Foppe, K., Ghaeli, N., Gu, X., Hanage, W.P.,
786 Huang, K.H., Lee, W.L., Matus, M., McElroy, K.A., Nagler, J., Rhode, S.F., Santillana, M.,
787 Tucker, J.A., Wuertz, S., Zhao, S., Thompson, J., Alm, E.J., 2022b. SARS-CoV-2 RNA
788 concentrations in wastewater foreshadow dynamics and clinical presentation of new COVID-
789 19 cases. Sci. Total Environ. 805, 150121.
790 https://doi.org/10.1016/J.SCITOTENV.2021.150121
791 Wu, F., Xiao, A., Zhang, J., Moniz, K., Endo, N., Armas, F., Bushman, M., Chai, P.R., Duvallet,
792 C., Erickson, T.B., Foppe, K., Ghaeli, N., Gu, X., Hanage, W.P., Huang, K.H., Lee, W.L.,
793 McElroy, K.A., Rhode, S.F., Matus, M., Wuertz, S., Thompson, J., Alm, E.J., 2021.
794 Wastewater surveillance of SARS-CoV-2 across 40 U.S. states from February to June 2020.
795 Water Res. 202, 117400. https://doi.org/10.1016/J.WATRES.2021.117400
796 Xiao, A., Wu, F., Bushman, M., Zhang, J., Imakaev, M., Chai, P.R., Duvallet, C., Endo, N.,

797 Erickson, T.B., Armas, F., Arnold, B., Chen, H., Chandra, F., Ghaeli, N., Gu, X., Hanage,
798 W.P., Lee, W.L., Matus, M., McElroy, K.A., Moniz, K., Rhode, S.F., Thompson, J., Alm,
799 E.J., 2022. Metrics to relate COVID-19 wastewater data to clinical testing dynamics. Water
800 Res. 212, 118070. https://doi.org/10.1016/J.WATRES.2022.118070
801 Zhan, Q., Babler, K.M., Sharkey, M.E., Amirali, A., Beaver, C.C., Boone, M.M., Comerford, S.,
802 Cooper, D., Cortizas, E.M., Currall, B.B., Foox, J., Grills, G.S., Kobetz, E., Kumar, N., Laine,
803 J., Lamar, W.E., Mantero, A.M.A., Mason, C.E., Reding, B.D., Robertson, M., Roca, M.A.,
804 Ryon, K., Schürer, S.C., Shukla, B.S., Solle, N.S., Stevenson, M., Jr, J.J.T., Thomas, C.,
805 Thomas, T., Vidović, D., Williams, S.L., Yin, X., Solo-Gabriele, H.M., 2022. Relationships
806 between SARS-CoV-2 in Wastewater and COVID-19 Clinical Cases and Hospitalizations,
807 with and without Normalization against Indicators of Human Waste. ACS ES&T Water.
808 https://doi.org/10.1021/ACSESTWATER.2C00045
809