

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43

The role of wastewater-based epidemiology for SARS-CoV-2 in developing countries: cumulative evidence from South Africa supports sentinel site surveillance to guide public health decision-making

Chinwe Iwu-Jaja^{1*}, Nkosenhle Lindo Ndlovu^{1*}, Said Rachida¹, Mukhlid Yousif^{1,2}, Setshaba Taukobong¹, Mokgaetji Macheke¹, Laurette Mhlanga³, Cari van Schalkwyk³, Juliet Pulliam³, Tom Moultrie⁴, Wouter le Roux⁵, Lisa Schaefer⁶, Gina Pockock⁷, Leanne Coetzee⁷, Janet Mans⁸, Faizal Bux⁹, Leanne Pillay⁹, Dariah de Villiers¹⁰, AP du Toit¹⁰, Don Jambo¹¹, Annancietar Gomba¹¹, Shaun Groenink¹², Neil Madgewick¹³, Martie van der Walt¹⁴, Awelani Mutshembele¹⁴, Natascha Berkowitz¹⁵, Melinda Suchard¹, Kerrigan McCarthy^{1,2,5} for the SACCESS network

1. Centre for Vaccines and Immunology, National Institute for Communicable Diseases, South Africa
2. Department of Virology, School of Pathology, University of the Witwatersrand, Johannesburg, South Africa
3. South African Centre for Epidemiological Modelling and Analysis, Stellenbosch University, Cape Town, South Africa
4. Centre for Actuarial Research, University of Cape Town, South Africa
5. School of Public Health, University of the Witwatersrand, Johannesburg, South Africa
6. Water Centre, Council for Scientific and Industrial Research, Pretoria, South Africa
7. Waterlab (Pty) Ltd, Pretoria, South Africa
8. University of Pretoria, Department of Medical Virology, Pretoria, South Africa
9. Institute for Water and Wastewater Technology, Durban University of Technology, Durban, South Africa
10. Lumegen Laboratories (Pty) Ltd, Potchefstroom, South Africa
11. National Institute for Occupational Health, South Africa
12. Greenhill Laboratories, South Africa
13. Praecautio, South Africa
14. South African Medical Research Council –Tuberculosis Platform
15. City of Cape Town, South Africa

*joint first authors

Corresponding authors: Chinwe Iwu-Jaja chinwei@nicd.ac.za; Nkosenhle Lindo Ndlovu nkosenhleh@nicd.ac.za

44

45 **Abstract**

46 Background: The World Health Organisation recommends wastewater based epidemiology (WBE) for SARS-
47 CoV-2 as a complementary tool for monitoring population-level epidemiological features of the COVID-19
48 pandemic. Yet, uptake of WBE in low-to-middle income countries (LMIC) is low. We report on findings from
49 SARS-CoV-2 WBE surveillance network in South Africa, and make recommendations regarding implementation
50 of WBE in LMICs

51 Methods: Seven laboratories using different test methodology, quantified influent wastewater collected from
52 87 wastewater treatment plants (WWTPs) located in all nine South African provinces for SARS-CoV-2 from 01
53 June 2021 – 31 May 2022 inclusive, during the 3rd and 4th waves of COVID-19. Regression analysis with
54 district laboratory-confirmed SARS-CoV-2 case loads, controlling for district, size of plant and testing frequency
55 was determined. The sensitivity and specificity of ‘rules’ based on WBE data to predict an epidemic wave
56 based on SARS-CoV-2 wastewater levels were determined.

57 Results: Among 2158 wastewater samples, 543/648 (85%) samples taken during a wave tested positive for
58 SARS-CoV-2 compared with 842 positive tests from 1512 (55%) samples taken during the interwave period.
59 Overall, the regression-co-efficient was 0,66 (95% confidence interval=0,6-0,72, $R^2=0.59$), but ranged from
60 0.14-0.87 by testing laboratory. Early warning of the 4th wave of SARS-CoV-2 in Gauteng Province in
61 November-December 2021 was demonstrated. A 50% increase in log-copies SARS-CoV-2 compared with a
62 rolling mean over the previous 5 weeks was the most sensitive predictive rule (58%) to predict a new wave.

63 Conclusion: Variation in the strength of correlation across testing laboratories, and redundancy of findings
64 across co-located testing plants, suggests that test methodology should be standardised and that surveillance
65 networks may utilise a sentinel site model without compromising the value of WBE findings for public health
66 decision-making. Further research is needed to identify optimal test frequency and the need for normalisation
67 to population size, so as to identify predictive and interpretive rules to support early warning and public health
68 action. Our findings support investment in WBE for SARS-CoV-2 surveillance in low and middle-income
69 countries.

70

71 **Keywords:** SARS-CoV-2, COVID-19, wastewater based epidemiology, Surveillance, South Africa

72

73 **Background**

74 The World Health Organisation (WHO) recommends wastewater based epidemiology (WBE) for SARS-CoV-2 as
75 a complementary tool for monitoring population-level epidemiological features of the COVID-19 pandemic
76 including the presence or absence of infection, changing trends over time and variant monitoring [1].

77 Wastewater based epidemiology provides certain advantages over case-based surveillance [2]. These include
78 the ability to monitor large numbers of persons by testing single samples and the potential to estimate the
79 burden of disease in that community from quantitative data[3], [4, 5]. WBE eliminates inherent biases in
80 disease burden estimation arising through differential health seeking by infected persons, variable access to
81 clinical laboratory testing and variation in health care practitioner's propensity to test [1]. WBE may provide
82 early warning of increases in case-incidence [6, 7] enabling prompt public health action to be taken to prevent
83 outbreaks and their consequences [5]. Furthermore, WBE is inherently cost-saving relative to patient-based
84 surveillance systems. Most likely, as a consequence of these advantages, many countries are implementing
85 WBE for SARS-CoV-2. Only one year into the pandemic, a scoping review of the literature showed that 26
86 countries had already reported the presence of SARS CoV-2 in wastewater including countries in Europe,
87 Americas, Asia, Oceania, and Africa [8], and by January 2023, 72 countries regularly publish results of SARS-
88 CoV-2 testing in wastewater on 164 dashboards across the world [9]. It is clear that WBE for SARS-CoV-2
89 contributes to the epidemiological knowledge base that supports individual and public health decision making
90 regarding SARS-CoV-2 prevention, detection and response.

91 Despite these benefits, and recommendations from developing countries researchers, few African countries
92 have adopted WBE surveillance. Of 164 SARS-CoV-2 WBE dashboards, only two are found in Africa (both in
93 South Africa) [9]. In developing countries, structural challenges such as low number of wastewater treatment
94 facilities, low service levels or non-functionality of plants and absent or limited testing resources stand in the
95 way of WBE implementation [10]. Challenges shared with developed countries such as uncertainty regarding
96 how to integrate WBE findings into existing surveillance programmes [11] and how to interpret WBE results
97 with regard to actionable thresholds [12], have most likely contributed to reluctance to invest in WBE
98 surveillance programmes.

99 Against this backdrop, the South African Collaborative COVID-19 Surveillance System (SACCESS) network was
100 established in 2021, with the aim of developing standard test methodology, identifying and addressing
101 challenges in qualitative, quantitative detection and RNA sequencing of SARS-CoV-2 in wastewater’[13]. Sites
102 for testing were identified on pragmatic and logistical considerations, and included testing of influent
103 wastewater from almost all treatment plants in densely populated urban metropolitan areas, and single plants
104 in central cities or towns in rural provinces. The SACCESS network actively monitored SARSCoV-2 in
105 wastewater across the country from the third wave of the pandemic in South Africa, where four distinct waves
106 of COVID-19 pandemic have occurred, peaking in June 2020, December 2020, July 2021, and December 2021
107 respectively [14].

108 In this paper, we describe the detection and quantification of SARS-CoV-2 from wastewater treatment plants
109 (WWTP) across larger South African metropolitan areas and in single plants from rural provinces from June
110 2021 to May 2022 (spanning the third and fourth COVID-19 waves). We further explored the relationship
111 between SARS-CoV-2 levels, and district laboratory-confirmed COVID-19 case load, while adjusting for
112 laboratory test method, geographical location, WWTP plant size, testing frequency and timing of testing during
113 or between SARS-CoV-2 waves. We determined the ability of SARS-CoV-2 levels in wastewater to predict a new
114 wave of COVID-19. We reflect on lessons learned to strengthen implementation and design of WBE
115 surveillance networks in middle and low-income countries.

116 **Methods**

117 *Development of the SACCESS network and extent of surveillance*

118 Commencing in 2018, the NICD tests wastewater for poliovirus as part of the National Department of Health’s
119 polio surveillance programme in accordance with WHO guidance [15]. In 2020, the NICD, and a number of
120 other research, academic and private laboratories commenced with independent projects involving qualitative
121 and quantitative detection of SARS-CoV-2 in local WWTPs. A proof of concept was published by the SAMRC in
122 2021 [16].

123 In August 2020, laboratories doing this work collectively established the SACCESS network to support
124 communication of challenges and sharing best practice[13]. Through funding and partnership with the Water
125 Research Commission (WRC), the NICD and partners identified WWTPs for sampling and SARS-CoV-2 testing to

126 maximise coverage across metros and sentinel sites in provinces with smaller populations. A number of
127 webinars were held to share methodologies for concentration, extraction and PCR detection of SARS-CoV-2
128 [17, 18]. Testing of WWTPs commenced between June 2020 and March 2021 as funding became available.
129 Overall, 87 wastewater treatment sites across nine provinces in SA were tested by the SACCESS network in
130 2021 (Table S1). This included 40 treatment sites in Gauteng, 5 in Eastern Cape, 5 in Western Cape, 9 in Free
131 State, 12 in KwaZulu-Natal, 2 in Limpopo, 3 in Mpumalanga, 2 in Northern Cape province, and 3 in Northwest
132 province. WWTPs were assigned to specific laboratories, and these laboratories consistently tested the same
133 plants over time.

134 *Laboratory test methods*

135 Each of seven participating laboratories collected samples, concentrated and extracted nucleic acid, and
136 amplified and quantified SARS-CoV-2 RNA using the general principles followed by the NICD methodology,
137 which is fully described here, but using their own RNA extraction and PCR detection methods listed in Table 1.
138 This approach was chosen on account of the number of laboratories participating in the network, the number
139 of plants we wished to test, the different SARS-CoV-2 test kits already used by laboratories, and stock
140 shortages during COVID. The NICD sample collection and processing is as follows: Influent grab samples of one
141 litre volume were collected at treatment facilities during morning peak hours and maintained below 5°C
142 during transportation to the testing facility. No data on flow rate nor volume were collected. At the NICD, 200
143 mL of sewage sample was centrifuged at 4650 x g (4°C) for 30 minutes. Without disturbing the pellet, 70mL of
144 the supernatant was centrifuged at 3500g for 15 min through the Centricon® Plus-70 centrifugal ultra-filter.
145 The 70mL supernatant was concentrated into a final volume of approximately 1mL. A sewage concentrate
146 volume of 140uL was spiked with 5µL of the commercial internal control (IC) from the Allplex™ 2019-nCoV
147 (Seegene) to monitor RNA inhibitory effects during RNA extraction and RT-qPCR processes. Spiked
148 concentrates plus IC underwent RNA extraction using the QIAmp® viral RNA mini kit (Qiagen) according to the
149 manufacturer's instructions. Extracted RNA was eluted into 50µL of AVE buffer and tested immediately or
150 stored at -20°C.

151 To prepare for quantification the RNA EDX standard® (Bio-Rad) was double extracted (150 µL and 200 µL) using
152 the QIAmp® viral RNA mini kit (Qiagen) according to manufacturer's instructions. These extracts were filtered
153 through 1 spin column and eluted into 50µL of AVE buffer. Serial dilutions of the standard were performed at

154 1:4096, 1:1024, 1:256, 1:64, 1:16, and 1:4. Absolute genome copies for each dilution were calculated as: 2.34,
155 9.38, 37.50, 150, 600, and 2400 respectively. Four 8.5 µL triplicate sets of the standard RNA were prepared for
156 each dilution and stored at -20 °C for simultaneous use with extracted samples.

157 Finally, viral RNA was detected and quantified in samples using Taqman-based RT-PCR on the 7500 real-time
158 PCR system (Applied Bio systems) using the commercial Allplex™ 2019-nCoV (Seegene) with EDX standard RNA
159 included in each testing run. The PCR reaction mix consisted of 5 µL 2019-nCoV MOM, 5 µL RNase-free H₂O, 5
160 µL 5X Real-time One-step Buffer, 2 µL Real-time One-step Enzyme. Cycling conditions were as follows: preheat
161 at 50 °C for 20 min, denature at 95 °C for 15 min, 45 cycles of amplification with 95 °C for 15s and 58 °C for
162 30s. To increase confidence samples were tested in triplicate, and mean values determined. No participating
163 laboratory normalised for population size.

164 *Inter-laboratory comparisons and quality assurance*

165 Viral concentrates of 10 samples were randomly selected from the database of each participating laboratory.
166 The concentrates were transported at -20°C to the NICD laboratory where nucleic acid was extracted and
167 SARS-CoV-2 detection was carried out in triplicate. All laboratories achieved the target concordance of 80%
168 with NICD results. In quantification quality assessment panels comprising spiked and unspiked raw wastewater
169 samples and concentrates prepared by the NICD, four of seven laboratories achieved 80% extraction efficiency

170 *Statistical analysis*

171 We determined the relationship between SARS-CoV-2 case load and SARS-CoV-2 levels in wastewater through
172 correlation and regression analyses as follows. We used the NICD anonymised line list of laboratory-confirmed
173 SARS-CoV-2 cases geo-located to district or sub-district of South Africa. We allocated cumulative case-counts
174 by epi-week and district (or sub-district where available) to corresponding log-transformed weekly SARS-CoV-2
175 levels (in genome copies per millilitre) from samples obtained from WWTP situated in that district (or sub-
176 district). Where quantification results were below the level of quantification of the test methodology used by
177 each laboratory, we omitted these values from correlation and regression analyses. We conducted correlation
178 and regression analyses for results tested by different laboratories together and independently (Table 2).
179 Regarding correlation, we used Spearman's correlation test for linear correlations due to the non-normality of
180 data. We conducted simple regression using the case loads of laboratory-confirmed SARS-CoV-2 as the

181 independent variable and multiple regression analyses adjusting for testing laboratory, plant size, frequency of
182 testing (weekly, fortnightly or monthly) and location of plant (by district). We excluded results where
183 wastewater tested negative for SARS-CoV-2 and where the levels of SARS-CoV-2 in genome copies per millilitre
184 were less than the laboratory's reported level of quantification. For all analyses, a p value <0.05 was
185 considered to be statistically significant. All analyses and graphs were done using Stata (StataCorp, LLC, USA, v
186 17).

187 To investigate the potential that quantitative changes in SARS-CoV-2 wastewater levels could predict a new
188 wave of infection, we identified wastewater results in log-copies of SARS-CoV-2 from samples obtained from
189 13 WWTP tested by laboratory 1 only. We selected laboratory 1 only to eliminate the impact of variation in
190 test methodology on genome copies/ml and because it had representative sites across South Africa. We paired
191 these by epidemiological week with the positivity rates of SARS-CoV-2 tests performed on persons presenting
192 for COVID-19 investigations in the health districts where these plants are located. We calculated the sensitivity
193 and specificity of a 'wastewater test' to detect a new wave as defined by the 'gold-standard', which we defined
194 as a positivity rate amongst clinical PCR tests exceeding 10% on the upward trajectory of a wave. We chose to
195 utilise the clinical positivity rate to define our gold standard, rather than the National Department of Health's
196 (NDoH) definition of a wave based on incidence rates [14], as positivity rate may more sensitively detect a rise
197 in population burden of disease and is more robust in changes to testing behaviour. We defined the
198 wastewater 'test' as a Y% increase in log-gene copies in the week of interest, compared to the average of the
199 last Z weeks, where Z = 1 to 6 weeks and Y varied from 5% to 200%. In the main analyses of predictive
200 indicators, we calculated the sensitivity of the 'test' as the fraction of times our 'wastewater test' in the
201 epidemiological week of interest, detected the crossing of the 10% threshold X= 1-, 2-, 3- or 4-weeks' ahead. In
202 supplementary analyses (see supplementary material), we also investigated tests based on any of the N
203 previous weeks. The downward trajectory of a wave defines the 'gold-standard' for the specificity calculation.
204 Regarding specificity, our 'test' should not predict a wave of new infection between the time points that
205 positivity drops below 10% for the first time and the lowest positivity rate between two waves. Sensitivity
206 calculations were based on as many available data points as possible (between 9 and 22 data points). Each of
207 the 15 sites contributed at most one data point for each of two waves. Specificity calculations were based on
208 between 103 and 123 data points, since more than one-time point in the downward trajectory was

209 considered. However, data was missing during some weeks, which reduced the number of data points
 210 available for calculations. These methods are further described in the Supplemental Methods for details.

211

212

213

214 **Table 1: Laboratory methods used for sampling, concentration, RNA extraction volume and method SARS-**
 215 **CoV-2 amplification and detection by partner laboratories**

Name of laboratory	Sampling method	Virus concentration	Nucleic acid extraction (volume of concentrate used for extraction)	SARS-CoV-2 RT-PCR & amplification (volume of RNA extract used for PCR)	Molecular analysis platform	Level (lower limit) of quantification (genome copies/ml)
Lab 1	Grab	Polyethylene Glycol precipitation	Omega Bio-tek ENZA total RNA Kit II (50 µL)	2019-nCoV CDC EUA Kit (4 µL)	Qiagen Rotor-Gene 6000 (5-plex) (Qiagen)	2,06
Lab 2	Grab	Ultrafiltration (Centricon® Plus-70 centrifugal ultra-filter device)	QIAamp® viral RNA mini kit, Qiagen (140 µL)	RT-ddPCR ^b using CDC 2019-nCoV_N2 Primers, Fam Labelled, double quenched probes (5 µL)	QX200 AutoDG Droplet Digital PCR System (Bio-rad)	200
Lab 3	Grab	Ultrafiltration (Amicon® Ultra-15 Centrifugal Filter Unit)	Omega Bio-Tek Mag-Bind® Viral DNA/RNA 96 Kit (50 µL)	CDC 2019-Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel (2,5 µL)	Rotor-Gene Q (Qiagen)	64
Lab 4	Passive	Passive sampler and resuspension in phosphate buffered saline	MN DNA/RNA pathogen extraction Kit (VNP)	TaqPath COVID-19 CE-IVD RT-PCR Kit (Thermo Fisher) (VNP)	QuantStudio 5 (Applied Biosystems)	40
Lab 5	Grab	Ultrafiltration (Centricon® Plus-70 centrifugal ultra-filter device)	QIAmp® viral RNA mini kit (Qiagen) (140 µL)	Allplex™ 2019-nCoV (Seegene) and the RNA EDX standard (Bio-Rad) (8 µL)	7500 real-time PCR system (Applied Biosystems)	1,49
Lab 6	Grab	Skimmed milk flocculation	MagMAX Viral/Pathogen Nucleic Acid Isolation Kit (200 µL)	TaqPath COVID-19 CE-IVD RT-PCR Kit (Thermo Fisher) (2 µL)	QuantStudio™ 5 Real-Time PCR System 96-well, 0.1 mL, desktop (Applied Biosystems)	2,33
Lab 7	Grab	None – sample is	ZymoBiomics RNA	Allplex™ 2019-nCoV Assay	QuantStudio 5 (Applied Biosystems)	Not provided

		centrifuged then supernatant analysed	Extraction Kit (1000 µL)	and the EDX SARS-CoV-2 standard (Bio-Rad) (8 µL)	Biosystems)	
Lab 8	Grab	Skimmed milk flocculation	QIAamp® Ultrasens® Virus kit (1000 µL)	Allplex™ 2019-nCoV Assay and 2019_nCoV_N positive control plasmid (Integrated DNA Technologies, Inc, Coralville, IA) (8 µL)	QuantStudio™ 5 Real-Time PCR System (Applied Biosystems)	16

216 VNP=volume used for reaction not provided

217 **Results**

218 A total of 2,158 wastewater samples were tested by the SACCESS network between 01 June 2021 – 31 May
 219 2022 inclusive, coinciding with the 3rd and 4th waves of the COVID-19 pandemic. Of these, 1388 (64%) tested
 220 positive for SARS-CoV-2. During the 3rd and 4th waves, 638 tests were conducted and 543 (85%) were positive.
 221 During the interwave periods, 1,512 samples were tested and of these, 842 (55%) were positive. Amongst all
 222 positive tests 207/1,388 were below the limit of quantification of the test method used by the testing
 223 laboratory and were excluded from correlation and regression analyses.

224 Figures 1A-F show the times series of laboratory confirmed cases by district and log genome copies of SARS-
 225 CoV-2/mL of wastewater- for WWTPs across all study metros for the entire review period, excluding those
 226 WWTPs with three or fewer data points after elimination of results below the level of quantification. On visual
 227 inspection of time series graphs, high concentrations were mostly seen during the waves of the pandemic, and
 228 low concentrations during the interwave period. The levels of SARS-CoV-2 from sentinel plants in Gauteng
 229 subdistricts that were tested by the NICD from epidemiological week 40 (just before the onset of fourth wave)
 230 are illustrated in Figure 2 and demonstrate a rise in levels which precede the increase in cases.

231 Table 2 lists the correlation and multivariable regression coefficients (adjusted for location, size of plant and
 232 testing frequency) between SARS-CoV-2 levels (in genome copies/ml) and weekly laboratory-confirmed case
 233 loads in the corresponding district/sub-district by testing laboratory. Table 3 presents results of the identical
 234 analyses but conducted for plants located in Gauteng Province only, where the clinical epidemiology of SARS-
 235 CoV-2 and case loads to which wastewater levels of SARS-CoV-2 were compared were similar by district in

236 Gauteng. Correlation co-efficient between SARS-CoV-2 genome copies/ml wastewater and laboratory-
 237 confirmed cases by district for each specific plant are shown in the Supplement Table S1.

238

239

240 **Table 2. Correlation coefficients and multivariable regression analyses (adjusted for location, size of plant**
 241 **and testing frequency) between SARS-CoV-2 levels (in genome copies/ml) in wastewater and weekly**
 242 **laboratory-confirmed clinical case load of SARS-CoV-2 from 01 June 2021 until 31 May 2022 from 87 plants**
 243 **across all nine South African provinces, for all samples and by testing laboratory.**

Testing laboratory	# samples tested	# positive (%)	Range (log genome copies/mL wastewater)	LOQ (Log genome copies/ml)	# below LOQ / # positive (%)	Correlation coefficient with lab-confirmed cases ¹ (*=p-value <0.001)	Multivariable variate regression with case load ³	
							Co-efficient (95% confidence interval)	Adjusted R-squared ³ (*p-value<0.001)
All	2158	1388 (64)	-1.84 - 9.5	various	207/1388	0.36	0.66 (0.60 - 0.72)	0.59*
Lab 1	199	148 (74)	-0.12 - 3.27	0.31	7/148	0.64*	0.87 (0.73 - 1.01)	0.45*
Lab 2	82	82 (100)	3.01 - 6.47	2.30	0/82 (0)	0.78*	0.73 (0.59 - 0.87)	0.57*
Lab 3	230	80 (34)	0- 3.12	1.81	42/80	0.53*	0.68 (0.52 - 0.85)	0.36*
Lab 4	296	210 (71)	-0.47-5.52	1.60	32/210	0.41*	1.23 (0.97 - 1.50)	0.33*
Lab 5	621	572 (92)	-0.54-3.43	1.52	27/572	0.68*	0.73 (0.66 - 0.80)	0.53*
Lab 6	168	133 (79)	-0.84-3.94	0.37	6/133	0.42*	0.53 (0.37 - 0.70)	0.24*
Lab 7	54	36 (67)	-0.33-9.5	NA	NA	NA	NA [#]	NA [#]
Lab 8	508	127 (25)	-1.84-4.32	1.20	93/127	0.19*	0.14 (0.06 - 0.21)	0.13

244 LOQ = Level of quantification. ¹excluding cases below level of quantification; ²Lab 7 was excluded from correlation and regression due to
 245 the absence of LOQ; ³adjusted for district, plant size and wastewater sampling frequency

246 **Table 3. Correlation coefficients and multivariable regression analyses (adjusted for location, size of plant**
 247 **and testing frequency) between SARS-CoV-2 levels (in genome copies/ml) in wastewater and weekly**
 248 **laboratory-confirmed clinical case load of SARS-CoV-2 from 44 wastewater treatment plants in Gauteng**
 249 **Province from 01 June 2021 until 31 May 2022.**

Testing laboratory- in Gauteng	# samples tested	# positive (%)	Range (log genome copies/mL wastewater)	Level of Quantification (Log genome copies/ml)	# below LOQ / # positive (%)	Correlation with lab-confirmed cases by district (*=p-value <0.05)	multivariable regression with case load ¹	
							Co-efficient (95% confidence interval)	Adjusted R ² (*p-value<0.001)
All	1050	670 (64)	-1.84 - 9.5		94/670	0.50*	0.54 (0.46 - 0.62)	0.39
Lab 1	199	148 (74)	-0.12 - 3.27	0.31	7/148	0.64*	0.87 (0.73 - 1.01)	0.45*
Lab 5	260	250 (96)	-0.54-3.2	1.52	9/572	0.70*	0.68 (0.58 - 0.78)	0.52*
Lab 6	168	133 (79)	-0.84-3.04	0.37	6/133	0.42*	0.53 (0.37 - 0.70)	0.24*
Lab 8	369	103 (28)	-1.84-4.33	1.20	72/127	0.25*	0.18 (0.08 - 0.29)	0.13*

250 ¹Adjusted for district, sampling frequency, testing lab, and plant size

251

252

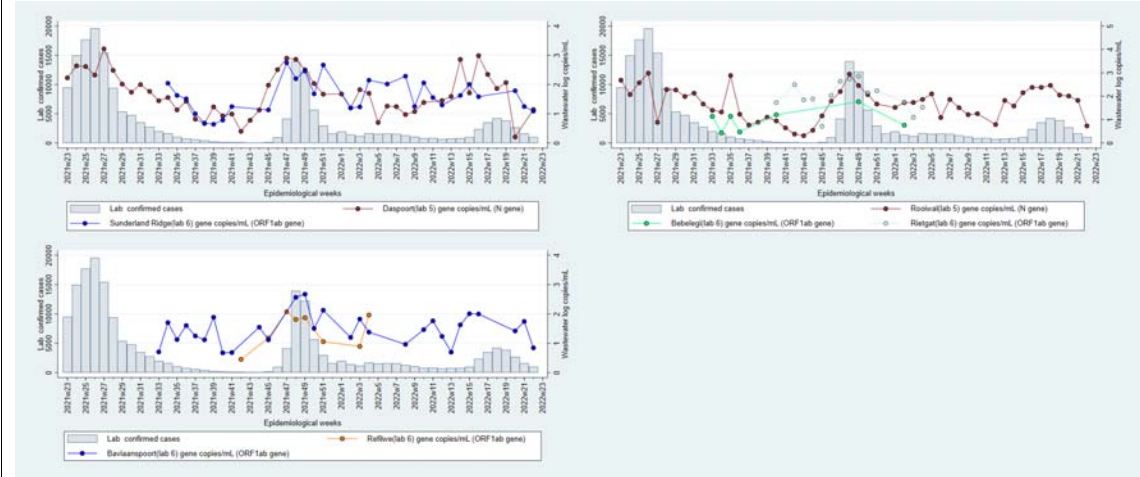
253

254

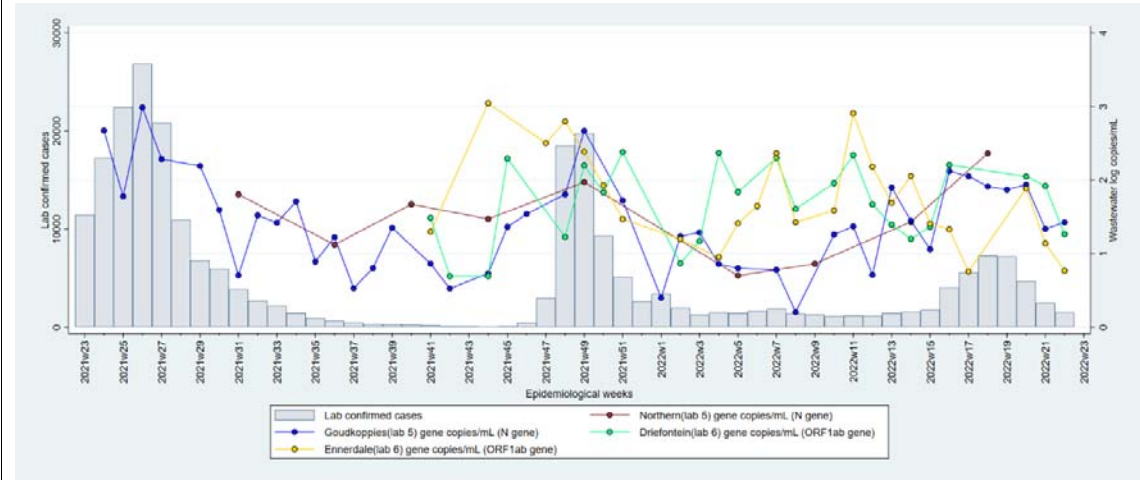
255

256

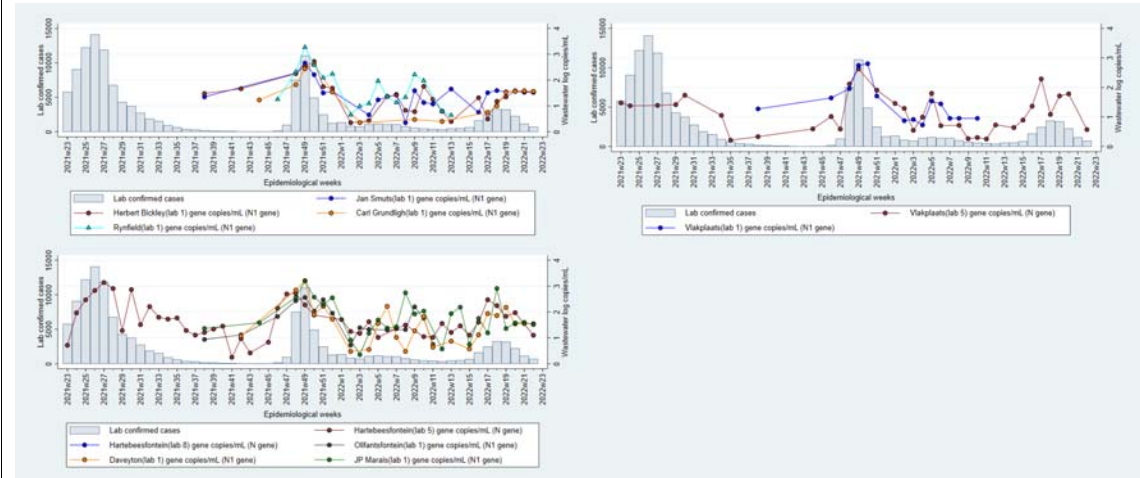
A. City of Tshwane



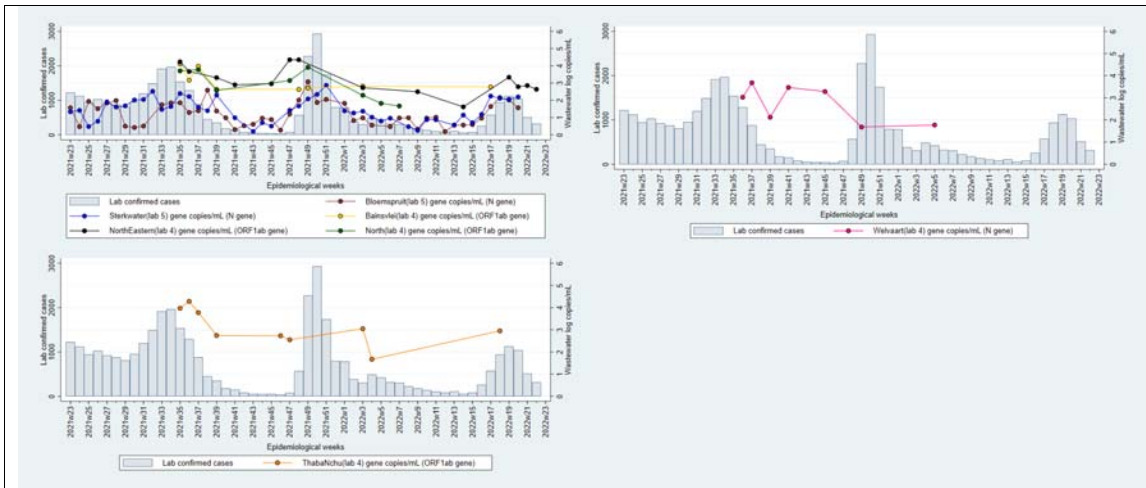
B. City of Johannesburg



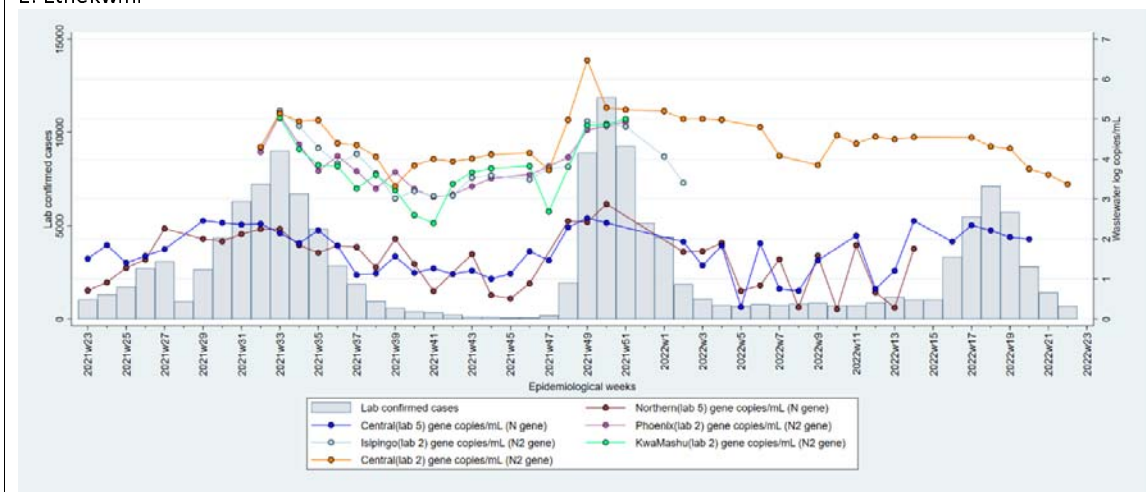
C. City of Ekurhuleni



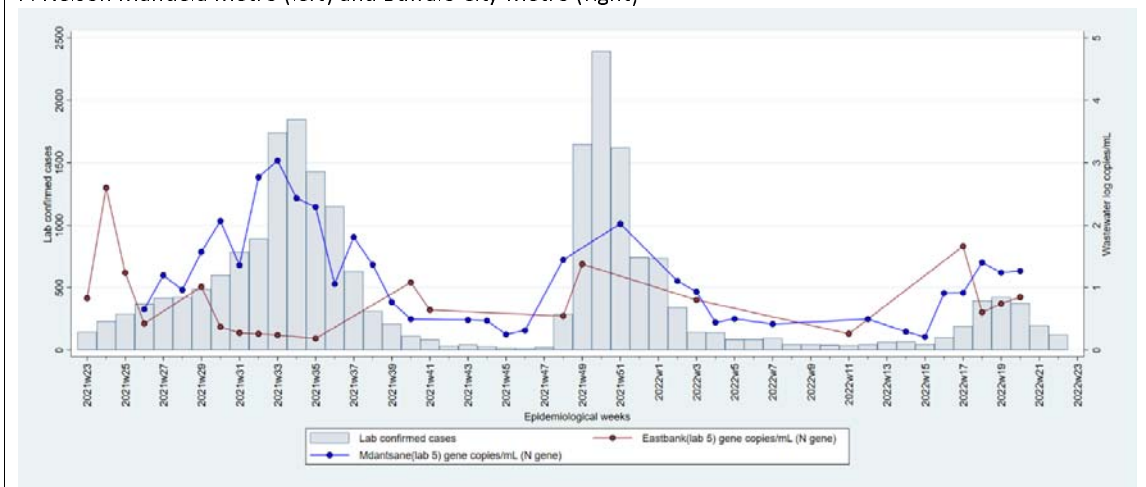
D. Mangaung

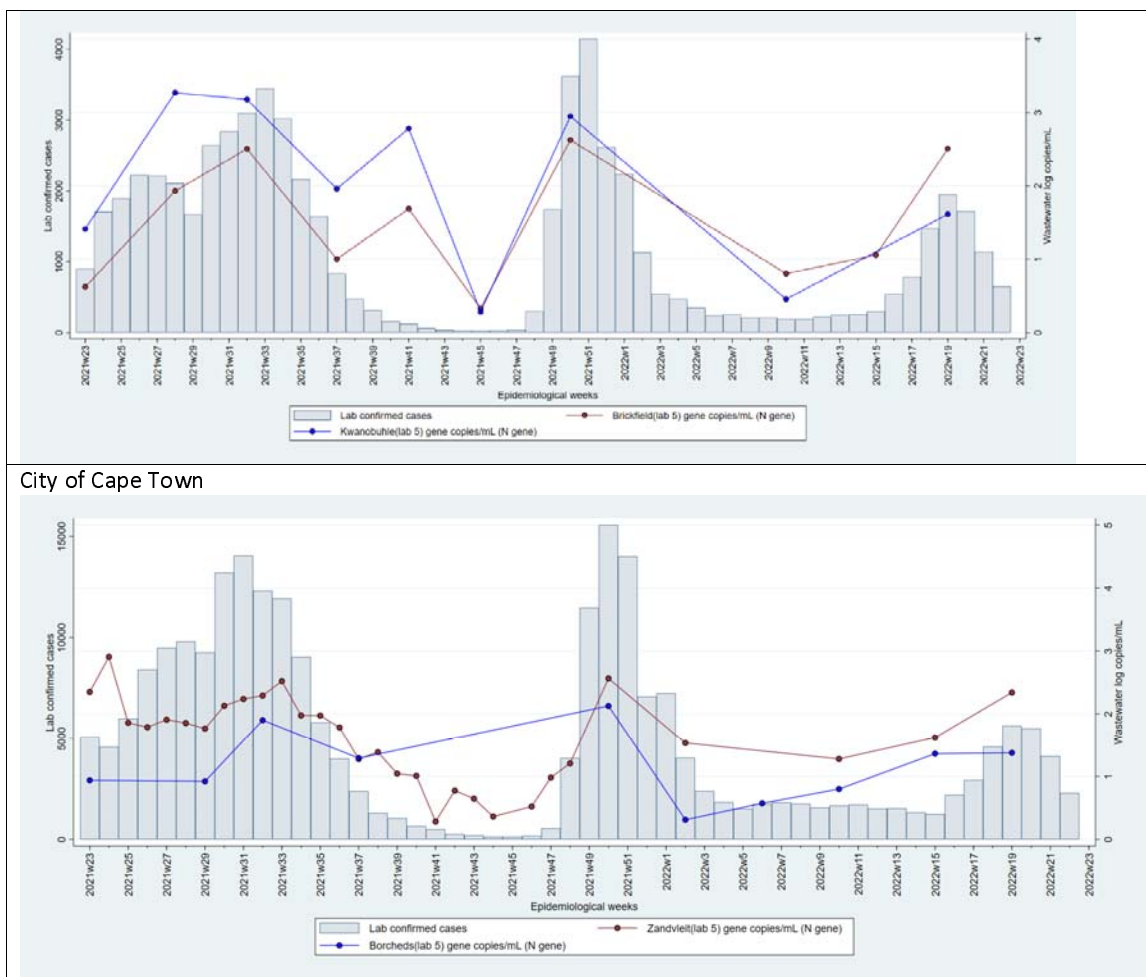


E. Ethekwini



F. Nelson Mandela Metro (left) and Buffalo City Metro (right)





257

258 Figure 1A-F. Laboratory-confirmed cases (left axis, grey bars) and log genome copies SARS-CoV-2 per millilitre
 259 of wastewater (right axis, lines), by epidemiological week. Colors represent different wastewater testing
 260 plants. Results below the level of quantification of each laboratory are omitted, as are wastewater treatment
 261 plants with 3 or fewer results. The testing laboratory (1-7) is indicated in the key.

262

263

264

265

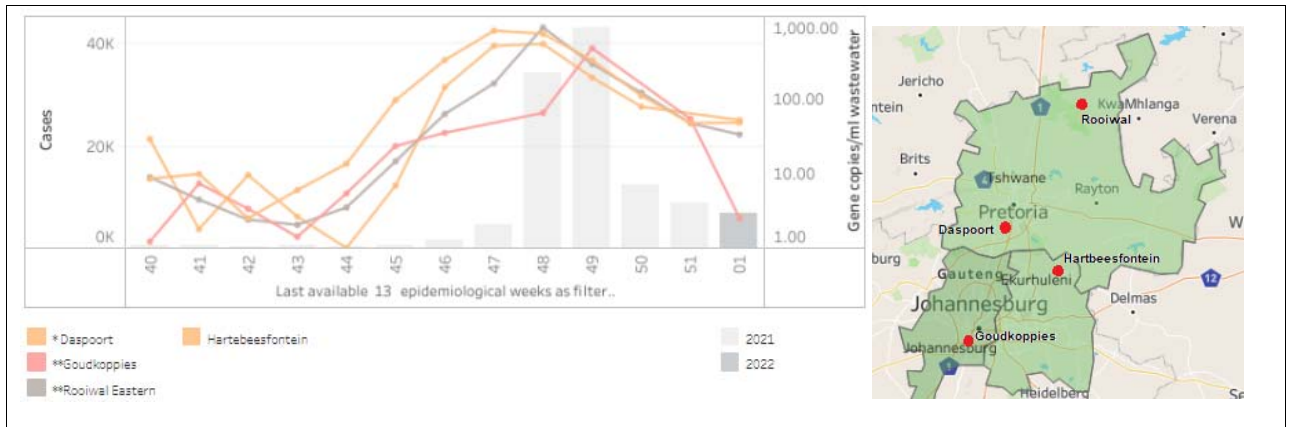
266

267

268

269

270

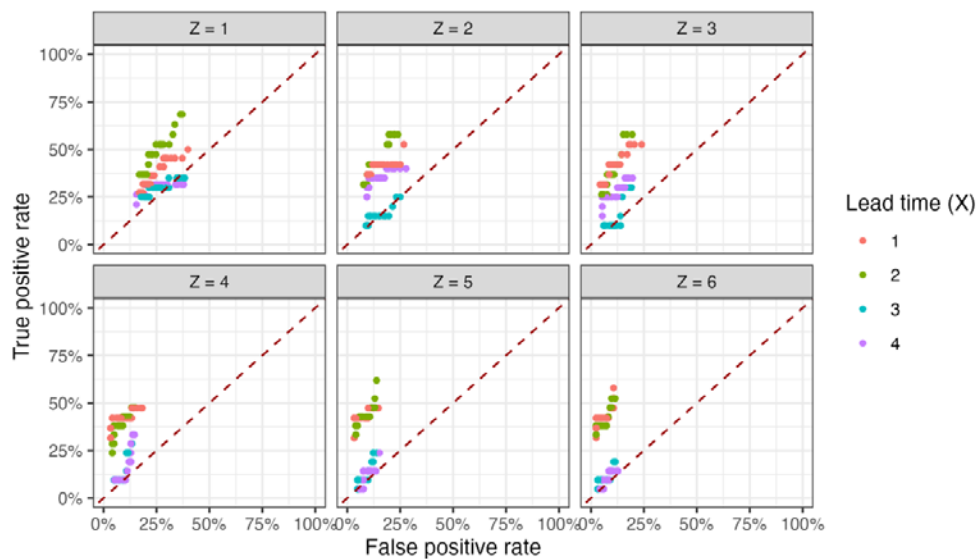


271 Figure 2. SARS-CoV-2 levels in wastewater (genome copies/ml) for four wastewater treatment plants in
 272 Gauteng (see inset map) and case load of laboratory-confirmed SARS-CoV-2 cases in Gauteng Province by
 273 epidemiological week 2021

274 There was a significant association between wastewater levels and case load in the multivariate linear
 275 regression after adjusting for testing laboratory, district, sampling frequency, and plantsize ($p < 0.001$, $R^2 = 0.6$).
 276 Similarly, a significant association was seen when the sample regression model was restricted to the Gauteng
 277 province where the clinical epidemiology of SARS-CoV-2 and case load to which wastewater levels of SARS-
 278 CoV-2 were compared were similar by district in Gauteng ($p < 0.001$, $R^2 = 0.4$).

279 Regarding a predictive rule, sensitivity was highest for the lead period of one and two weeks, and when the
 280 average log genome copies/ml included values up to six preceding weeks from the point of measurement. The
 281 time points at which the rule was generated that were further from the point of declaration of a wave (a
 282 clinical positivity rate of $> 10\%$) were generally less sensitive. None of the tests with specified lead times had
 283 sufficiently high sensitivity and specificity to be provide robust indicators of the start of a new wave (Figure 3).
 284 The most informative indicator we identified was whether or not the wastewater level in any of the previous 5
 285 weeks was at least 100% (2X) higher than the average of the previous 5 weeks; this indicator had a sensitivity
 286 of 79% and a specificity of 78% (Figure S1).

287



288

289

290 **Figure 3:** True positive rate (sensitivity) and false positive rate (1-specificity) measures of a defined
291 rule vs gold standard to establish the ability of wastewater levels of SARS-CoV-2 to predict an
292 epidemic wave. The 'gold-standard' for sensitivity is defined by the epidemiological week when the
293 positivity rate of clinical testing exceeds 10% on the upward trajectory of a wave. The 'gold-
294 standard' for specificity is defined on the downward trajectory of a wave, at the time point when the
295 positivity rate drops below 10% for the first time and the lowest positivity rate between two waves.
296 The test used in this figure is defined as whether or not the SARS-CoV-2 log-gene copies in
297 wastewater X weeks ago exceeded the average log-gene copies of the previous Z weeks by a
298 percentage varied from 5% to 200%.

299

300

301 Discussion

302 Results of quantitative SARS-CoV-2 testing across South African wastewater treatment plants have
303 demonstrated good correlation with district laboratory-confirmed COVID-19 case load, and provided early
304 warning of the 4th wave of SARS-CoV-2 in Gauteng Province in November-December 2021. Our results
305 demonstrated variation in the strength of correlation across testing laboratories, and redundancy of findings
306 across co-located testing plants, suggesting that test methodology should be standardised and that
307 surveillance networks may utilise a sentinel site model without compromising the value of WBE findings for
308 public health decision-making. Further research is needed to identify optimal test frequency so as to identify
309 predictive and interpretive rules to support early warning and public health action. Our findings support
310 investment in WBE for SARS-CoV-2 surveillance in low and middle-income countries.

311 Our data series, comprising 1388 observations over 2 waves of COVID-19, illustrate that the levels of SARS-
312 CoV-2 in wastewater follow trends in laboratory-confirmed SARS-CoV-2 case load, and demonstrate high
313 correlation and strong regression analysis indicators between case load and genome copies/ml from single
314 WWTPs over time (Supplementary table). These findings concur with global observations regarding the
315 usefulness of WBE of SARS-CoV-2 as an indirect measure of population burden of disease [7] and support the
316 recommendations of the WHO interim guidance on Environmental surveillance for SARS-CoV-2, that WBE
317 provides insight into the distribution and time trends in SARS-CoV-2 levels at community level [1]. Our findings,
318 generated at a time when SARS-CoV-2 clinical testing was widespread and accessible, point to the usefulness
319 of WBE to provide insight into the distribution of cases and relative burden of disease as clinical testing rates
320 decline and hospitalisation rates decrease with subsequent waves of infection [19].

321 The variation in the strength of correlation between wastewater levels of SARS-CoV-2 and laboratory-
322 confirmed case load between plants tested by different laboratories, even when plants were co-located in the
323 same health district where timing of the wave and burden of clinical disease to which the wastewater levels
324 were compared, were constant. It is important to understand and address variation in SARS-CoV-2 levels, as
325 variation may contribute to weaker correlation with clinical burden of disease and may therefore undermine
326 the potential usefulness of WBE data.

327 A key source of variation in our context arose from the different sample collection, extraction and detection
328 methods, and different wastewater sample volumes used by participating laboratories to determine
329 quantitative SARS-CoV-2 levels. Passive sampling, as utilised by one of our laboratories, will trap particulate
330 matter, and will report consistently higher wastewater levels compared with grab samples, the method
331 preferred by most of our laboratories, as RNA seems to concentrate more in solid phase of the wastewater
332 matrix [20], [21]. Differences in concentration and RNA extraction techniques and the volume of concentrate
333 that undergoes RNA extraction will retain more or less template RNA for detection[22]. Finally, the different
334 genes of SARS-CoV-2 that are detected and amplified by different RT-PCR detection kits, may be present at
335 variable amounts in wastewater because of differential decay of viral fragments in a hostile environment
336 [Personal communication, Doug Manuel, Coronavirus Variants, Rapid Response Network, Canada]. We also
337 observed variable extraction efficiencies in the quality assessment panels (data not shown), suggesting that
338 some methodologies are more sensitive and precise than others, a finding observed elsewhere [23]. Special

339 attention should be given to digital PCR technologies, as we observed higher levels, stronger correlation and
340 closer tracking of clinical cases with results from the laboratory using this technology [23]

341 Sources of variation in SARS-CoV-2 levels in wastewater other than those that may be ascribed to test
342 methodology [24, 25] may also have accounted for variation in strength of correlation with clinical case load.
343 In our network, these may include the following: Firstly, the district boundaries to which laboratory-confirmed
344 case loads were allocated, do not correspond with sewage reticulation networks. We were unable to
345 reconfigure cases to provide data from populations resident within the sewage catchment areas. If different
346 districts reported different case loads correlation of results from a single plant serving populations across
347 district boundaries may be weaker. Secondly, the WWTPs from where samples were taken were large, serving
348 a median of 100,000 persons (range 4,000-1,200,000). The geographical areas served by these WWTPS were
349 socio-economically diverse. It is well established that high population density and poorer socio-economic
350 conditions are associated with higher SARS-CoV-2 transmission and therefore higher disease burden,
351 hospitalisation and mortality rates [26]. Yet these areas may be under-represented in sewage networks
352 because of informal settlements, or dilution with effluent from more well-resourced areas may artificially
353 lower SARS-CoV-2 levels. Thirdly, the high proportion of asymptomatic disease [27], and differential access to
354 SARS-CoV-2 testing may have led to discontinuity between wastewater levels and clinical cases [28]. Fourthly,
355 and particularly problematic in larger networks, differential SARS-CoV-2 RNA decay rates within the sewage
356 network [26], destruction of RNA through mixing with industrial effluent [19], and dilution of RNA through
357 environmental sources such as rain-water may alter the relationship to clinical case load. Finally, but less likely
358 in large wastewater treatment plants, variation in population size over time through in-and outward migration
359 may artificially change SARS-CoV-2 levels in wastewater. These sources of variation may explain why some
360 testing laboratories, particularly those testing plants in the rural provinces with smaller and less well-resourced
361 communities demonstrated lower correlation with case load. These sources of variation may also explain why
362 our data did not support our attempt to create a highly sensitive and generalizable predictive rule for a new
363 clinical wave.

364 We chose not to obtain additional contextual data pertaining to wastewater flow rates and volumes, nor did
365 we attempt to normalise data to population load, all of which would have allowed us to estimate burden of
366 disease due to COVID-19 in sewer catchment areas. This is a key differentiating factor in our approach

367 compared with well-resourced contexts. The Netherlands, for example, publishes the number of SARS-CoV-2
368 particles per 100,000 persons by health district [29]. Our decision was based on a number of factors, most
369 notably that only 60% of South African population is served by waterborne sewage, and a smaller proportion
370 of that served by wastewater treatment networks, the balance using septic tanks [30]. Other factors include
371 the difficulty of standardising PCR methods to support normalisation, unavailability of sewage reticulation
372 network maps and therefore uncertainties regarding precise geographical areas being tested, and
373 uncertainties regarding decay rates and degradation of RNA within complex industrial and residential sewage
374 networks.

375 Under the current circumstances, and with experience learned from our network, we reflect on the suitability
376 of the current configuration of wastewater testing SARS-CoV-2 for surveillance purposes. The sources of
377 variation (including laboratory test methodology) that lead to differing correlation and regression indicators
378 between levels and case-load , and the relatively high proportion of that is not served by wastewater
379 treatment plants, suggest as we anticipated, that estimation of precise population burden of disease in our
380 context is presently unattainable. Further, testing large numbers of samples from geographically disparate
381 plants creates high workloads and sample transport costs. Regarding SARS-CoV-2 levels, we observed that
382 plants co-located in the same district revealed similar temporal trends, suggesting redundancy in the data to
383 support public health decision-making authorities. It is our opinion therefore, that extensive networks that test
384 influent from large wastewater treatment plants across entire geographical areas may not be useful in low and
385 middle income settings from a public health perspective.

386 Therefore, given these limitations in our middle-income context, we propose a sentinel site WBE surveillance
387 model with centralised testing to eliminate variation due to different laboratory test methods. Such a network
388 serves public health decision-making needs by providing relative population burden, temporal trends and
389 geographic distribution of viral infection and variants. These data may be triangulated with results from clinical
390 testing to provide a fuller picture of disease epidemiology. In the context of a such a model, with carefully
391 chosen sites that ensure maximal population representativeness (including wastewater treatment plants,
392 environmental sources, rivers downstream of informal settlements) it may be possible with further research,
393 to determine of the population burden of disease. Key research questions include the added value of more
394 frequent testing normalising for population size, measuring wastewater flow-rates, and correcting for

395 environmental factors such as ambient temperature and dilution with rainfall in defined sewage reticulation
396 networks where populations are well characterised. These findings may serve public health decision-making at
397 a local level, and may inform generalisability of results from sentinel surveillance sites. Presently, despite its
398 many advantages and potential, WBE cannot suffice as a stand-alone surveillance methodology in our or any
399 context. Its key limitation is the inability to provide insight into disease severity. Therefore, WBE findings
400 should always be interpreted with other measures of population health, including hospitalisation and mortality
401 data.

402 **Conclusion**

403 We have presented findings from the wastewater-based surveillance for SARS-CoV-2 during the 3rd and 4th
404 waves of SARS-CoV-2 in South Africa which demonstrate both the value and limitations of wastewater-based
405 epidemiology in low and middle income countries. Further, our work has highlighted key areas for
406 investigation to support interpretability of quantitative SARS-CoV-2 levels in wastewater with reference to
407 burden of disease, and provided evidence to support the optimal configuration of WBE surveillance networks.
408 Given trends in viral evolution and immune evasion, SARS-CoV-2 will remain a concern for population health.
409 Investment in WBE as a surveillance tool will continue to support provision of data to address key
410 uncertainties in decision-making aimed at preserving population health.

411 **Declarations**

412 ***Ethics approval and consent to participate***

413 Not applicable

414 ***Availability of data and materials***

415 Data used for this study is available on request

416 ***Competing interests***

417 All authors declare no competing interests

418 ***Funding***

419 This study was funded by the Water Research Commission and National Institute for Communicable Diseases

420 **Authors' contributions**

421 Study was conceived by KM, MS and the SACCESS network; sample collection, processing, laboratory testing
422 and generation of results were done by NN, SR, MY, WL, LS, GP, LC, JM, FB, LP, DV, AD, DJ, AG, SG, NM, MV,
423 AM, NB; data visualisation and analyses were done by CI, NN, KM, LM, CS, JP and TM; all authors contributed
424 to the writing of various sections of the manuscript; CI, NN and KM compiled the first draft of the manuscript;
425 all authors read and approved the final version of the manuscript

426 **Acknowledgements:**

427 We wish to acknowledge the following in no particular order:

- 428 • The NICD Centre for Respiratory Disease and Meningitis are thanked for their assistance in setting up
429 and troubleshooting PCR testing, and ongoing supportive collaboration
- 430 • The staff from participating labs
- 431 • The contributions of local government and wastewater treatment staff to sample collection and
432 transport is acknowledged and appreciated.
- 433 • Students who supported with sample collections and processing the samples: Mr Thoriso Mooa, Ms
434 Unarine Matodzi, Ms Phiwinhlanhla Nkosi – SAMRC-TB Platform
- 435 • The Water Research Commission, for their vision and support
- 436 • The NICD SARS-CoV-2 epidemiology and IT team members for sharing district and sub-district case
437 burdens in order to generate graphs. These team members include Andronica Moipone Shonhiwa,
438 Genevieve Ntshoe, Joy Ebonwu, Lactatia Motsuku, Liliwe Shuping, Mazvita Muchengeti, Jackie
439 Kleynhans, Gillian Hunt, Victor Odhiambo Olago, Husna Ismail, Nevashan Govender, Ann Mathews,
440 Vivien Essel, Veerle Msimang, Tendesayi Kufa-Chakezha, Nkengafac Villyen Motaze, Natalie Mayet,
441 Tebogo Mmaborwa Matjokotja, Mzimasi Neti, Tracy Arendse, Teresa Lamola, Itumeleng Matiea,
442 Darren Muganhiri, Babongile Ndlovu, Khuliso Ravhuhali, Emelda Ramutshila, Salaminah Mhlanga,
443 Akhona Mzoneli, Nimesh Naran, Trisha Whitbread, Mpho Moeti, Chidozie Iwu, Eva Mathatha,
444 Fhatuwani Gavhi, Masingita Makamu, Matimba Makhubele, Simbulele Mdleleni, Tsumbedzo
445 Mukange, Trevor Bell, Lincoln Darwin, Fazil McKenna, Ndivhuwo Munava, Muzammil Raza Bano,
446 Themba Ngobeni
- 447 • The Staff of SACCESS network laboratories, for their assistance in generating these results

448 References

- 449 1. World Health Organization. Environmental surveillance for SARS-CoV-2 to complement public health
450 surveillance - Interim Guidance. 2022; April:1–47. [https://www.who.int/publications/i/item/WHO-HEP-ECH-](https://www.who.int/publications/i/item/WHO-HEP-ECH-WSH-2022.1)
451 [WSH-2022.1](https://www.who.int/publications/i/item/WHO-HEP-ECH-WSH-2022.1). Accessed 15 Sep 2022.
- 452 2. Radu E, Masseron A, Amman F, Schedl A, Agerer B, Endler L, et al. Emergence of SARS-CoV-2 Alpha lineage
453 and its correlation with quantitative wastewater-based epidemiology data. *Water Res.* 2022;215
454 March:118257.
- 455 3. Claro ICM, Cabral AD, Augusto MR, Duran AFA, Graciosa MCP, Fonseca FLA, et al. Long-term monitoring of
456 SARS-CoV-2 RNA in wastewater in Brazil: A more responsive and economical approach. *Water Res.* 2021;203
457 July:117534.
- 458 4. Weidhaas J, Aanderud ZT, Roper DK, VanDerslice J, Gaddis EB, Ostermiller J, et al. Correlation of SARS-CoV-2
459 RNA in wastewater with COVID-19 disease burden in sewersheds. *Sci Total Environ.* 2021;775:145790.
- 460 5. Fazli M, Sklar S, Porter MD, French BA, Shakeri H. Wastewater-Based Epidemiological Modeling for
461 Continuous Surveillance of COVID-19 Outbreak. *medRxiv.* 2021;:2021.10.19.21265221.
- 462 6. Radu E, Masseron A, Amman F, Schedl A, Agerer B, Endler L, et al. Emergence of SARS-CoV-2 Alpha lineage
463 and its correlation with quantitative wastewater-based epidemiology data. *Water Res.* 2022;215 January.
- 464 7. Shah S, Gwee SXW, Ng JQX, Lau N, Koh J, Pang J. Wastewater surveillance to infer COVID-19 transmission: A
465 systematic review. *Science of the Total Environment.* 2022;804:150060.
- 466 8. Bonanno Ferraro G, Veneri C, Mancini P, Iaconelli M, Suffredini E, Bonadonna L, et al. A State-of-the-Art
467 Scoping Review on SARS-CoV-2 in Sewage Focusing on the Potential of Wastewater Surveillance for the
468 Monitoring of the COVID-19 Pandemic. *Food and Environmental Virology.* 2021;1:3.
- 469 9. COVIDPoops19 Dashboard. 2021. <https://www.covid19wbec.org/covidpoops19>. Accessed 15 Sep 2022.
- 470 10. Panchal D, Tripathy P, Prakash O, Sharma A, Pal S. SARS-CoV-2: Fate in water environments and sewage
471 surveillance as an early warning system. *Water Sci Technol.* 2021;84:1–15.
- 472 11. Robins K, Leonard AFC, Farkas K, Graham DW, Jones DL, Kasprzyk-Hordern B, et al. Research needs for
473 optimising wastewater-based epidemiology monitoring for public health protection. *J Water Health.*
474 2022;20:1284–313.
- 475 12. Safford HR, Shapiro K, Bischel HN. Wastewater analysis can be a powerful public health tool—if it's done
476 sensibly. *Proc Natl Acad Sci U S A.* 2022;119:1–5.
- 477 13. Archer E, Preiser W, Suchard M. The SACCESS network for COVID-19 wastewater surveillance: a national
478 collaboration for public health responsiveness. *South African Heal Rev.* 2021;:215–23.
- 479 14. National Institute for Communicable Diseases. Proposed definition of COVID-19 wave in South Africa.
480 Communicable Diseases Communiqué . 2021. [https://www.nicd.ac.za/wp-](https://www.nicd.ac.za/wp-content/uploads/2021/11/Proposed-definition-of-COVID-19-wave-in-South-Africa.pdf)
481 [content/uploads/2021/11/Proposed-definition-of-COVID-19-wave-in-South-Africa.pdf](https://www.nicd.ac.za/wp-content/uploads/2021/11/Proposed-definition-of-COVID-19-wave-in-South-Africa.pdf). Accessed 21 Mar 2022.
- 482 15. WHO. Guidelines for environmental surveillance of poliovirus circulation. *Vaccine and Biologicals.*
483 2003;03:28. <https://apps.who.int/iris/handle/10665/67854>. Accessed 27 Jan 2023.
- 484 16. Johnson R, Muller CJF, Ghoor S, Louw J, Archer E, Surujlal-Naicker S, et al. Qualitative and quantitative
485 detection of SARS-CoV-2 RNA in untreated wastewater in Western Cape Province, South Africa. *South African*
486 *Med J.* 2021;111:198–202.
- 487 17. SAMRC COVID-19 prevention research programme: wastewater surveillance for SARS-CoV-2. Wastewater
488 sampling guide (1 February 2021). [https://www.samrc.ac.za/wbe/SARS-CoV-](https://www.samrc.ac.za/wbe/SARS-CoV-2WastewaterCollectionManual.pdf)
489 [2WastewaterCollectionManual.pdf](https://www.samrc.ac.za/wbe/SARS-CoV-2WastewaterCollectionManual.pdf). Accessed 15 Sep 2022.
- 490 18. Bhagwan J. A compendium of emerging South African testing methodologies for detecting of SARS-CoV-2
491 RNA in wastewater surveillance. 2020. www.wrc.org.za. Accessed 15 Sep 2022.

- 492 19. Amoah ID, Abunama T, Awolusi OO, Pillay L, Pillay K, Kumari S, et al. Effect of selected wastewater
493 characteristics on estimation of SARS-CoV-2 viral load in wastewater. *Environ Res.* 2022;203 August
494 2021:111877.
- 495 20. Gholipour S, Ghalhari MR, Nikaeen M, Rabbani D, Pakzad P, Miranzadeh MB. Occurrence of viruses in
496 sewage sludge: A systematic review. *Sci Diabetes Self-Management Care.* 2022;824:153886.
- 497 21. Bivins A, Kaya D, Ahmed W, Brown J, Butler C, Greaves J, et al. Passive sampling to scale wastewater
498 surveillance of infectious disease: Lessons learned from COVID-19. *Sci Total Environ.* 2022;835:155347.
- 499 22. Bolanaki E, Kottaridi C, Dedepsidis E, Kyriakopoulou Z, Pliaka V, Pratti A, et al. Direct extraction and
500 molecular characterization of enteroviruses genomes from human faecal samples. *Mol Cell Probes.*
501 2008;22:156–61.
- 502 23. Flood MT, D'Souza N, Rose JB, Aw TG. Methods Evaluation for Rapid Concentration and Quantification of
503 SARS-CoV-2 in Raw Wastewater Using Droplet Digital and Quantitative RT-PCR. *Food Environ Virol.*
504 2021;13:303–15.
- 505 24. Wade MJ, Lo Jacomo A, Armenise E, Brown MR, Bunce JT, Cameron GJ, et al. Understanding and managing
506 uncertainty and variability for wastewater monitoring beyond the pandemic: Lessons learned from the United
507 Kingdom national COVID-19 surveillance programmes. *J Hazard Mater.* 2022;424 August 2021.
- 508 25. Ahmed W, Simpson SL, Bertsch PM, Bibby K, Bivins A, Blackall LL, et al. Minimizing errors in RT-PCR
509 detection and quantification of SARS-CoV-2 RNA for wastewater surveillance. *Sci Total Environ.* 2022;805.
- 510 26. Ghosh AK, Venkatraman S, Soroka O, Reshetnyak E, Rajan M, An A, et al. Association between overcrowded
511 households, multigenerational households, and COVID-19: a cohort study. *Public Health.* 2021;198:273–9.
- 512 27. Paleker M, Tembo Y, Davies M-A, Mahomed H, Pienaar D, Madhi SA, et al. Public health action. *Public Heal*
513 *Action.* 2021;11:58–60.
- 514 28. Lancaster E, Byrd K, Ai Y, Lee J. Socioeconomic status correlations with confirmed COVID-19 cases and
515 SARS-CoV-2 wastewater concentrations in small-medium sized communities. *Environ Res.* 2022;215:114290.
- 516 29. Virus particles in wastewater. Coronavirus Dashboard. Netherlands.
517 <https://coronadashboard.government.nl/landelijk/rioolwater>. Accessed 27 Jan 2023.
- 518 30. Overview of wastewater treatment in South Africa. Association for Water and Rural Development
519 (AWARD). <https://award.org.za/>. Accessed 27 Jan 2023.
- 520