

1 **Title**

2 **Monitoring SARS-CoV-2 in wastewater during New York City's second wave of COVID-**
3 **19: Sewershed-level trends and relationships to publicly available clinical testing data**

4
5 **Authors**

6 Catherine Hoar,^a Francoise Chauvin,^b Alexander Clare,^b Hope McGibbon,^b Esmeraldo Castro,^b
7 Samantha Patinella,^b Dimitrios Katehis,^b John J. Dennehy,^{c,d} Monica Trujillo,^e Davida S.
8 Smyth,^{f,g} Andrea I. Silverman^{a*}

9
10 ^aDepartment of Civil and Urban Engineering, New York University Tandon School of
11 Engineering Brooklyn, NY, USA

12 ^bNew York City Department of Environmental Protection, New York, NY, USA

13 ^cBiology Department, Queens College, The City University of New York, Queens, NY USA

14 ^dBiology Doctoral Program, The Graduate Center, The City University of New York, New York,
15 NY, USA

16 ^eDepartment of Biology, Queensborough Community College, The City University of New
17 York, Bayside, NY, USA

18 ^fDepartment of Natural Sciences and Mathematics, Eugene Lang College of Liberal Arts at The
19 New School, New York, NY, USA

20 ^gpresent affiliation: Department of Life Sciences, Texas A&M University San Antonio, San
21 Antonio, Texas, USA

22 *corresponding author: Andrea Silverman, andrea.silverman@nyu.edu

23

24

25

26 **Abstract**

27

28 New York City's ongoing wastewater monitoring program tracked trends in sewershed-level
29 SARS-CoV-2 loads starting in the fall of 2020, just before the start of the City's second wave of
30 the COVID-19 outbreak. During a five-month study period, from November 8, 2020 to April 11,
31 2021, viral loads in influent wastewater from each of New York City's 14 wastewater treatment
32 plants were measured and compared to new laboratory-confirmed COVID-19 cases for the
33 populations in each corresponding sewershed, estimated from publicly available clinical testing
34 data. We found significant positive correlations between viral loads in wastewater and new
35 COVID-19 cases. The strength of the correlations varied depending on the sewershed, with
36 Spearman's rank correlation coefficients ranging between 0.38 and 0.81 (mean = 0.55). Based on
37 a linear regression analysis of a combined data set for New York City, we found that a 1 log₁₀
38 change in the SARS-CoV-2 viral load in wastewater corresponded to a 0.6 log₁₀ change in the
39 number of new laboratory-confirmed COVID-19 cases/day in a sewershed. An estimated
40 minimum detectable case rate between 2 - 8 cases/day/100,000 people was associated with the
41 method limit of detection in wastewater. This work offers a preliminary assessment of the
42 relationship between wastewater monitoring data and clinical testing data in New York City.
43 While routine monitoring and method optimization continue, information on the development of
44 New York City's ongoing wastewater monitoring program may provide insights for similar
45 wastewater-based epidemiology efforts in the future.

46

47 Introduction

48
49 In March 2020, New York City became an epicenter of the coronavirus disease 2019 (COVID-
50 19) pandemic. In response to this first wave of COVID-19 cases, the New York City Department
51 of Environmental Protection (NYC DEP) - the city agency responsible for wastewater collection
52 and treatment - launched a wastewater monitoring program with the goal of tracking sewershed-
53 level trends in the concentration of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-
54 CoV-2), the virus that causes COVID-19. The program was developed in partnership with
55 researchers at New York University, Queens College, Queensborough Community College, and
56 The New School, with all routine analysis conducted in the NYC DEP's existing microbiology
57 laboratory under the management of the NYC DEP.

58
59 Wastewater-based epidemiology (WBE) programs for COVID-19, including the one in New
60 York City (NYC), were established on the premise that SARS-CoV-2 virions are excreted in the
61 human waste of individuals infected with SARS-CoV-2 and that the resulting concentrations of
62 viral RNA measured in wastewater are indicative of disease incidence or prevalence in the
63 contributing sewershed. Significant associations between SARS-CoV-2 RNA concentrations
64 measured in wastewater and metrics of COVID-19 disease incidence--including case rates--have
65 been shown at scales ranging from single buildings to entire sewersheds.¹⁻³ Early reports from
66 WBE programs suggested promising predictive applications that could help inform COVID-19
67 response measures,^{4,5} sparking widespread interest in SARS-CoV-2 monitoring programs around
68 the world.^{6,7} While the extent to which wastewater data is a leading indicator of trends in
69 COVID-19 incidence ahead of clinical data may vary depending on clinical testing rates,^{8,9} WBE
70 data do offer the advantage of providing information representative of entire populations, free
71 from clinical testing-related biases. In NYC, where communities of color and high-poverty areas
72 were disproportionately impacted by the first wave of the COVID-19 pandemic,¹⁰ testing rates
73 varied spatially, with significant demographic-based disparities.¹¹ In situations where clinical
74 testing does not adequately sample vulnerable populations, WBE may help inform modifications
75 to testing strategies and provide supplemental information regarding COVID-19 trends.
76 Wastewater monitoring is therefore a potential tool to identify new outbreaks of COVID-19 after
77 high clinical testing rates associated with major "waves" of disease incidence have subsided or
78 when resources and technical capacity for extensive clinical testing of individuals are limited.

79
80 These opportunities make WBE an attractive option for many municipalities, including NYC, to
81 confirm findings from clinical testing about population-level COVID-19 dynamics and to
82 monitor for new outbreaks in instances when testing is inadequate. In August 2020, the NYC
83 DEP's SARS-CoV-2 wastewater monitoring program began routine analysis of influent
84 wastewater collected from NYC's 14 wastewater treatment plants (referred to as wastewater
85 resource recovery facilities (WRRF) by the NYC DEP) (SI Table S1), capturing data during the
86 region's second wave of COVID-19 cases, which started in the fall of 2020. The sewershed
87 catchment areas contributing to each of the 14 WRRFs vary markedly in size, serving
88 populations ranging from approximately 120,000 to 1.2 million residents. To assess the
89 relationship between NYC sewershed-level SARS-CoV-2 RNA concentrations and confirmed
90 cases of COVID-19 within each sewershed, wastewater data were compared to publicly available
91 case data provided by the NYC Department of Health and Mental Hygiene (DOHMH). In
92 presenting findings from the NYC DEP, we also aim to provide insights into the development of

93 a sustainable wastewater monitoring program designed for long-term, routine tracking of trends
94 in virus loads for multiple sewersheds serving a large urban population.

95

96 **Methods**

97

98 *Sample collection and processing*

99 24-h flow-weighted composite influent wastewater samples were collected from each of NYC's
100 14 WRRFs twice weekly beginning August 31, 2020. From January 31, 2021 to April 18, 2021
101 sampling was reduced to once weekly. Each composite sample consisted of eight grab samples
102 collected every three hours beginning at 7:00 AM on the sampling date. Samples were
103 transported on ice and stored at 4 °C until processing, which started within twelve hours after the
104 final grab sample was collected. For each sampling date, one of the 14 samples was analyzed in
105 duplicate and the remainder were analyzed as single samples; facilities were selected for
106 duplicate analysis on a rotating basis. A method blank containing Type I deionized water was
107 included with each set of samples to confirm the absence of contamination during sample
108 processing. Detailed descriptions of materials, methods, and data analysis are provided in the SI.
109 In brief, 40-mL aliquots of the 24-h composite samples were first pasteurized (60 °C, 90 min),
110 and then centrifuged (5000 x g, 4 °C, 10 min) to remove solids. The supernatant was filtered
111 (0.22 µm, cellulose acetate) and then subjected to virus concentration using polyethylene glycol
112 (PEG) precipitation (addition of 4.0 g PEG and 0.9 g NaCl followed by overnight incubation at 4
113 °C, and centrifugation at 12,000 x g at 4 °C for 120 min to pellet viruses).¹² The supernatant was
114 discarded and RNA was extracted from the concentrated PEG pellet using the Qiagen QiaAmp
115 Viral RNA Mini Kit with modifications (described in the SI).

116

117 *SARS-CoV-2 quantification by RT-qPCR*

118 A one-step RT-qPCR assay was used to quantify copies of the SARS-CoV-2 nucleocapsid (N)
119 gene, targeting the N1 region (CDC RUO Primers and Probes, Integrated DNA Technologies¹³)
120 in triplicate reactions on a StepOnePlus Real-Time PCR System (Thermo Fisher Scientific).
121 Synthetic SARS-CoV-2 RNA covering > 99.9% of the viral genome (Twist Bioscience Control
122 1, GENBANK ID MT007544.1) served as both a positive control and standard used in a decimal
123 serial dilution for quantification of N1 gene copies.

124

125 The limit of detection (LOD) and limit of quantification (LOQ) for the assay were estimated
126 from replicate standard curves as described by Forootan et al. 2017¹⁴ and found to be 4,500
127 copies/L and 15,000 copies/L, respectively. Note that these LOD and LOQ values as well as
128 calculated sample concentrations are relative to the approximate concentration of the synthetic
129 RNA control reported by the manufacturer, as absolute quantification of the RNA control was
130 not feasible when sample analysis began. Note that quantification of the RNA control through
131 digital PCR is underway. N1 concentrations--including those of the LOD and LOQ--reported in
132 the current version of this work may therefore be updated in future versions to reflect the
133 quantified concentration of the RT-qPCR standard. Nonetheless, while the approach described
134 herein limits direct comparison of N1 concentrations to those found in other studies, it does not
135 alter trends and comparisons across facilities examined within this study. In addition, we elected
136 to use a pooled standard curve to quantify samples on all plates to ameliorate variability in
137 standard preparation by different analysts from plate to plate. A description of the analysis used
138 to motivate this decision is presented in the SI (Figure S1). The absence of contamination during

139 RT-qPCR preparation was confirmed through no template controls included on all RT-qPCR
140 plates. Only samples quantified above the LOQ were included in subsequent analysis. From
141 September 8, 2020 to June 8, 2021, samples were collected from each facility on 72 sampling
142 dates, with samples from only two dates associated with method blanks having N1
143 concentrations above the LOD; samples collected on these two dates were flagged as
144 contaminated and were not included in subsequent analysis.

145
146 An attenuated bovine coronavirus (BCoV) (Calf-Guard® Bovine Rota-Coronavirus Vaccine,
147 Zoetis) was used as a process control.^{15,16} BCoV was inoculated into samples after the
148 pasteurization step (details provided in the SI). A one-step RT-qPCR assay, adapted from
149 previously published assays,¹⁵⁻¹⁷ targeting the transmembrane-protein gene of BCoV was used to
150 qualitatively assess BCoV recovery for each sample using an aliquot of the extracted RNA
151 (primers and probes purchased from Integrated DNA Technologies). Detection of BCoV was
152 used to confirm whether viruses were recovered in samples for which the N1 target was not
153 detected. Additional details regarding the RT-qPCR assays, standard curves, and QA/QC
154 procedures are provided in the SI.

155 *Data analysis*

156
157 The concentration of the N1 RNA target in wastewater (C_{WW}) was determined for each sample in
158 units of N1 gene copies (GC)/L according to Equation 1, where N_r is the number of N1 GC
159 measured by RT-qPCR, $V_{RNA,s}$ is the volume of RNA extracted from each sample (60 μ L),
160 $V_{RNA,r}$ is the volume of template RNA added to the RT-qPCR reaction (5 μ L), and V_s is the
161 volume of wastewater sample analyzed (0.04 L).

$$162 \quad 163 \quad C_{WW} = (N_r \times V_{RNA,s}) / (V_{RNA,r} \times V_s) \quad \text{Equation 1}$$

164
165 The resulting C_{WW} was then normalized by the associated daily influent wastewater flow rate
166 (i.e., the flow rate in the same facility on the same day) to calculate the SARS-CoV-2 viral
167 loading rate (L_{WW}) in units of N1 GC/day (Equation 2). Given that 60% of the NYC sewer
168 system is a combined stormwater-sewer system, flow-based normalization was used to account
169 for differences in per capita water usage and variability in wastewater flow rates caused by non-
170 domestic water inputs (e.g., rain events), which can affect measured virus concentrations. In
171 Equation 2, Q is the daily flow rate at the facility in millions of gallons per day (MGD), and CF
172 is the conversion factor required to convert from liters to million gallons (3.78541×10^6 L/MG).
173 Continuous measurements of flow rate were conducted at each facility using either magnetic
174 flow meters or flow measuring weirs (with uncertainty in measurements of $\sim 5\%$). Average daily
175 flow rates had been measured at each facility prior to the establishment of the SARS-CoV-2
176 monitoring program, and thus required no additional analysis burden, making it a logistically
177 advantageous option for normalization of virus measurements.

$$178 \quad 179 \quad L_{WW} = C_{WW} \times Q \times CF \quad \text{Equation 2}$$

180 *Statistical analyses of relationships between SARS-CoV-2 loads in wastewater and laboratory-* 181 *confirmed COVID-19 cases*

182 Relationships between SARS-CoV-2 wastewater data in each sewershed and laboratory-
183 confirmed COVID-19 cases for the associated sewershed population were evaluated through
184

185 correlation and linear regression analyses. Clinical data were obtained from publicly available
186 data provided by the NYC DOHMH.¹⁸ In particular, the data set “last7days-by-modzcta.csv”,
187 which was posted online daily, was used to obtain daily reports of the cumulative clinical
188 molecular testing results over the previous seven days for each modified ZIP code tabulation area
189 (MODZCTA) in NYC.¹⁸ Specifically, data on the total clinical COVID-19 tests administered and
190 the total number of positive tests (not including individuals who previously tested positive),
191 reported based on date of specimen collection, were obtained. Note that molecular tests included
192 diagnostic PCR tests and did not include antigen or antibody tests. This data set was used to
193 calculate 7-day averages of new COVID-19 cases (i.e., positive molecular tests) per day,
194 organized by the last date in the 7-day range. For example, the 7-day average reported on
195 February 14 represents the daily average of new cases calculated based on the total number of
196 positive molecular tests collected from February 8 to February 14. Data were available starting
197 on November 7, 2020, with data from March 15, 2021 to March 21, 2021 omitted due to
198 technical issues related to data transmission during this period (Figure S.2). While alternative
199 data sets were available with cumulative new COVID-19 case counts prior to November 2020,
200 these data were organized by the date that test results were reported, as opposed to date of
201 specimen collection, and were therefore not recommended by NYC DOHMH for use in
202 calculating the number of daily new COVID-19 cases.¹⁸

203
204 Each of the 177 MODZCTAs were assigned to one of NYC’s 14 sewersheds. Of the 177
205 MODZCTAs, 44 straddled multiple sewershed areas and were assigned to only the sewershed in
206 which it had the greatest overlapping land area. Total new cases in each sewershed each day
207 were calculated by summing new cases in the MODZCTA assigned to that sewershed. The same
208 data set was used to calculate 7-day averages of COVID-19 testing rates (i.e., the number of tests
209 administered divided by the total population) and the percentages of COVID-19 tests that were
210 positive for each sewershed (Figure S.2).

211
212 Spearman correlations between SARS-CoV-2 viral loading rates in wastewater (N1 GC/day) and
213 7-day averages of new daily COVID-19 cases were determined for each individual sewershed for
214 a five-month study period (November 8, 2020 to April 11, 2021). Correlations were also
215 determined for a combined data set that included each data pair (i.e., SARS-CoV-2 viral loading
216 rates and 7-day average of new COVID-19 cases on each date) for all facilities, excluding the
217 Port Richmond and Oakwood Beach WRRFs (see the Results and Discussion section). For the
218 combined data, correlations were also evaluated after removing data pairs associated with
219 potentially inadequate clinical testing rates: data for dates with percentages of positive molecular
220 tests (7-day average) that exceeded 10% in the sewershed were excluded. A general benchmark
221 suggested by the World Health Organization in the Spring of 2020 indicated that clinical testing
222 is less likely to represent all infections in a population when the percentage of positive tests
223 exceeds approximately 10%;^{19,20} we therefore excluded these data in an effort to best
224 approximate the incidence of SARS-CoV-2 infections.

225
226 To assess whether trends in SARS-CoV-2 viral loading rates in wastewater preceded trends in
227 clinical testing data, correlations between the two data sets were also evaluated for each
228 sewershed with the clinical data shifted back in time with lags ranging from 0 to 21 days. For
229 this analysis, additional clinical data from April 12, 2021 to May 2, 2021 was included to
230 maintain a constant number of data pairs for each number of lag days applied.

231
232 Simple linear regressions were performed using \log_{10} -transformed SARS-CoV-2 viral loading
233 rates (N1 GC/day) and \log_{10} -transformed 7-day averages of new COVID-19 cases (new COVID-
234 19 cases/day) for each individual sewershed as well as for the combined data set. The combined
235 data set was assessed with and without the testing rate filter described above. Linear regressions
236 were used to estimate the equivalent number of cases/day/100,000 people associated with the
237 method LOD (C_{LOD}), equal to 4,500 N1 GC/L. This estimate was calculated for each facility
238 using individual, sewershed-specific linear regressions and using the linear regression for the
239 combined data set. First, the LOD was converted to a SARS-CoV-2 viral loading rate in
240 wastewater ($L_{WW,LOD}$) for each sewershed in units of N1 GC/day using Equation 3, where Q_{avg}
241 is the average of daily flow rates at the facility over the study period (Table S.1), in MGD.

$$242$$
$$243 L_{WW,LOD} = C_{LOD} \times Q_{avg} \times CF \quad \text{Equation 3}$$
$$244$$

245 $L_{WW,LOD}$ for each sewershed were then input to the linear regressions determined for each
246 sewershed to estimate the number of new COVID-19 cases/day associated with the SARS-CoV-
247 2 method LOD ($Case_{LOD}$), using Equation 4, where m and b are the slope and y-intercept of the
248 linear regression line, respectively (presented for each sewershed in the Results and Discussion
249 section). An example estimation is illustrated graphically in Figure S.6. Resulting $Case_{LOD}$ values
250 were normalized per 100,000 people using MODZCTA-level population estimates from the
251 NYC DOHMH NYC Coronavirus Disease 2019 (COVID-19) Data.¹⁸

$$252$$
$$253 \log_{10}(Case_{LOD}) = m \times \log_{10}(L_{WW,LOD}) + b \quad \text{Equation 4}$$
$$254$$

255 As described above, quantification of the RT-qPCR standard for the N1 target is underway.
256 Future updates to the N1 standard concentration will change the reported method LOD, in units
257 of N1 GC/L. However, because all sample concentrations will also be adjusted to reflect the
258 updated standard concentration, we anticipate that the resulting relationships between the
259 wastewater data and the clinical data (including the associated $Case_{LOD}$) should remain similar to
260 what is reported herein.

261
262 Statistical analyses were performed using R, and figures were created using GraphPad Prism.^{21,22}

263 264 265 **Results and Discussion**

266 267 *Methodological considerations for SARS-CoV-2 quantification in wastewater*

268 The public health emergency caused by the emergence of COVID-19 required the expedited
269 development of NYC DEP's SARS-CoV-2 wastewater monitoring program. As such, several
270 methodological choices for virus quantification were considered, and the ultimate standard
271 operating procedure (SOP) described herein was developed reflecting NYC DEP's program
272 goals of monitoring trends in SARS-CoV-2 viral loads in wastewater, accounting for equipment
273 availability, existing expertise of personnel, and considerations of material procurement.
274 Selections were also made to minimize analyst-based variability. For example, commercially-
275 available kits for RNA extraction were considered over alternatives that may be more sensitive to
276 analyst skill and consistency. Data analysis and internally-developed QA/QC guidelines were

277 established in line with programmatic goals. Additional methodological considerations, such as
278 the inclusion of a filtration step in sample preparation, are discussed in the SI.

279
280 Long-term routine monitoring to assess virus trends through quantification with RT-qPCR
281 requires reliable comparison of data originating from different RT-qPCR plates prepared by
282 different analysts, which presents several challenges. First, in the absence of a formally
283 quantified standard for the N1 RNA target, this program relied on the use of a synthetic RNA
284 control. An approximate concentration of this RNA control was provided by the manufacturer,
285 but was found to differ between lots purchased at different times. In addition, standard curves for
286 routine RT-qPCR assays were prepared by different analysts on different days, with separate
287 serial dilutions of standards performed for each individual RT-qPCR plate. To account for any
288 resulting variability caused by these aspects of the RT-qPCR quantification method, we
289 quantified the concentration of each RNA control lot relative to the original lot used and applied
290 a pooled standard curve for quantification of all samples (Figure S.1). Challenges associated with
291 RT-qPCR-based quantification using a standard curve highlight the benefits of alternative
292 methods, such as digital PCR for absolute RNA quantification, which eliminates the need for a
293 standard curve and may offer more sensitive detection for environmental samples.²³ Nonetheless,
294 the methodology employed in this work allowed us to compare relative viral loads and
295 confidently assess of trends of SARS-CoV-2 in wastewater over time.

296
297

298 *SARS-CoV-2 viral loads in influent wastewater*

299 SARS-CoV-2 viral loads in NYC's 14 sewersheds between September 8, 2020 and June 8, 2021
300 were determined from quantifiable N1 gene copy (GC) concentrations in influent samples and
301 are presented normalized by sewershed population (Table S.1²⁴) in Figure 1. Maximum
302 population-normalized SARS-CoV-2 viral loads for each facility during this period ranged from
303 1.6×10^8 to 6.8×10^8 N1 GC/day/population, with many of these values occurring around the
304 time when a peak in COVID-19 cases was observed (January 2021). Note that in September of
305 2020, prior to the increase in COVID-19 cases associated with NYC's second wave of the
306 outbreak, N1 concentrations in wastewater remained below the LOQ in several sewersheds.

307
308 Visual inspection of trends in SARS-CoV-2 quantities in wastewater and new laboratory-
309 confirmed COVID-19 cases indicates an association between the wastewater and clinical data.
310 The strength of this association varied across sewersheds, as reflected in results from statistical
311 analysis presented in the next section. Additionally, most sewersheds exhibited peaks for both
312 data sets in January 2021 (Figure 1), with two notable exceptions being Oakwood Beach and
313 Port Richmond, discussed below. Sewersheds with lower incidence rates of COVID-19 (e.g.,
314 Red Hook WRRF) generally had lower per capita SARS-CoV-2 viral loads in wastewater than
315 those with higher incidence rates of COVID-19 (e.g., Hunts Point WRRF).

316
317 SARS-CoV-2 viral loads in the Coney Island WRRF influent in September 2020 and October
318 2020 displayed a high degree of variability, with some measured virus loads that were greater
319 than those in all other sewersheds during that period, despite a consistent processing method
320 applied for all samples and confirmed COVID-19 case rates that were consistently low across
321 NYC (Figure 1). While there were relatively low rates of clinical testing in New York City in
322 September 2020 and COVID-19 clusters emerged in some neighborhoods served by the Coney

323 Island WRRF at that time,²⁵ it is unclear if these factors contributed to the high viral loads
324 measured in some Coney Island WRRF samples. For example, COVID-19 clusters were also
325 identified in other sewersheds at this time, yet did not result in high SARS-CoV-2 loads in
326 influent samples collected from other WRRFs, and it is difficult to determine whether clinical
327 testing was adequate. It should also be noted that given its large geographic resolution,
328 sewershed-level monitoring may not fully capture the effect of disease clusters (such as those
329 identified at high spatiotemporal resolution using clinical data²⁶) that may be relatively small
330 compared to the sewershed or may straddle multiple sewersheds. Though not examined in this
331 work, differences in wastewater quality or sewershed characteristics may also have contributed
332 to the observed variability.

333
334 A smaller extent of variability in measured SARS-CoV-2 viral loads was observed to varying
335 degrees across all facilities and can stem from several sources. Evaluation of duplicate samples
336 analyzed during the study period allowed for an assessment of potential variability due to sample
337 processing and RNA quantification. Relative standard deviations for N1 concentrations of
338 duplicate samples (i.e., the standard deviation of concentrations from duplicate samples, each
339 with triplicate RT-qPCR reactions, as a percent of the average concentration) ranged from 3% to
340 44% (mean = 17%, median = 14%); these values are comparable to those reported elsewhere for
341 measurement of N1 concentrations in influent wastewater.^{16,27} Aside from methodological
342 sources of variability, potential sources of variability or uncertainty include (1) dilution of
343 wastewater from non-domestic water inputs and variations in domestic water use habits, (2)
344 wastewater chemical composition, which may interfere with sample processing or RNA
345 quantification methods, (3) variability in SARS-CoV-2 shedding intensity and duration for
346 infected individuals²⁸⁻³⁰ and (4) the extent and consistency of viral RNA degradation in
347 sewers.^{27,31}

348
349 To account for variability in wastewater flow rates and minimize the effect of (1), viral loads
350 calculated using measured wastewater flow rates (Equation 2) were used for analysis instead of
351 N1 concentrations. Preliminary tests with an RT-qPCR inhibition control assay during method
352 optimization were used to assess the impact of factor (2) and indicated minimal inhibition (data
353 not shown). Regular assessment of inhibition with additional control assays was not feasible
354 during routine monitoring due to resource constraints. In addition, dilution of RNA, a strategy
355 used to reduce PCR inhibition, was avoided in order to maintain consistency in sample
356 processing, given that viral concentrations in samples collected during periods of low COVID-19
357 case rates were susceptible to dilution below the limits of quantification or detection. While not
358 included in this work, assessment of viral recovery and wastewater matrix effects should be
359 considered for future research aiming to characterize uncertainty in WBE data. Although beyond
360 the scope of this work, identifying and characterizing external factors related to (3) and (4) is the
361 focus of ongoing SARS-CoV-2 WBE research efforts. Considering these uncertainties and
362 variabilities in wastewater data, which likely increase with scale,³² we did not attempt to quantify
363 the number of SARS-CoV-2 infections in each sewershed based on wastewater data, but instead
364 explored the relationship between viral quantities in wastewater and publicly available clinical
365 data to assess trends and associations, and examine differences between sewersheds.

366
367 As mentioned above, SARS-CoV-2 viral loads in wastewater from the Port Richmond and
368 Oakwood Beach WRRFs (both located in the borough of Staten Island) did not capture the peak

369 in COVID-19 cases that was observed in January 2021 across all sewersheds. In the Port
370 Richmond and Oakwood Beach sewersheds there was a marked increase in COVID-19 cases in
371 December 2020 that was accompanied by an associated peak in the SARS-CoV-2 viral load in
372 wastewater during this time. However, as new COVID-19 cases in Staten Island increased by
373 60% in January 2021, the virus loads in wastewater stayed constant or decreased. Compared to
374 sewersheds in the other boroughs, those in Staten Island had relatively high clinical test
375 positivity in December and January (7-14%), despite having an average testing rate (i.e., number
376 of clinical tests administered per capita) for the study period that was greater than that of over
377 half of the other sewersheds (Figure S.2). This observation suggests that testing may not have
378 adequately captured all infections in Staten Island during this period. While inadequate clinical
379 testing rates could potentially reduce the accuracy of the observed relationships between clinical
380 and wastewater data for these sewersheds, it does not explain the lower-than-expected SARS-
381 CoV-2 viral loads measured in Staten Island wastewater in January 2020. A more likely
382 explanation could stem from the composition or operation of the wastewater system in the
383 borough. For example, a portion of the Staten Island population is not served by the sewer
384 system and instead uses septic systems. As such, a segment of this population does not contribute
385 to the sewer system, and viruses excreted by these residents would not have been present in the
386 influent wastewater at the Oakwood Beach and Port Richmond WRRFs. Nonetheless, given that
387 the population served by septic systems on Staten Island is thought to be smaller than those
388 served by the sewer system, it is unlikely that this hypothesis can entirely explain the
389 discrepancy between measured SARS-CoV-2 viral load and new COVID-19 cases. In addition,
390 much of Staten Island uses separated rather than combined stormwater-sewer systems, which
391 could potentially impact the wastewater matrix and influence viral recovery during concentration
392 and quantification steps in sample analysis. Because of these discrepancies, the Staten Island
393 sewersheds were excluded from analysis of the combined data set and the estimation of
394 minimum COVID-19 case rates associated with the LOD.

395
396 By early June 2021, city-wide weekly averages of the percentage of positive COVID-19 clinical
397 tests declined below 1%, and over 50% of NYC residents had received at least one dose of a
398 COVID-19 vaccine.^{18,33} To minimize the potential impact of mass vaccination on the evaluation
399 of relationships between case rates and SARS-CoV-2 concentrations in wastewater presented in
400 this work, we chose to conduct the statistical analyses described in the following section for a
401 period ending in early April, shortly after New York State extended vaccination availability to
402 individuals of 16 years and older.

403

404 *Relationships between SARS-CoV-2 viral loads in wastewater and new laboratory-confirmed* 405 *COVID-19 cases*

406 Significant positive correlations between SARS-CoV-2 viral loads in wastewater and new
407 laboratory-confirmed COVID-19 cases in the corresponding populations were found for all
408 individual sewersheds and for the combined data set (Spearman, $p < 0.05$), indicating, as
409 expected, that an increase in COVID-19 cases was associated with an increase in SARS-CoV-2
410 concentrations in wastewater (Figure 2). Correlation coefficients (ρ) for the individual
411 sewersheds ranged from 0.38 (Coney Island WRRF) to 0.81 (Wards Island WRRF), with an
412 average of 0.55. Similar correlation coefficients between SARS-CoV-2 wastewater
413 concentrations and clinical case data have been reported elsewhere.^{16,34} Note that analysis of
414 correlations between virus concentrations (N1 GC/L, as opposed to virus loads) and new

415 COVID-19 case rates (cases/day/100,000, as opposed to cases/day) yielded similar results (Table
416 S.3). The correlation coefficient for the combined data set ($\rho = 0.82$) was higher than for any of
417 the individual sewersheds (Figure 3.a).

418
419 Minimal differences were observed in the magnitudes of the Spearman's rank correlation
420 coefficients between clinical COVID-19 case data and SARS-CoV-2 viral loads in wastewater
421 for the data sets with and without lag times applied (Figure S.4). Furthermore, correlations for
422 several sewersheds--including the Wards Island WRRF--were strongest without a time lag
423 between the two data sets. Previous studies, applying a variety of assessment methods, have
424 suggested lag times between clinical testing and wastewater data ranging on the order of days to
425 weeks, while others have indicated that the SARS-CoV-2 concentration in wastewater is not a
426 leading indicator of COVID-19 diagnosis.⁹ Inconsistent findings for lag times may be attributed
427 to whether clinical data are presented by the date of specimen collection or the date that results
428 are reported, as well as the adequacy of COVID-19 testing rates, which vary in different regions
429 and shift across time. Clinical data collected during periods with low testing rates are less likely
430 to capture all infections in a region, and individuals may be more likely to be tested after
431 symptom onset, at a time when viral shedding in feces may have already begun. These
432 conditions can result in a lag behind wastewater monitoring data, which provides viral load
433 information independent from clinical testing rates. Data for this work was collected during a
434 time when testing rates were significantly higher than those during the first wave of the
435 pandemic in NYC, and weekly median turnaround times for test results were 1 to 2 days.¹⁸
436 Furthermore, we could not confidently rule out that the small improvements in correlations
437 observed when applying a lag time for some sewersheds was an artifact of variability in the
438 measured wastewater data. A rigorous assessment of lag time would also need to account for
439 contributions of previous as well as newly infected individuals to viral loads in wastewater,
440 which was beyond the scope of this work. For these reasons, we considered data without a time
441 lag for subsequent comparisons and linear regression analysis.

442
443 Because the nonparametric Spearman's rank correlation was used for this analysis, results
444 suggest that there is, at minimum, a monotonic, direct relationship between SARS-CoV-2
445 quantified in wastewater and clinically confirmed COVID-19 cases. Linear relationships
446 between the two \log_{10} -transformed datasets were assessed through analysis of linear regressions,
447 with the best fit found for the Wards Island WRRF ($R^2 = 0.65$) and some of the poorest fits found
448 for the sewersheds in Staten Island (Figure 2). Inconsistent relationships between sewershed-
449 level SARS-CoV-2 viral loads in wastewater and COVID-19 cases observed across sewersheds
450 may be due to differences in the sewer systems for each sewershed, including sewershed areas,
451 residence times of wastewater in the sewer system, the presence of non-domestic wastewater
452 inputs, proportions of the population made up by transient individuals or commuters, and per
453 capita water use. Differences could also be related to clinical testing rates for each sewershed,
454 though no significant correlation was found between the slopes of the linear regression lines and
455 the average testing rates for the study period for each sewershed (Spearman, $p > 0.05$). Similarly,
456 no significant correlations were found between the slopes of the linear regression lines and (1)
457 average wastewater flow rate, (2) sewershed population, or (3) average per capita wastewater
458 flow rate (Spearman, $p > 0.05$), which was expected given that N1 concentrations were
459 normalized by flow rate. Nonetheless, the linear regression found using the combined data set
460 had a strong fit ($R^2 = 0.70$) relative to the fits of regressions for the individual sewersheds.

461
462 Understanding the utility of SARS-CoV-2 wastewater monitoring data has largely involved
463 comparison of viral concentrations in wastewater to COVID-19 case counts based on clinical
464 testing.³⁵ Because the accuracy of confirmed case rates as a measure of the number of infected
465 individuals is dependent on COVID-19 testing rates, this comparison must be made with a
466 consideration of clinical testing biases. Moreover, if multiple clinical data types are available,
467 one must determine which is most appropriate for comparison to wastewater data. The analysis
468 applied herein utilized a data set containing 7-day averages of new COVID-19 cases based on
469 testing in each approximated sewershed area. Uncertainties surrounding such clinical testing data
470 include (1) whether there were regional biases in testing results (Figure S.2), potentially due to
471 testing disparities;¹¹ (2) whether testing rates were adequate and what constitutes adequate
472 testing; and (3) how long before specimen collection infected individuals contracted COVID-19
473 and started shedding the virus. Others have reported correlations of wastewater data with
474 COVID-19 surveillance data sets other than clinical case rates, such as clinical test positivity or
475 hospitalization rates.² Hospital admissions data, although not without its own biases,³⁶ may be an
476 alternative epidemiological metric to compare to or to validate wastewater monitoring data if
477 significant inadequacies in clinical testing are suspected. While hospitalization data at the
478 MODZCTA level were not publicly available for NYC, visual comparison at the borough level
479 indicates that trends in daily hospitalizations generally reflect trends in case rates for sewersheds
480 within each borough (Figure S.3). The limitations of clinical testing are in fact a major driver for
481 the application of WBE, which aims to provide community-level information free from clinical
482 testing bias.³⁷⁻³⁹ Continued population-level monitoring from wastewater data could become
483 increasingly useful in areas where clinical testing rates decline or resources for clinical testing
484 are limited.

485
486 Linear regressions for the combined data set are presented in Figure 3 with data collected on
487 dates with over 10% positive COVID-19 testing rates removed. Removing data associated with
488 potentially inadequate testing from the combined data set did not significantly change the
489 regression (Analysis of Covariance, $p > 0.05$) compared to the full data set without filtering
490 (Figure S.5). After the peak in COVID-19 cases in NYC in January 2021, there was a decline in
491 cases across all sewersheds. To assess whether the relationship between SARS-CoV-2 loads in
492 wastewater and new clinical COVID-19 cases was significantly different during the period of
493 declining cases from that during the period when cases were increasing, we compared separate
494 linear regressions for the data associated with the rise in case rates (data prior to January 2021)
495 and the decline in case rates (data after January 2021). No significant differences were found
496 between the slopes of the linear regression lines determined using the full combined data set and
497 the data separated based on time period.

498
499 The slope of the linear regression line for the full combined data set was found to be 0.6,
500 indicating that a 1 \log_{10} change in the number of N1 GC/day corresponded to a 0.6 \log_{10} change
501 in the number of new laboratory-confirmed COVID-19 cases/day in a sewershed. Metrics such
502 as these are derived from relative changes in viral load, and therefore do not require absolute
503 quantification of viral concentrations in wastewater, allowing for comparison to other studies and
504 alleviating challenges related to absolute quantification of standard curves. However, this metric
505 comparing SARS-CoV-2 loads and daily new COVID-19 cases has not been consistently
506 reported in studies monitoring SARS-CoV-2 in influent wastewater. Harmonizing data analysis

507 strategies to include such a metric would improve efforts to compare results across different
508 locations. The slope of 0.6 observed herein is greater than that reported previously by Wolfe et
509 al. (slope = 0.24), who compared SARS-CoV-2 concentrations measured in primary wastewater
510 settled solids and COVID-19 incidence in seven publicly owned treatment works located across
511 the United States, including one of the NYC facilities described in this work.³⁵ In addition to
512 analyzing a different type of sample for SARS-CoV-2 concentrations (i.e., primary settled solids
513 versus influent wastewater), the analysis used by Wolfe et al. (2021) differed from that herein in
514 that they normalized measured SARS-CoV-2 concentrations in wastewater solids by
515 concentrations of pepper mild mottle virus (PMMoV). The differences in the slopes may be due
516 to either of these factors, to variations in the relationship between SARS-CoV-2 wastewater
517 loads and COVID-19 cases in different regions, or to a difference in the overall sensitivity of the
518 methodology applied by Wolfe et al.

519
520 At present, limitations regarding the accuracy of COVID-19 clinical testing data and
521 uncertainties related to SARS-CoV-2 measurements in wastewater--including SARS-CoV-2
522 shedding rates and RNA stability in different sewersheds--preclude development and validation
523 of a universal, quantitative model to predict disease incidence based on viral RNA concentrations
524 in wastewater. Ongoing research continues to expand our understanding of critical model
525 parameters and factors contributing to uncertainty, owing particularly to SARS-CoV-2
526 monitoring work completed at smaller scales (e.g., building-level),⁴⁰ from which information
527 about the contributing population can be obtained more easily than from larger sewersheds. An
528 attempt to quantify COVID-19 case rates in NYC's sewersheds based on wastewater data at this
529 time would be inaccurate, and is not currently recommended for application in the realm of
530 public health.⁴¹ However, based on our analysis and others, there is utility in using wastewater
531 data to monitor trends in COVID-19 incidence.

532
533 *Estimated case rates associated with method LOD*

534 The utility of SARS-CoV-2 wastewater data depends on whether virions are present in
535 wastewater at detectable concentrations (i.e., above the LOD and LOQ). It is therefore useful to
536 approximate the minimum number of contributing COVID-19 cases per day required for
537 detection of the SARS-CoV-2 N1 gene target in wastewater using the methodology described
538 here. When estimated using individual, sewershed-specific linear regressions (Figure 2), the
539 minimum new COVID-19 case rate that corresponds to the method LOD varied for each
540 sewershed, ranging between 2 and 8 cases/day/100,000 people (Table S.4). Minimum detectable
541 case rates were also estimated for each sewershed using the linear regression from the combined
542 data set and the average daily influent flow rates for each WRRF during the study period. These
543 estimates fell within the same range as those derived from sewershed-specific linear regressions
544 (Table S.4).

545
546 The minimum detectable case rate estimates presented here should be taken as order-of-
547 magnitude approximations rather than absolute quantities, especially considering the varying
548 strength of the linear relationships between data for certain sewersheds (e.g., data sets for Coney
549 Island, Bowery Bay, Oakwood Beach, and Port Richmond WRRFs had Pearson correlation
550 coefficients below 0.5). Furthermore, these findings hold only for the specific SARS-CoV-2
551 quantification methodology applied herein, and may not be transferable to locations with
552 different per capita wastewater flow rates, even if testing rates and case rates are similar to those

553 described here. The estimates may also be limited by the assumption that the dominant source of
554 the SARS-CoV-2 viral load in the wastewater is from recent cases as opposed to prolonged fecal
555 shedding, which is consistent with assumptions made in previous studies.^{35,42} Furthermore,
556 variability in virus shedding rates were not considered for the simple linear models in our study.
557 The relationships found are also limited by the accuracy of clinical testing data, as discussed
558 above.

559
560 As COVID-19 cases declined in NYC in the spring and early summer of 2021, the estimated
561 minimum detectable COVID-19 case rates were reached in most sewersheds by May and June
562 2021. As such, we expected that SARS-CoV-2 viral loads in wastewater would have decreased
563 to below the LOQ and LOD at this time. However, viral RNA was still detectable in influent
564 wastewater collected from all sewersheds in mid June 2021 (Figure 4). While this discrepancy
565 may be explained by the limitations described above, it may also be due to decreasing COVID-
566 19 testing rates, which could result in reduced diagnosis of individuals with asymptomatic
567 infections, who are less likely to seek out COVID-19 tests. The average COVID-19 testing rate
568 in NYC during the period from May 2, 2021 to June 8, 2021 decreased 30% from the average in
569 January 2021. Additionally, widespread vaccination of adults in New York may have resulted in
570 asymptomatic and mild infections that were not diagnosed. While individuals with asymptomatic
571 SARS-CoV-2 infections may not be captured by clinical testing, viral shedding by asymptomatic
572 individuals would still contribute to the viral load in wastewater, given that SARS-CoV-2 has
573 been detected in fecal samples associated with asymptomatic or mild cases of COVID-19.⁴³⁻⁴⁵
574 Viral loads may have also been elevated in wastewater because of prolonged fecal shedding of
575 the virus. Finally, it is possible that the linear relationship found in this work does not hold at low
576 SARS-CoV-2 infection levels as the study period used for statistical analysis included only case
577 rates above the minimum detectable case rates estimated for each sewershed.

578
579
580 The estimated minimum numbers of COVID-19 cases required before SARS-CoV-2 can be
581 detected in wastewater from NYC sewersheds are associated with considerable disease incidence
582 that may be captured if some degree of clinical testing continues. Nonetheless, these estimates
583 could aid public health agencies in understanding what COVID-19 incidence to expect if SARS-
584 CoV-2 loads measured in wastewater influent cross the threshold from being below the detection
585 limit to being detected. Improvements to analytical methods that lower the LOD⁴⁶⁻⁴⁸ would
586 expand the utility of WBE in indicating low levels of disease incidence.

587 588 **Conclusion**

589
590 Critical choices made at the beginning of the development of NYC's SARS-CoV-2 wastewater
591 monitoring program proved beneficial for the long-term wastewater monitoring goals for NYC,
592 and highlight strategies that may be useful for agencies interested in implementing wastewater
593 monitoring programs for emerging pathogens. First, collaborating parties--including academic
594 partners and NYC DEP personnel--worked together to develop a monitoring program centered
595 around NYC DEP's priorities. Second, sample analysis was conducted in a NYC DEP
596 microbiology laboratory, which allowed the program to take advantage of existing equipment,
597 expertise, protocols, and resources related to wastewater analysis, as well as existing wastewater
598 sampling and transport protocols and infrastructure. Doing so expedited the initiation of the

599 wastewater monitoring program and supported virus analysis capacity building within the NYC
600 DEP. With this structure, routine monitoring began in parallel with training and continued
601 method optimization. Consequently, protocol adjustments responded to practical challenges as
602 well as technical ones, taking into account laboratory infrastructure and equipment that would
603 ultimately be used for the ongoing monitoring program. This also made for a rich training
604 experience, in which analysts shared insights from hands-on experience, contributed to workflow
605 decisions, and were exposed to the empirical reasoning behind methodological choices. Direct
606 communication between wastewater treatment facility operators and laboratory personnel
607 maximized use of the NYC DEP's extensive knowledge base and data, which aided in
608 troubleshooting.

609
610 As WBE programs for wastewater-related viruses evolve to meet future challenges, continued
611 research is needed to better understand the mechanisms by which virus concentration, extraction,
612 and quantification methods work, and the factors that influence the efficiency of each step; this
613 knowledge can subsequently inform method optimization, standardization, and the accounting of
614 methodological uncertainty. Since the implementation of the SARS-CoV-2 wastewater
615 monitoring program in NYC, several studies have begun to evaluate and compare different
616 sample processing strategies, including one interlaboratory study which included the
617 methodology used herein.⁴⁸⁻⁵⁰ A clear characterization of the limitations and benefits of
618 methodological choices for virus enumeration is critical for not only assessing previously
619 collected data but also comparing results between WBE programs implemented by different
620 parties, and informing future efforts in the WBE field. For example, varied priorities, resources,
621 and expertise in different WBE programs may foster the continued use of many different
622 methods rather than the adoption of one universal method. Additionally, poorly characterized
623 variability in WBE data stands in the way of the critical goal of relating viral loads in wastewater
624 to disease dynamics. Clear characterization of uncertainties related to analytical methodologies
625 would therefore facilitate interpretation of wastewater data by public health agencies.⁵¹
626 Nonetheless, results from NYC's monitoring program show that relative trends in SARS-CoV-2
627 loads in wastewater can be evaluated and associated with trends in clinical testing data, and
628 therefore can potentially contribute to situational awareness of disease incidence in large urban
629 sewersheds.
630

631 **Conflicts of Interest**

632 There are no conflicts of interest to declare.

633

634 **Acknowledgements**

635

636 Funding for this work was provided by the New York City Department of Environmental
637 Protection and the Alfred P. Sloan Foundation.

638

639 An extensive team at the NYC DEP made this monitoring program possible, including Samantha
640 MacBride, Peter Williamsen, Gina Behnke, Jasmin Torres, and Jorge Villacis; members of the
641 NYC DEP Microbiology Lab, including William Kelly, Naudet Joasil, Patrick Hoyes, Donovan
642 Johnson, Manzura Kopusov, Oren Sachs, and Samantha Cruickshank; the NYC DEP
643 transportation team, including Lateef Franklin, Samuel Young, and John Congemi; and Abeba
644 Negatu, Patrick Jagessar, Max Verastegui and their process control laboratory teams at NYC
645 DEP.

646

647 Several researchers at CUNY provided support and assistance for protocol development,
648 optimization, and training, including Sherin Kannoly, Kaung Myat “Zach” San, Kristen Cheung,
649 Anna Gao, Michelle Markman, Nanami Kubota, and Irene Hoxie.

650

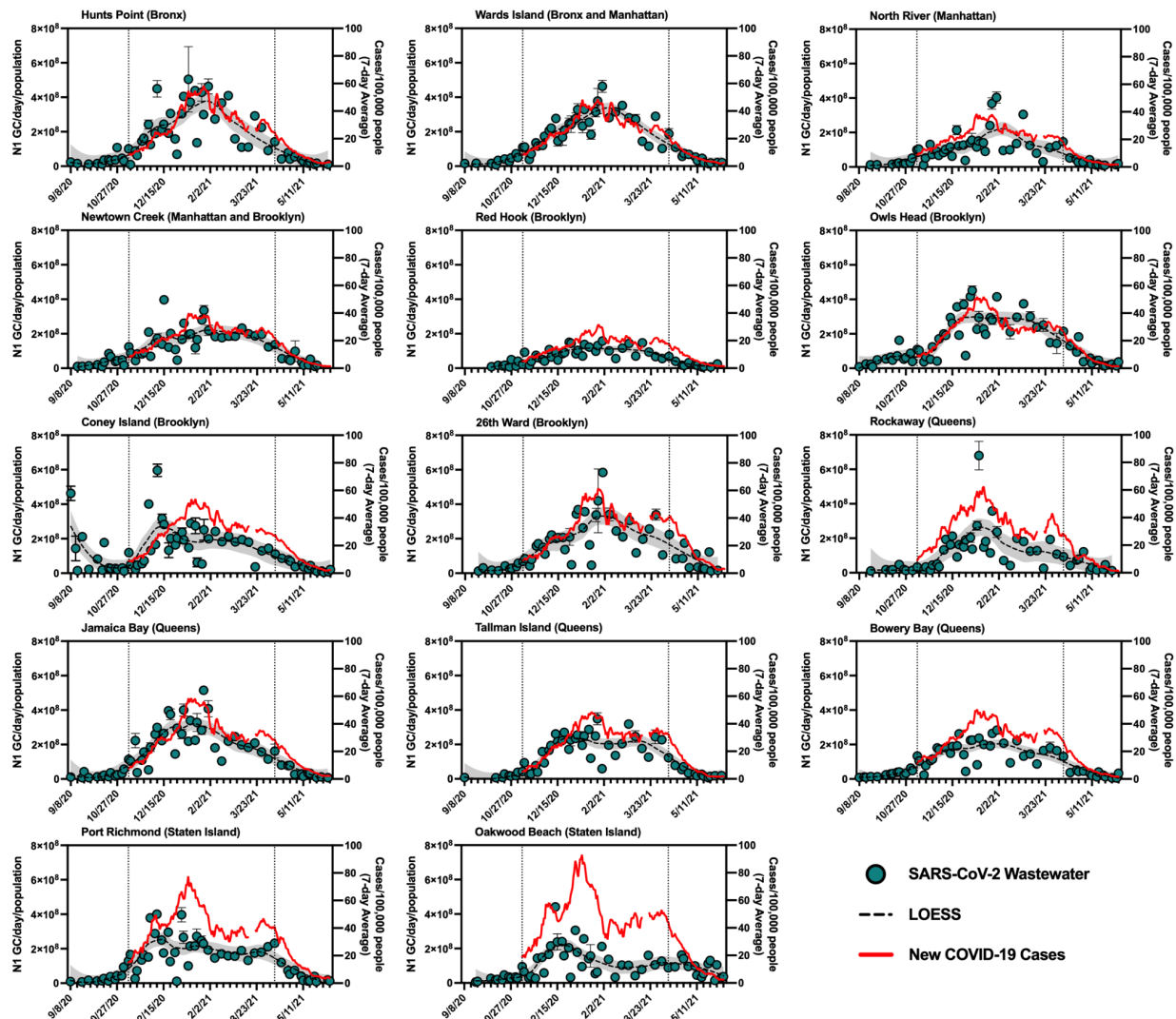
651 We thank Alexandria Boehm (Stanford University) and Sandra Mclellan (University of
652 Wisconsin-Milwaukee) for their support and guidance during program development. We also
653 acknowledge the many insights gained from the interactions through the NSF Research
654 Coordination Network (RCN) on Wastewater Surveillance for SARS-CoV-2.

655

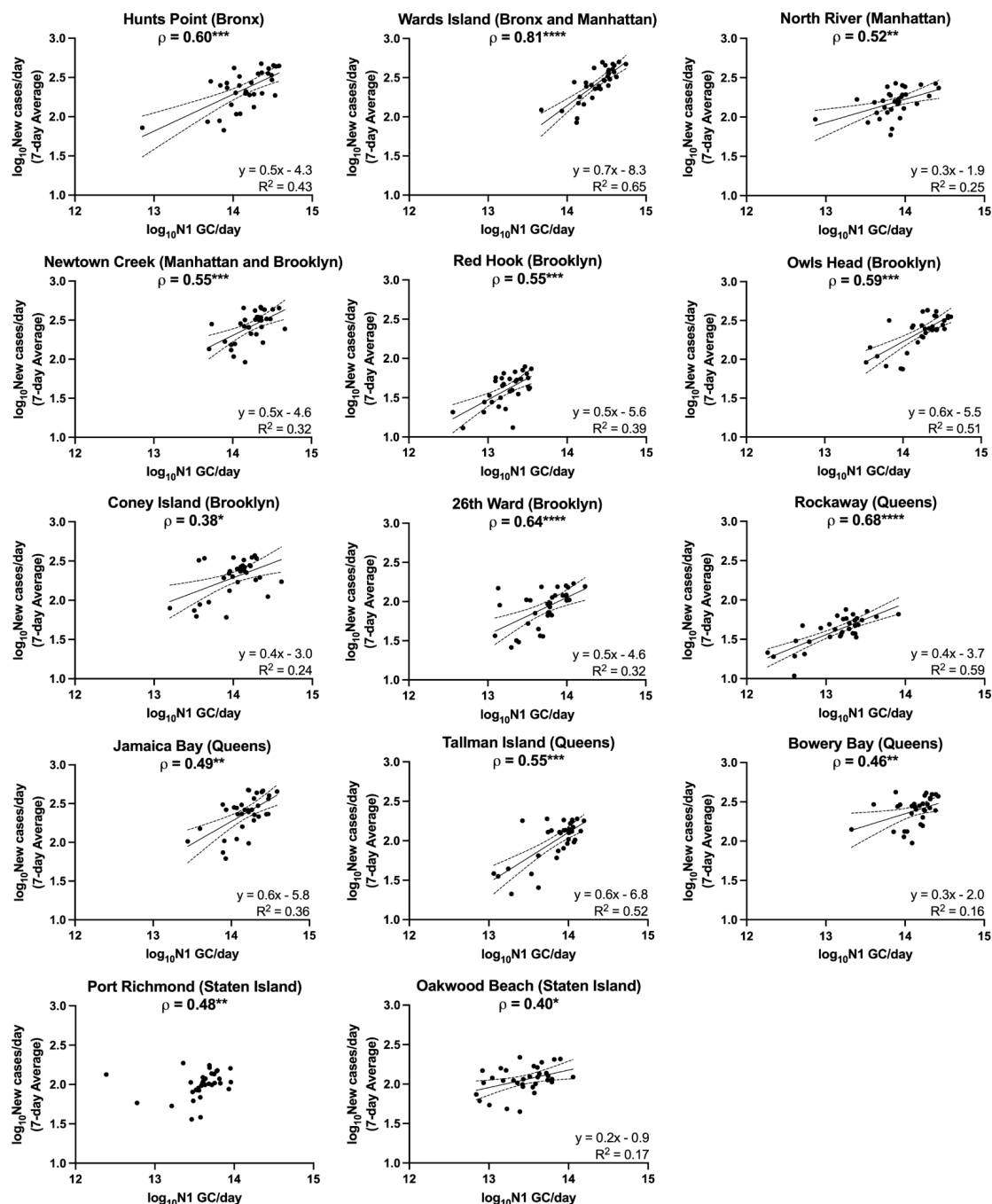
656 A script automating the download of New York City’s publicly available COVID-19 clinical
657 testing data was generously provided by Charlie Mydlarz (NYU Center for Urban Science and
658 Progress).

659 **Figures**

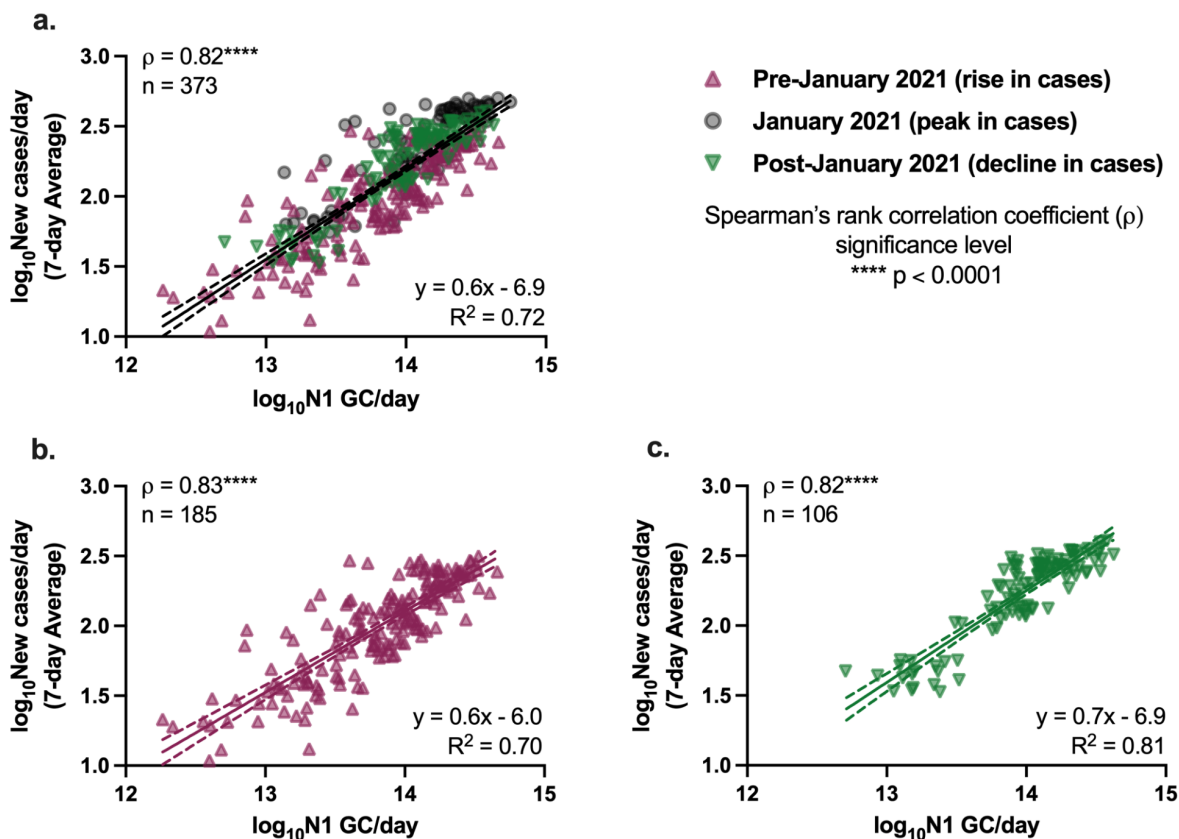
660 Note that the N1 concentrations reported in the following figures may be updated in future
 661 versions of this work to reflect the quantified concentration of the RT-qPCR standard, which is
 662 currently being quantified. These updates should not change observed trends reported here, as
 663 described in the main text.



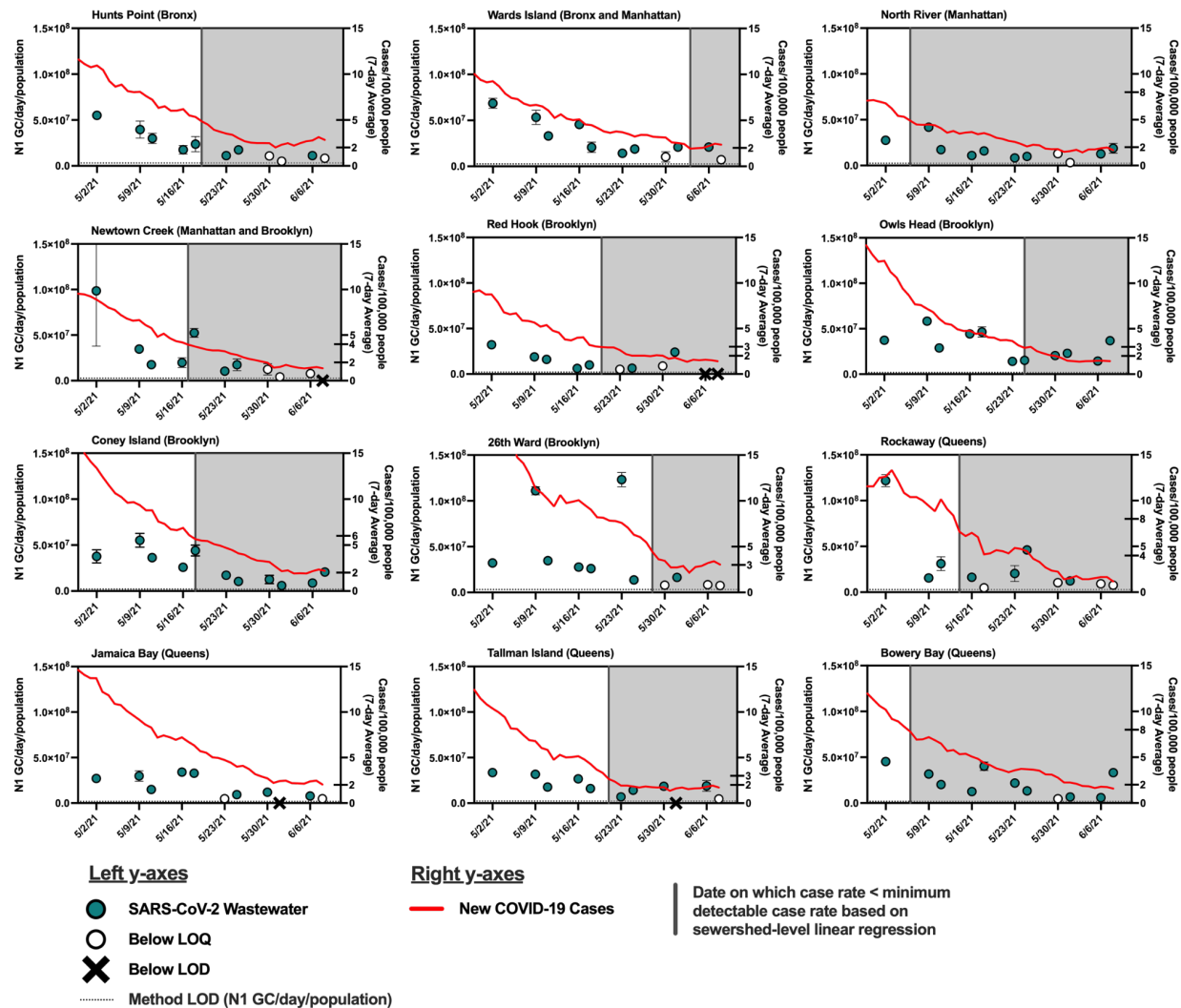
664 **Figure 1. Summary of SARS-CoV-2 wastewater data for New York City’s 14 sewersheds.**
 665 Data from September 8, 2020 to June 8, 2021 is shown, with the period for which statistical
 666 analysis was conducted (November 8, 2020 to April 11, 2021) bounded by vertical dotted lines.
 667 **Primary (left) y-axis, blue circles:** Influent SARS-CoV-2 viral loads normalized by sewershed
 668 populations. Error bars indicate standard deviations from triplicate RT-qPCR reactions as well as
 669 standard deviations of duplicate samples, where applicable. Dashed black lines represent LOESS
 670 curve fits (span = 0.4), with the 95% confidence intervals shaded in grey. **Secondary (right) y-**
 671 **axis, red line:** 7-day average of new COVID-19 cases/day/100,000 people in the previous 7 days
 672 normalized using MODZCTA-level population estimates from the NYC DOHMH’s NYC
 673 Coronavirus Disease 2019 (COVID-19) Data.¹⁸ Normalization by population was used for visual
 674 comparison across different sewersheds only and was not used for statistical analysis.
 675
 676



677
 678 **Figure 2. Linear regressions of \log_{10} -transformed SARS-CoV-2 viral loads in wastewater**
 679 **(N1 GC/day) and \log_{10} -transformed 7-day averages of new COVID-19 cases/day for each**
 680 **sewershed in New York City. Linear regressions (solid lines) and associated 95% confidence**
 681 **intervals (dashed lines) are shown along with goodness of fit R^2 values for those data sets with**
 682 **significantly non-zero slopes. Note that linear regression for Port Richmond has been excluded**
 683 **as the slope was not significantly non-zero (see SI). The Spearman's rank correlation coefficient**
 684 **(ρ) between N1 GC/day and new COVID-19 cases/day is shown at the top of each sewershed**
 685 **plot, with significance levels indicated (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p <$**
 686 **0.0001).**



687
688 **Figure 3. Linear regressions of \log_{10} -transformed flow-normalized SARS-CoV-2 viral loads**
689 **in wastewater (N1 GC/day) and \log_{10} -transformed 7-day averages of new COVID-19**
690 **cases/day for (a) the combined data set, (b) data from the combined data set associated with**
691 **a rise in cases, and (c) data from the combined data set associated with a decline in cases.**
692 Data associated with potentially inadequate testing (i.e., over 10% positive tests) are not included
693 in this analysis. Linear regressions (solid lines) and associated 95% confidence intervals (dashed
694 lines) are shown along with goodness of fit R^2 values and Spearman's rank correlation
695 coefficients (ρ) between N1 GC/day and new COVID-19 cases/day.
696



697
 698 **Figure 4. SARS-CoV-2 wastewater data and COVID-19 case data from May 2, 2021 to**
 699 **June 8, 2021.** The date on which the case rate first fell below the estimated minimum detectable
 700 case rate (based on the sewershed-level linear regression) is indicated with a solid vertical line
 701 for each sewershed. Shaded regions indicate the time period during which case rates were below
 702 the estimated minimum detectable case rate. **Primary (left) y-axis, blue circles:** Influent SARS-
 703 CoV-2 viral loads normalized by sewershed populations. Error bars indicate standard deviations
 704 from triplicate RT-qPCR reactions as well as standard deviations of duplicate samples, where
 705 applicable. Open circles represent N1 concentrations below the limit of quantification (LOQ).
 706 Samples below the limit of detection (LOD, shown with a horizontal dotted line) are denoted
 707 with an “X.” **Secondary (right) y-axis, red line:** 7-day average of new COVID-19
 708 cases/day/100,000 people in the previous 7 days. Estimated minimum detectable case rates (new
 709 cases/day/100,000) needed to detect SARS-CoV-2 in wastewater, based on linear regressions
 710 derived from sewershed-level data and the combined data set, are indicated with tick marks
 711 across the y-axes.

712
 713
 714

715 **References**

- 716 1 G. Medema, F. Been, L. Heijnen and S. Petterson, Implementation of environmental
717 surveillance for SARS-CoV-2 virus to support public health decisions: Opportunities and
718 challenges, *Curr Opin Environ Sci Health*, 2020, **17**, 49–71.
- 719 2 J. Peccia, A. Zulli, D. E. Brackney, N. D. Grubaugh, E. H. Kaplan, A. Casanovas-Massana, A.
720 I. Ko, A. A. Malik, D. Wang, M. Wang, J. L. Warren, D. M. Weinberger, W. Arnold and S. B.
721 Omer, Measurement of SARS-CoV-2 RNA in wastewater tracks community infection
722 dynamics, *Nature Biotechnology*, 2020, **38**, 1164–1167.
- 723 3 L. C. Scott, A. Aubee, L. Babahaji, K. Vigil, S. Tims and T. G. Aw, Targeted wastewater
724 surveillance of SARS-CoV-2 on a University Campus for COVID-19 outbreak detection and
725 mitigation, *Environmental Research*, 2021, 111374.
- 726 4 F. Wu, A. Xiao, J. Zhang, K. Moniz, N. Endo, F. Armas, R. Bonneau, M. A. Brown, M.
727 Bushman, P. R. Chai, C. Duvallet, T. B. Erickson, K. Foppe, N. Ghaeli, X. Gu, W. P. Hanage,
728 K. H. Huang, W. L. Lee, M. Matus, K. A. McElroy, J. Nagler, S. F. Rhode, M. Santillana, J.
729 A. Tucker, S. Wuertz, S. Zhao, J. Thompson and E. J. Alm, SARS-CoV-2 titers in wastewater
730 foreshadow dynamics and clinical presentation of new COVID-19 cases, *medRxiv*, 2020,
731 DOI:10.1101/2020.06.15.20117747.
- 732 5 J. Peccia, A. Zulli, D. E. Brackney, N. D. Grubaugh, E. H. Kaplan, A. Casanovas-Massana, A.
733 I. Ko, A. A. Malik, D. Wang, M. Wang, J. L. Warren, D. M. Weinberger and S. B. Omer,
734 SARS-CoV-2 RNA concentrations in primary municipal sewage sludge as a leading indicator
735 of COVID-19 outbreak dynamics, *medRxiv*, 2020, 2020.05.19.20105999.
- 736 6 F. Kreier, The myriad ways sewage surveillance is helping fight COVID around the world,
737 *Nature*, 2021, DOI:10.1038/d41586-021-01234-1.
- 738 7 World Health Organization, *Status of environmental surveillance for SARS-CoV-2 virus:*
739 *Scientific Brief*, 2020.
- 740 8 X. Fernandez-Cassi, A. Scheidegger, C. Bänziger, F. Cariti, A. Tuñas Corzon, P.
741 Ganesanandamoorthy, J. C. Lemaitre, C. Ort, T. R. Julian and T. Kohn, Wastewater
742 monitoring outperforms case numbers as a tool to track COVID-19 incidence dynamics when
743 test positivity rates are high, *Water Research*, 2021, **200**, 117252.
- 744 9 S. W. Olesen, M. Imakaev and C. Duvallet, Making waves: Defining the lead time of
745 wastewater-based epidemiology for COVID-19, *Water Res*, 2021, **202**, 117433.
- 746 10 C. N. Thompson, COVID-19 Outbreak — New York City, February 29–June 1, 2020, *MMWR*
747 *Morb Mortal Wkly Rep*, 2020, DOI:10.15585/mmwr.mm6946a2.
- 748 11 W. Lieberman-Cribbin, S. Tuminello, R. M. Flores and E. Taioli, Disparities in COVID-19
749 Testing and Positivity in New York City, *American Journal of Preventive Medicine*, 2020, **59**,
750 326–332.
- 751 12 M. Trujillo, K. Cheung, A. Gao, I. Hoxie, S. Kannoly, N. Kubota, K. M. San, D. S. Smyth and
752 J. J. Dennehy, Protocol for safe, affordable, and reproducible isolation and quantitation of
753 SARS-CoV-2 RNA from wastewater, *PLOS ONE*, 2021, **16**, e0257454.
- 754 13 X. Lu, L. Wang, S. K. Sakthivel, B. Whitaker, J. Murray, S. Kamili, B. Lynch, L. Malapati, S.
755 A. Burke, J. Harcourt, A. Tamin, N. J. Thornburg, J. M. Villanueva and S. Lindstrom, US
756 CDC Real-Time Reverse Transcription PCR Panel for Detection of Severe Acute Respiratory
757 Syndrome Coronavirus 2, *Emerg Infect Dis*, 2020, **26**, 1654–1665.
- 758 14 A. Forootan, R. Sjöback, J. Björkman, B. Sjögreen, L. Linz and M. Kubista, Methods to
759 determine limit of detection and limit of quantification in quantitative real-time PCR (qPCR),
760 *Biomol Detect Quantif*, 2017, **12**, 1–6.

- 761 15 S. Loeb, One-Step RT-ddPCR for Detection of SARS-CoV-2, Bovine Coronavirus, and
762 PMMoV RNA in RNA Derived from Wastewater or Primary Settled Solids, *protocols.io*,
763 2020, DOI:10.17504/protocols.io.bi6vkhe6.
- 764 16 S. Feng, A. Roguet, J. S. McClary-Gutierrez, R. J. Newton, N. Kloczko, J. G. Meiman and S.
765 L. McLellan, Evaluation of Sampling, Analysis, and Normalization Methods for SARS-CoV-
766 2 Concentrations in Wastewater to Assess COVID-19 Burdens in Wisconsin Communities,
767 *ACS EST Water*, 2021, **1**, 1955–1965.
- 768 17 N. Decaro, G. Elia, M. Campolo, C. Desario, V. Mari, A. Radogna, M. L. Colaianni, F.
769 Cirone, M. Tempesta and C. Buonavoglia, Detection of bovine coronavirus using a TaqMan-
770 based real-time RT-PCR assay, *J Virol Methods*, 2008, **151**, 167–171.
- 771 18 nychealth/coronavirus-data, <https://github.com/nychealth/coronavirus-data>, (accessed 20 May
772 2021).
- 773 19 World Health Organization, COVID-19 - virtual press conference - March 30, 2020,
774 [https://www.who.int/docs/default-source/coronaviruse/transcripts/who-audio-emergencies-](https://www.who.int/docs/default-source/coronaviruse/transcripts/who-audio-emergencies-coronavirus-press-conference-full-30mar2020.pdf?sfvrsn=6b68bc4a_2)
775 [coronavirus-press-conference-full-30mar2020.pdf?sfvrsn=6b68bc4a_2](https://www.who.int/docs/default-source/coronaviruse/transcripts/who-audio-emergencies-coronavirus-press-conference-full-30mar2020.pdf?sfvrsn=6b68bc4a_2).
- 776 20 A. Aubrey, Which States Are Doing Enough Testing? This Benchmark Helps Settle The
777 Debate, *npr*, 2020 [https://www.npr.org/sections/health-shots/2020/04/22/840526338/is-the-u-](https://www.npr.org/sections/health-shots/2020/04/22/840526338/is-the-u-s-testing-enough-for-covid-19-as-debate-rages-on-heres-how-to-know)
778 [s-testing-enough-for-covid-19-as-debate-rages-on-heres-how-to-know](https://www.npr.org/sections/health-shots/2020/04/22/840526338/is-the-u-s-testing-enough-for-covid-19-as-debate-rages-on-heres-how-to-know), (accessed 25 May
779 2021).
- 780 21 R Core Team, *R: A language and environment for statistical computing*. R Foundation for
781 Statistical Computing, Vienna, Austria, 2019.
- 782 22 *GraphPad Prism version 9.1.1 for macOS*, GraphPad Software, La Jolla California USA,
783 www.graphpad.com.
- 784 23 R. Gonzalez, A. Larson, H. Thompson, E. Carter and X. F. Cassi, Redesigning SARS-CoV-2
785 clinical RT-qPCR assays for wastewater RT-ddPCR, *medRxiv*, 2021, 2021.03.02.21252754.
- 786 24 2050 SED Forecasts, [https://www.nymtc.org/DATA-AND-MODELING/SED-](https://www.nymtc.org/DATA-AND-MODELING/SED-Forecasts/2050-Forecasts)
787 [Forecasts/2050-Forecasts](https://www.nymtc.org/DATA-AND-MODELING/SED-Forecasts/2050-Forecasts), (accessed 12 April 2021).
- 788 25 New York City Department of Health and Mental Hygiene, Press Notice About COVID-19
789 Areas of Concern: Tuesday, September 22, 2020,
790 <https://www1.nyc.gov/assets/doh/downloads/pdf/covid/dear-reporter-letter-09222020.pdf>.
- 791 26 S. K. Greene, E. R. Peterson, D. Balan, L. Jones, G. M. Culp, A. D. Fine and M. Kulldorff,
792 Detecting COVID-19 Clusters at High Spatiotemporal Resolution, New York City, New
793 York, USA, June–July 2020 - Volume 27, Number 5—May 2021 - Emerging Infectious
794 Diseases journal - CDC, 2021, DOI:10.3201/eid2705.203583.
- 795 27 X. Li, S. Zhang, J. Shi, S. P. Luby and G. Jiang, Uncertainties in estimating SARS-CoV-2
796 prevalence by wastewater-based epidemiology, *Chemical Engineering Journal*, 2021, **415**,
797 129039.
- 798 28 D. L. Jones, M. Q. Baluja, D. W. Graham, A. Corbishley, J. E. McDonald, S. K. Malham, L.
799 S. Hillary, T. R. Connor, W. H. Gaze, I. B. Moura, M. H. Wilcox and K. Farkas, Shedding of
800 SARS-CoV-2 in feces and urine and its potential role in person-to-person transmission and the
801 environment-based spread of COVID-19, *Science of The Total Environment*, 2020, **749**,
802 141364.
- 803 29 S. Mallett, A. J. Allen, S. Graziadio, S. A. Taylor, N. S. Sakai, K. Green, J. Suklan, C. Hyde,
804 B. Shinkins, Z. Zhelev, J. Peters, P. J. Turner, N. W. Roberts, L. F. di Ruffano, R. Wolff, P.
805 Whiting, A. Winter, G. Bhatnagar, B. D. Nicholson and S. Halligan, At what times during
806 infection is SARS-CoV-2 detectable and no longer detectable using RT-PCR-based tests? A

- 807 systematic review of individual participant data, *BMC Medicine*, 2020, **18**, 346.
- 808 30 M. Cevik, M. Tate, O. Lloyd, A. E. Maraolo, J. Schafers and A. Ho, SARS-CoV-2, SARS-
809 CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a
810 systematic review and meta-analysis, *The Lancet Microbe*, 2021, **2**, e13–e22.
- 811 31 A. Bivins, J. Greaves, R. Fischer, K. C. Yinda, W. Ahmed, M. Kitajima, V. J. Munster and K.
812 Bibby, Persistence of SARS-CoV-2 in Water and Wastewater, *Environ. Sci. Technol. Lett.*,
813 2020, **7**, 937–942.
- 814 32 D. A. Larsen and K. R. Wigginton, Tracking COVID-19 with wastewater, *Nature*
815 *Biotechnology*, 2020, **38**, 1151–1153.
- 816 33 *COVID-19 Vaccination Reporting*, NYC Department of Health and Mental Hygiene, 2021.
- 817 34 J. Weidhaas, Z. T. Aanderud, D. K. Roper, J. VanDerslice, E. B. Gaddis, J. Ostermiller, K.
818 Hoffman, R. Jamal, P. Heck, Y. Zhang, K. Torgersen, J. V. Laan and N. LaCross, Correlation
819 of SARS-CoV-2 RNA in wastewater with COVID-19 disease burden in sewersheds, *Sci Total*
820 *Environ*, 2021, **775**, 145790.
- 821 35 M. K. Wolfe, A. Archana, D. Catoe, M. M. Coffman, S. Dorevich, K. E. Graham, S. Kim, L.
822 M. Grijalva, L. Roldan-Hernandez, A. I. Silverman, N. Sinnott-Armstrong, D. J. Vugia, A. T.
823 Yu, W. Zambrana, K. R. Wigginton and A. B. Boehm, Scaling of SARS-CoV-2 RNA in
824 Settled Solids from Multiple Wastewater Treatment Plants to Compare Incidence Rates of
825 Laboratory-Confirmed COVID-19 in Their Sewersheds, *Environ. Sci. Technol. Lett.*, 2021,
826 DOI:10.1021/acs.estlett.1c00184.
- 827 36 K. Sherratt, S. Abbott, S. R. Meakin, J. Hellewell, J. D. Munday, N. Bosse, M. Jit and S.
828 Funk, Exploring surveillance data biases when estimating the reproduction number: with
829 insights into subpopulation transmission of COVID-19 in England, *Philos Trans R Soc Lond*
830 *B Biol Sci*, 2021, DOI:10.1098/rstb.2020.0283.
- 831 37 M. Murakami, A. Hata, R. Honda and T. Watanabe, Letter to the Editor: Wastewater-Based
832 Epidemiology Can Overcome Representativeness and Stigma Issues Related to COVID-19,
833 *Environ. Sci. Technol.*, 2020, **54**, 5311–5311.
- 834 38 A. Zahedi, P. Monis, D. Deere and U. Ryan, Wastewater-based epidemiology—surveillance
835 and early detection of waterborne pathogens with a focus on SARS-CoV-2, *Cryptosporidium*
836 and *Giardia*, *Parasitol Res*, 2021, DOI:10.1007/s00436-020-07023-5.
- 837 39 N. Sims and B. Kasprzyk-Hordern, Future perspectives of wastewater-based epidemiology:
838 Monitoring infectious disease spread and resistance to the community level, *Environ Int*,
839 2020, **139**, 105689.
- 840 40 B. W. Schmitz, G. K. Innes, S. M. Prasek, W. Q. Betancourt, E. R. Stark, A. R. Foster, A. G.
841 Abraham, C. P. Gerba and I. L. Pepper, Enumerating asymptomatic COVID-19 cases and
842 estimating SARS-CoV-2 fecal shedding rates via wastewater-based epidemiology, *Science of*
843 *The Total Environment*, 2021, **801**, 149794.
- 844 41 CDC, National Wastewater Surveillance System, <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/wastewater-surveillance.html>, (accessed 1 June 2021).
- 845
846 42 D. Gerrity, K. Papp, M. Stoker, A. Sims and W. Frehner, Early-pandemic wastewater
847 surveillance of SARS-CoV-2 in Southern Nevada: Methodology, occurrence, and
848 incidence/prevalence considerations, *Water Research X*, 2021, **10**, 100086.
- 849 43 S. Park, C.-W. Lee, D.-I. Park, H.-Y. Woo, H. S. Cheong, H. C. Shin, K. Ahn, M.-J. Kwon
850 and E.-J. Joo, Detection of SARS-CoV-2 in Fecal Samples From Patients With Asymptomatic
851 and Mild COVID-19 in Korea, *Clinical Gastroenterology and Hepatology*, 2021, **19**, 1387-
852 1394.e2.

- 853 44 A. Mesoraca, K. Margiotti, A. Viola, A. Cima, D. Sparacino and C. Giorlandino, Evaluation
854 of SARS-CoV-2 viral RNA in fecal samples, *Virology*, 2020, **17**, 86.
- 855 45 X. Jiang, M. Luo, Z. Zou, X. Wang, C. Chen and J. Qiu, Asymptomatic SARS-CoV-2
856 infected case with viral detection positive in stool but negative in nasopharyngeal samples
857 lasts for 42 days, *Journal of Medical Virology*, 2020, **92**, 1807–1809.
- 858 46 W. Ahmed, P. M. Bertsch, A. Bivins, K. Bibby, K. Farkas, A. Gathercole, E. Haramoto, P.
859 Gyawali, A. Korajkic, B. R. McMinn, J. F. Mueller, S. L. Simpson, W. J. M. Smith, E. M.
860 Symonds, K. V. Thomas, R. Verhagen and M. Kitajima, Comparison of virus concentration
861 methods for the RT-qPCR-based recovery of murine hepatitis virus, a surrogate for SARS-
862 CoV-2 from untreated wastewater, *Science of The Total Environment*, 2020, **739**, 139960.
- 863 47 S. E. Philo, E. K. Keim, R. Swanstrom, A. Q. W. Ong, E. A. Burnor, A. L. Kossik, J. C.
864 Harrison, B. A. Demeke, N. A. Zhou, N. K. Beck, J. H. Shirai and J. S. Meschke, A
865 comparison of SARS-CoV-2 wastewater concentration methods for environmental
866 surveillance, *Sci Total Environ*, 2021, **760**, 144215.
- 867 48 A. Pérez-Cataluña, E. Cuevas-Ferrando, W. Randazzo, I. Falcó, A. Allende and G. Sánchez,
868 Comparing analytical methods to detect SARS-CoV-2 in wastewater, *Science of The Total
869 Environment*, 2021, **758**, 143870.
- 870 49 B. M. Pecson, E. Darby, C. N. Haas, Y. M. Amha, M. Bartolo, R. Danielson, Y. Dearborn, G.
871 D. Giovanni, C. Ferguson, S. Fevig, E. Gaddis, D. Gray, G. Lukasik, B. Mull, L. Olivas, A.
872 Olivieri, Y. Qu and S.-C.-2 I. Consortium, Reproducibility and sensitivity of 36 methods to
873 quantify the SARS-CoV-2 genetic signal in raw wastewater: findings from an interlaboratory
874 methods evaluation in the U.S., *Environmental Science: Water Research & Technology*, 2021,
875 **7**, 504–520.
- 876 50 N. Alygizakis, A. N. Markou, N. I. Rousis, A. Galani, M. Avgeris, P. G. Adamopoulos, A.
877 Scorilas, E. S. Lianidou, D. Paraskevis, S. Tsiodras, A. Tsakris, M.-A. Dimopoulos and N. S.
878 Thomaidis, Analytical methodologies for the detection of SARS-CoV-2 in wastewater:
879 Protocols and future perspectives, *TrAC Trends in Analytical Chemistry*, 2021, **134**, 116125.
- 880 51 J. S. McClary-Gutierrez, M. C. Mattioli, P. Marcenac, A. I. Silverman, A. B. Boehm, K.
881 Bibby, M. Balliet, F. L. de los Reyes, D. Gerrity, J. F. Griffith, P. A. Holden, D. Katehis, G.
882 Kester, N. LaCross, E. K. Lipp, J. Meiman, R. T. Noble, D. Brossard and S. L. McLellan,
883 SARS-CoV-2 Wastewater Surveillance for Public Health Action - Volume 27, Number 9—
884 September 2021 - Emerging Infectious Diseases journal - CDC, 2021,
885 DOI:10.3201/eid2709.210753.
886