

Methodological Approach for Wastewater Based Epidemiological Studies for SARS-CoV-2

Harishankar Kopperi^{1,3†}, Athmakuri Tharak^{1†}, Manupati Hemalatha^{1,3†}, Uday Kiran^{2,3,†},
C. G. Gokulan^{2†}, Rakesh K Mishra^{2,3,*}, S Venkata Mohan^{1,3,*}

¹Bioengineering and Environmental Sciences Lab, Department of Energy and Environmental
Engineering (DEEE), CSIR-Indian Institute of Chemical Technology (CSIR-IICT),
Hyderabad-500007, India

²CSIR-Centre for Cellular and Molecular Biology (CSIR-CCMB), Hyderabad-500007, India,

³Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, India

†These authors contributed equally to the work

*Corresponding authors

Correspondence should be addressed to vmohan_s@yahoo.com or mishra@ccmb.res.in

Abstract

Post COVID-19 outbreak, wastewater-based epidemiology (WBE) studies as surveillance system is becoming an emerging interest due to its functional advantage as tool for early warning signal and to catalyze effective disease management strategies based on the community diagnosis. A comprehensive attempt was made in this study to define a methodological approach for conducting WBE studies in the framework of identifying/selection of surveillance sites, standardizing sampling policy, designing sampling protocols to improve sensitivity, adopting safety protocol, and interpreting the data. The methodology was applied to a community and studied its epidemiological status with reference to occurrence, persistence, and variation of SARS-CoV-2 genome load in wastewater system to understand the prevalence of infection. Hourly and daily grab samples were analyzed and compared with the composite samples over a surveillance window of 7 days. Based on the SARS-CoV-2 RNA copies/L, faeces shedding, and volume of sewage generated the infected individuals and the population who are in active phase in the studied community was estimated.

Keywords: Domestic wastewater; Grab sampling; Composite sampling; Sampling station; Frequency

1. Introduction

The persistence and replication of SARS-CoV-2 in the gastrointestinal (GI) tract and shedding through faeces is being established as a transmission route to the environment settings, which eventually discharges to the wastewater/sewage system (Wang et al., 2020 a,b; Xiao et al., 2020b; Zhang et al., 2020; Young et al., 2020; Woelfel et al., 2020; Zahedi et al., 2021; Venkata Mohan et al., 2021). Detection of SARS-CoV-2 genetic material in the sewage/wastewater documented a significant interest among the research fraternity in the framework of wastewater based epidemiological (WBE) studies (Kumar et al., 2020; Venkata Mohan et al., 2021; Hemalatha et al., 2021). Previously, the wastewater functioned as a surveillance system for poliovirus and Aichi virus (Lodder et al., 2012; Asghar et al., 2014). WBE studies also functioned as an important supplement to clinical surveillance in polio eradication and have the potential to inform the epidemiology of COVID-19 (Xagorarakis and O'Brien. 2020; Shaw et al., 2020). Unlike polio, the COVID-19 vaccine reached the mankind, relatively in very short span of time but with some challenges such as their efficacy among different people and time it takes to reach entire population. In this scenario, the testing of the massive population to contain the spread of the virus is a challenge and therefore, an alternative strategy to assess the disease spread and thereby efficiently manage the disease is critical (Xagorarakis and O'Brien. 2020; Shaw et al., 2020; Medema et al., 2020; Venkata Mohan et al., 2021). Wastewater surveillance is an unbiased tool that helps to establish an early-warning system that would be able to monitor the occurrence, spread and, severity of the infection at a community level and therefore help in early preventative measures and allocation of resources to potentially affected areas (Bibby et., 2013 & 2015; Casanova et al., 2015; Brainard et al., 2017; Torrey et al., 2019), which eventually minimize the outbreak and spread (Daughton et al., 2018; Lednicky et al., 2020; Venkata Mohan et al., 2021). Recent reports employed WBE-based approaches to detect SARS-CoV-2 in

domestic/sewage wastewater (Wu et al., 2020; Ahmed et al., 2020a; Wurtzer et al., 2020; La Rosa et al., 2020b; Medema et al., 2020; Usman et al., 2020; Hemalatha et al., 2021). Despite the evidence on the persistence of SARS-CoV-2 RNA in wastewater/sewage, the virus transmission to the community from wastewater infrastructure is yet to be established.

The wastewater/sewage complexity, the dilute nature of biomarker in wastewater and, the inability to pinpoint specific locations are some of limitations in establishing quantitative predictions of viral RNA load WBE in (Ahmed et al., 2020b; Mao et al., 2020; Hart et al., 2020). The accuracy of the assessment of viral load among the community is an important pre-requisite of any WBE Studies (Nakamura et al., 2015; Venkata Mohan et al., 2020). Interpretation with a limited number of positive wastewater sample is challenging (Asghar et al., 2014; O'Reilly et al., 2020). Setting minimum standards for surveillance sites, developing a standardized sampling policy, creating laboratory testing protocols to improve sensitivity and minimizing the risk of cross-contamination would be necessary for an informative mode of surveillance (Wu et al., 2020; Xagorarakis and O'Brien., 2020). Several approaches for sampling were used to increase the volume of a sample which might help to identify the virus in wastewater which makes intractable to handle samples and process in laboratories (O'Reilly et al., 2020). The nature, time and frequency of sampling apart from appropriate sampling station will play a decisive role in comprehensively representing the community in order to obtain reliable data. The sampling protocol adopted can also be used to analyze the pattern of viral loads at different time-frequency and finally provide a basis for regular collection at that particular sampling station. Safety understanding in sample collection, handling and processing is needed. Therefore, in this study, a comprehensive attempt was made to establish a methodological approach to represent a community with reference to the prevalence of infection in the framework of WBE studies in terms of identifying/selection of

surveillance sites, standardizing sampling policy, designing sampling protocols to improve sensitivity, adopting safety protocol, and interpreting the data.

2. Materials and Methods

2.1. Surveillance (sampling) station

The selected sampling station receives domestic wastewater from ~1.8 lakhs people residing in the sub-urban areas including Tarnaka, HMT Nagar, Lalaguda, and Nacharam (and partly Raghavendra Nagar) in Hyderabad, Telangana (State), India. Based on the available population data, the total domestic flow per day was estimated to be 18 million litre per day (MLD). Domestic wastewater samples were collected at the converging (terminal; downstream) point of all the lateral drains that further leads to the Sewage Treatment Plant (STP; 10 MLD capacity; 17.37°N 78.48°E; Nacharam, India) (Fig 1).

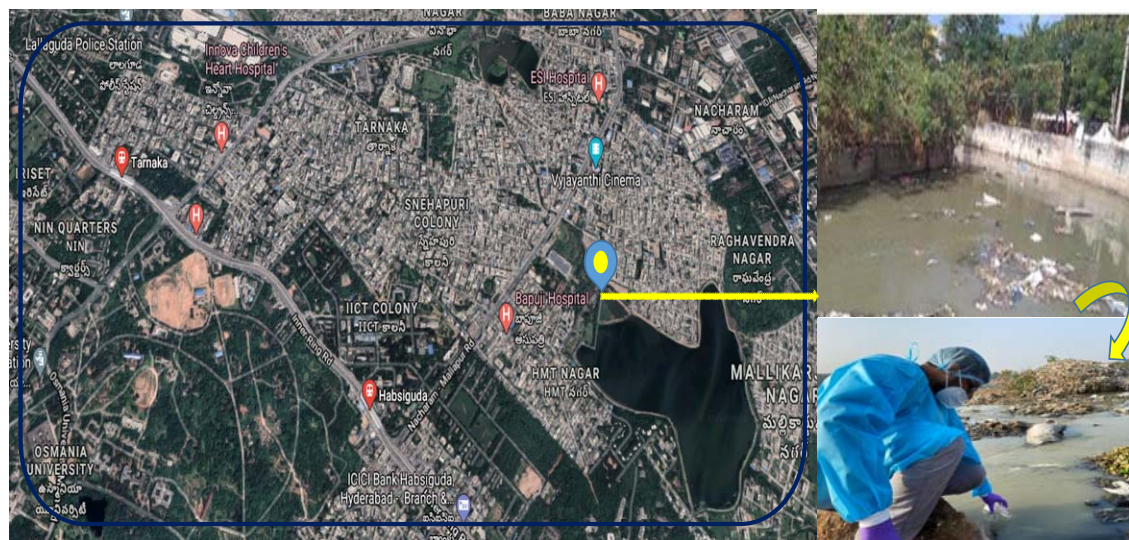


Figure 1: Sampling point selected for daily and hourly monitoring at the inlet of the STP in Nacharam, Hyderabad. (Map source: Google Map).

2.2. Sampling Protocol

2.2.1. Type of Samples

The sampling method (grab/composite samples) play a important role on WBE surveillance studies along with population demographics and local epidemiological factors (Alygizakis et al., 2020). An aggregated sample representing an entire community is more to accessible than pooled clinical samples (Murakami et al., 2020). Both grab and composite sampling of domestic wastewater was employed at the selected sampling station in a defined time frequency as detailed in Table 1. A discrete grab sampling was employed at a selected sampling site at multiple time intervals. Viral load, in general, varies over time. Therefore, a composite sample was also prepared by pooling together all grab samples obtained at various hours of the same day or daily samples obtained throughout the sampling week (Table 1). The composite sampling depicts the cumulative viral loads over the selected time-frequency (24 hours or 7 days). The two sampling approaches employed will help mathematically to evaluate the heterogeneity of viral RNA.

2.2.2. Frequency of sampling

A total number of 14 grab samples were collected for hourly monitoring starting from 5 am on 05-12-2020 to 4 am on 06-12-2020 (Table 1). A total of 7 wastewater samples was collected at 7 am from 05-12-2020 to 11-12-2020 for daily monitoring (Table 1). The time period between 5 am to 9 am represent maximum sewage flow at the selected sampling station. No rainfall was recorded during the sampling window of 7 days. The two composite samples were prepared by pooling the hourly and daily collected grab samples individually.

Table 1: Details of Sampling with reference to time of hourly and daily samples.

| Sample | Date | Time |
|-----------------------|-------------------|-------------------|
| Hourly Sample | 05-12-2020 | 5 am |
| | | 6 am |
| | | 7 am |
| | | 8 am |
| | | 9 am |
| | | 10 am |
| | | 12 pm |
| | | 2 pm |
| | | 4 pm |
| | | 6 pm |
| | | 8 pm |
| | | 11 pm |
| | | 06-12-2020 |
| | 4 am | |
| Daily Sampling | 05-12-2020 | 7 am |
| | 06-12-2020 | |
| | 07-12-2020 | |
| | 08-12-2020 | |
| | 09-12-2020 | |
| | 10-12-2020 | |
| | 11-12-2020 | |

2.2.3. Sampling Procedure & Safety Considerations

The grab samples were collected in a clean plastic container (disposable; 1.2 litres) containing 20 ml of sodium hypochlorite (0.1%) to inactivate the pathogens (Hemalatha et al., 2021). Adequate biosafety measures were taken during sampling and processing of the samples. Samples were collected by wearing PPE kit comprising of gloves, cover suite, eye safety glasses, N95 protective mask and shoes (Hemalatha et al., 2021). For sample collection, the sample container was slightly lowered in the opposite direction of flow with partial immersion. Grab sample volume of one litre was collected at one time with three replicates. Sample information was prepared in the form of field sheets (date and time) and position notified (GPS readings) and point codes, observations. After sampling, the exposed surface of the container was disinfected (isopropyl alcohol (70%)) and sealed in multi-layered

plastic covers, labelled, and transported (2-4 °C) to lab and stored at 4°C until further processing. Samples were processed within 12 h after sampling for SARS-CoV-2 detection. Prior to use, all utilities (including PPE kit) are sealed in bio-hazard bag and subjected for sterilization before disposal. The unused samples and materials were inactivated/disinfected before disposal.

2.3. Processing of Samples

All the sample processing and detection experiments were performed in a Biosafety level 2 (BSL-2) laboratories as outlined by Hemalatha et al., (2021). Initially, collected samples were filtered using 1 mm filter papers to remove the larger debris followed by secondary filtration using 0.2 µm filtration units (Nalgene® filtration system(vacuum)) to remove other fine particles and pathogens. The filtrate (60 mL) was concentrated to ~600 µl using 30 kDa Amicon® Ultra-15 (15 ml; Merck Millipore) by ultra-filtration (4 °C; 4000 rpm; 10 min). The concentrated samples (~600 µL) were aliquoted to 1.5 mL eppendorf vials and 150 µL of the sample was subjected for RNA extraction.

2.4. RNA extraction and RT-PCR

The RNA extraction was done using Viral RNA isolation kit (QIAamp, Qiagen) according to the manufacturer's protocol. Fosun COVID-19 RT-PCR Detection Kit (Shanghai Fosun Long March Medical Science Co., Ltd, China) approved by FDA (Food and Drug Administration, USA) was used to isolate SARS-CoV-2 RNA. The kit consists of primers and probes that target the envelope protein coding gene (E-gene; ROX), nucleocapsid gene (N-gene; JOE), and open reading frame1ab (ORF1ab; FAM) of SARS-CoV-2. The RT-PCR (QuantStudio™5) was performed as per manufacturer recommendations. The reaction program includes reverse transcription (50°C for 15 min) and the initial denaturation (95°C for 3 minutes) followed by 45 cycles at 95°C for 5 seconds and 60°C for 40 seconds. The signals of FAM (ORF1ab), JOE (N gene), ROX (E gene), and CY5 (Internal reference)

during the cycling stage was monitored. Positive and negative controls provided along with kit were also included in the reaction plates. All the samples were tested in triplicates. Assessment of RT-PCR kit efficiency, estimation of copy number calculation and virus recovery from sewage were performed as discussed elsewhere (Hemalatha et al., 2021). The filtration units, filter papers and sample bottles were discarded in biosafety bags followed decontamination.

2.5. Data Analysis

In order to calculate the number of RNA copies per litre of collected domestic wastewater samples, linear fit equation of the E-gene was employed by which RNA copies per litre wastewater can be calculated (Hemalatha et al., 2021).

Number of RNA copies was calculated based on the C_T values of E-gene (Equation 1).

$$\text{Log RNA copies for volume of RNA used for RT - PCR} = \frac{C_T \text{ of E gene} - 33.696}{-3.2889} \dots (1)$$

The number of infected people in the given community was calculated based on the average number of RNA copies present in the sewage as per the calculation reported by Ahmed et al., 2020a and Hellmer et al., 2014.

Method 1 (Ahmed et al., 2020a)

$$\text{No. of infected individuals} = \frac{\left(\frac{\text{RNA copies}}{\text{L water}}\right) * \left(\text{L water} \frac{\text{day}}{\text{day}}\right)}{\left(\frac{\text{g faeces}}{\text{day}}\right) * \left(\frac{\text{RNA copies}}{\text{g faeces}}\right)} \dots (2)$$

Faeces excreted/person/day = 128 g (Rose et al., 2015). One positive person sheds 10^7 RNA copies/g of faeces (maximum estimate) (Foladori et al., 2020, Bivins et al., 2020).

Method 2 (Hellmer et al., 2014)

$$\text{No. of infected individuals} = \frac{\text{No. of RNA copies per liter of waste water}}{\text{Contribution of RNA copies per person to total sewage water}} \dots\dots (3)$$

Number of RNA copies excreted per mL of faeces = 10^7 . Volume of faeces excreted = 120 mL (calculated by considering the density of human faeces is 1.07 g/mL) (Foladori et al., 2020).

Relative standard deviation (RSD) for C_T value of each gene (E-gene, N-gene, ORF1ab) was calculated using equation 4, where \bar{X} is the mean of C_T value and ‘S’ is the standard deviation.

$$\text{RSD} = 100 * S / \bar{X} \dots\dots (4)$$

Number of individuals who are in active Phase of Infection during the window period was calculated by considering the infected individuals, window period and infectious period of an infected individual.

$$\text{Individuals in active phase of infection} = \frac{\text{Infected individuals in the selected area}}{\text{Window period (35 days)/Infection period (14 days)}} \dots\dots(5)$$

3. Results and Discussion

3.1. Hourly Sampling for Detection of Viral Load in Sewage

SARS-CoV-2 genetic material was detected in all the hourly samples (n=14) with temporal variation in the viral load. Dynamic detection for viral genome was observed in the domestic

sewage indicated the spread of SARS-CoV-2 among the selected community. C_T value based on the average of each gene showed $28.62 \pm 1.34\%$ for E-Gene, $27.32 \pm 1.38\%$ for N-gene and $27.43 \pm 1.35\%$ for ORF1ab in the time frame window of 24 hours (Fig.2; Table 1). RNA copy number calculated based on E- gene (Hemalatha et al., 2021) showed a 24-hour average of 22,871 RNA copies/L in the hourly monitored samples (Table 1; Fig 2b). From 5 am to 9 am, the average RNA copies recorded was 45,456 RNA copies /L which was relatively higher compared to the average of 14 hours window (10 am to 1 am; 13,387 RNA copies/L).

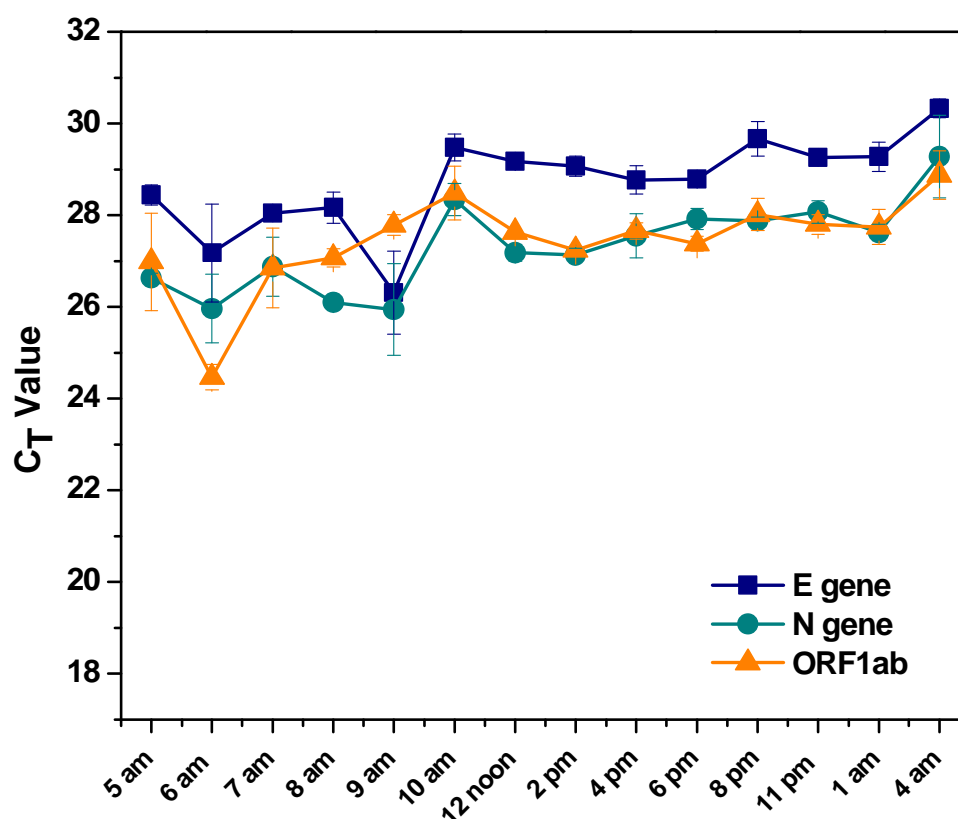


Figure 2: C_T values of E gene, N gene, and ORF1ab of samples collected hourly from selected point (All values represent, $\bar{x} \pm SD$).

Higher RNA copies were observed in the samples collected at 6 am (53531 RNA copies/L) and 9 am (98522 RNA copies/L). Between 10 am and 1 am (nearly 14 hours) the viral loads got stabilized between 9,340 and 17,556 RNA copies/L. The lowest viral load of 5,880 RNA copies/L was recorded with the sample collected at 4 am. Individual C_T values of three genes

also followed the trends with the average C_T and RNA copies. E gene C_T value varied between $26.31 \pm 5.74\%$ and $29.67 \pm 1.26\%$. Similarly, the C_T values of N gene and ORF1 ab was observed between $25.95 \pm 3.47\%$ to $28.07 \pm 0.88\%$ and $24.47 \pm 1.14\%$ to $28.48 \pm 2.06\%$ respectively. Higher viral load (genetic material) was observed in early hours i.e., 6 am to 9 am, where the peak domestic activity will happen normally during the 24 hours window. Higher detection of viral genome in morning hours might be attributed to the shedding through faeces, which represent the majority of the population in the community. Several considerations like sampling frequency, variation in sampling method, the concentration of disinfectant added, the flow rate of sewage, storage condition, downstream process, etc. effect the RNA copy numbers (Ahmad et al., 2020).

Table 1: SARS-CoV-2 RNA load with hourly domestic sewage samples

| Date | Time | E gene* | N gene* | ORF1ab* | RNA copies/L *** (using eq. 1) |
|----------------|--------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|
| 05-12-2020 | 5 am | $28.43 \pm 0.78\%$ | $26.63 \pm 0.40\%$ | $26.98 \pm 3.92\%$ | 22282 |
| | 6 am | $27.18 \pm 3.95\%$ | $25.96 \pm 2.90\%$ | $24.47 \pm 1.14\%$ | 53531 |
| | 7 am | $28.05 \pm 0.39\%$ | $26.88 \pm 2.38\%$ | $26.85 \pm 3.26\%$ | 29086 |
| | 8 am | $28.17 \pm 1.20\%$ | $26.10 \pm 0.34\%$ | $27.07 \pm 0.75\%$ | 26738 |
| | 9 am | $26.31 \pm 5.74\%$ | $25.95 \pm 3.47\%$ | $27.80 \pm 0.85\%$ | 98522 |
| | 10 am | $29.48 \pm 0.98\%$ | $28.34 \pm 1.22\%$ | $28.48 \pm 2.06\%$ | 10671 |
| | 12 noon | $29.17 \pm 0.15\%$ | $27.18 \pm 0.65\%$ | $27.63 \pm 0.10\%$ | 13262 |
| | 2 pm | $29.07 \pm 0.77\%$ | $27.13 \pm 0.55\%$ | $27.24 \pm 0.29\%$ | 14226 |
| | 4 pm | $28.77 \pm 1.08\%$ | $27.55 \pm 1.75\%$ | $27.66 \pm 0.66\%$ | 17556 |
| | 6 pm | $28.79 \pm 0.43\%$ | $27.92 \pm 0.82\%$ | $27.38 \pm 0.60\%$ | 17312 |
| | 8 pm | $29.67 \pm 1.26\%$ | $27.88 \pm 0.29\%$ | $28.02 \pm 1.25\%$ | 9340 |
| 11 pm | $29.26 \pm 0.27\%$ | $28.07 \pm 0.88\%$ | $27.80 \pm 0.79\%$ | 12451 | |
| 06-12-2020 | 1 am | $29.28 \pm 1.09\%$ | $27.62 \pm 0.64\%$ | $27.74 \pm 1.38\%$ | 12278 |
| | 4 am | $30.33 \pm 0.65\%$ | $29.28 \pm 3.09\%$ | $28.88 \pm 1.83\%$ | 5880 |
| Average | | $28.62 \pm 1.34\%$ | $27.32 \pm 1.38\%$ | $27.43 \pm 1.35\%$ | 18292 |

*Represent $\bar{X} + RSD$; ***RNA copies (based on E gene) were calculated based on the linear fit equation.

3.2. Understanding the dynamics of SARS-CoV-2 RNA load on daily basis

Based on the results obtained from the hourly monitoring, sampling time of 7 am was commonly chosen for daily monitoring to obtain the average distribution of the viral load.

SARS-CoV-2 genetic material was detected in all the seven daily monitored samples collected during the window period of one week (05-12-2020 to 11-12-2020) (Table 2; Fig 3). Detection of the viral genetic material over 7 days window period indicated the presence of SARS-CoV-2 RNA in domestic wastewater.

E-gene C_T values varied between $27.75 \pm 0.26\%$ and $31.37 \pm 2.80\%$ with an average of $29.12 \pm 0.96\%$, N-gene between $26.03 \pm 1.00\%$ and $30.35 \pm 2.26\%$ with an average of $27.74 \pm 1.60\%$ and ORF1ab between $26.49 \pm 0.21\%$ and $29.60 \pm 0.20\%$ with an average of $27.74 \pm 1.16\%$. Progressive increase in C_T values followed by decrement was observed in 7 days of sampling. The RNA copies/L ranged from 2,836 to 35,895 with an average of 17,982.

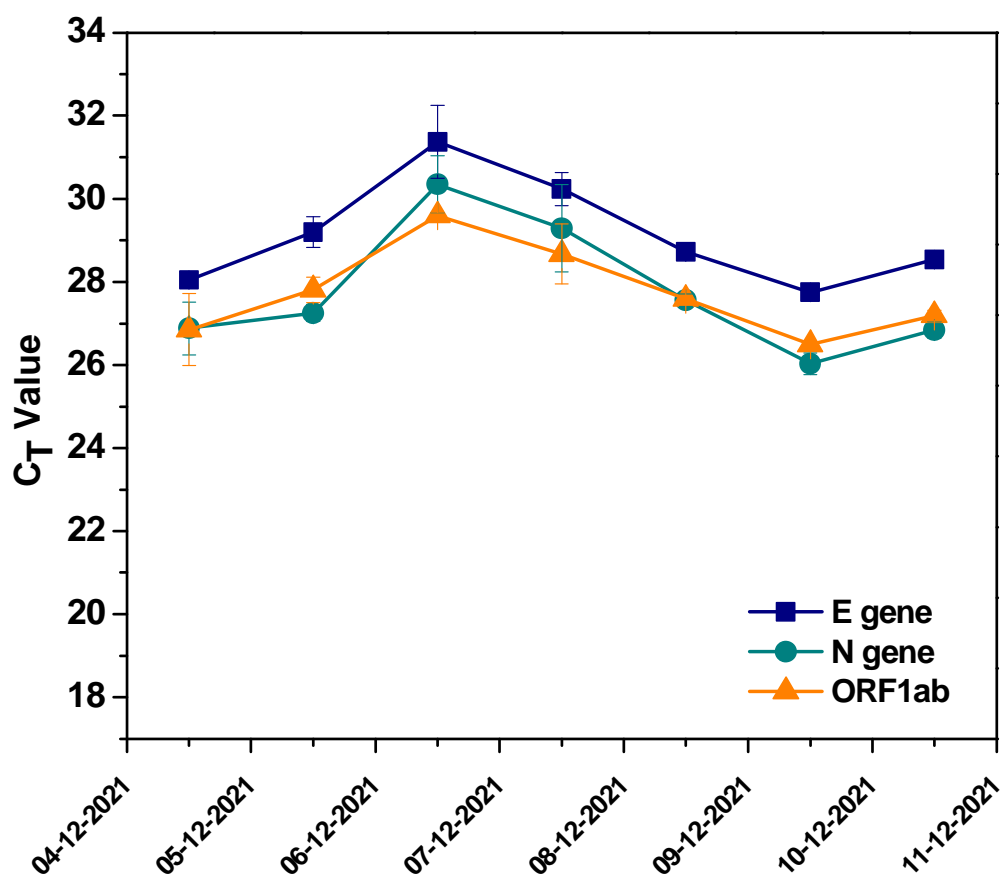


Figure 3: a) C_T values of E gene, N gene, and ORF1ab of samples collected hourly from selected point (All values represent, $\bar{x} \pm SD$).

Table 2: SARS-CoV-2 RNA C_T values detected for daily sample.

| Date | Time | E gene* | N gene* | ORF1ab* | RNA copies/L *** (using eq. 1) |
|----------------|------|--------------------|--------------------|--------------------|--------------------------------------|
| 05-12-2020 | 7am | 28.05±0.39% | 26.88±2.38% | 26.85±3.26% | 29086 |
| 06-12-2020 | 7am | 29.20±1.28% | 27.25±0.74% | 27.81±1.13% | 12986 |
| 07-12-2020 | 7am | 31.37±2.80% | 30.35±2.26% | 29.60±0.20% | 2836 |
| 08-12-2020 | 7am | 30.24±1.33% | 29.29±3.58% | 28.67±2.51% | 6263 |
| 09-12-2020 | 7am | 28.72±0.44% | 27.56±0.53% | 27.59±0.30% | 18182 |
| 10-12-2020 | 7am | 27.75±0.26% | 26.03±1.00% | 26.49±0.21% | 35895 |
| 11-12-2020 | 7am | 28.54±0.21% | 26.84±0.73% | 27.19±0.49% | 20628 |
| Average | | 29.12±0.96% | 27.74±1.60% | 27.74±1.16% | 17982 |

*Represent \bar{X} +RSD; ***RNA copies (based on E gene) were calculated based on the linear fit equation.

3.3. Composite Sample Analysis

To understand the consistency of sampling procedure two composite samples (hourly and daily) were analysed during the study period (Fig. 4; Table 3). The hourly composite sample was prepared by pooling all the hourly collected samples and similarly for daily composite samples (uniform sample volume). Hourly composite C_T of E-gene, N-gene and ORF1ab was detected to be 28.80±1.10%, 27.04±1.43% and 27.56±3.20% respectively. which is correlating well with the C_T values of the hourly sample (Table 2; Fig 4). Daily composite C_T of E-gene, N-gene and ORF1ab was detected to be 28.62±0.60%, 27.25±1.38% and 26.99±2.58% respectively. which is correlating well with the C_T values of the daily sample (Table 2; Fig 4). Quantitatively, RNA copies present in the hourly and daily composite samples was observed to be 19,503 RNA copies/L and 17,191 RNA copies/L, respectively, which correlates well with the average RNA copies of hourly and daily monitored samples 22,871 RNA copies/L and 17,982 RNA copies/L, respectively. Correlation between the composite and cumulative RNA copies represents the efficiency of sampling frequency.

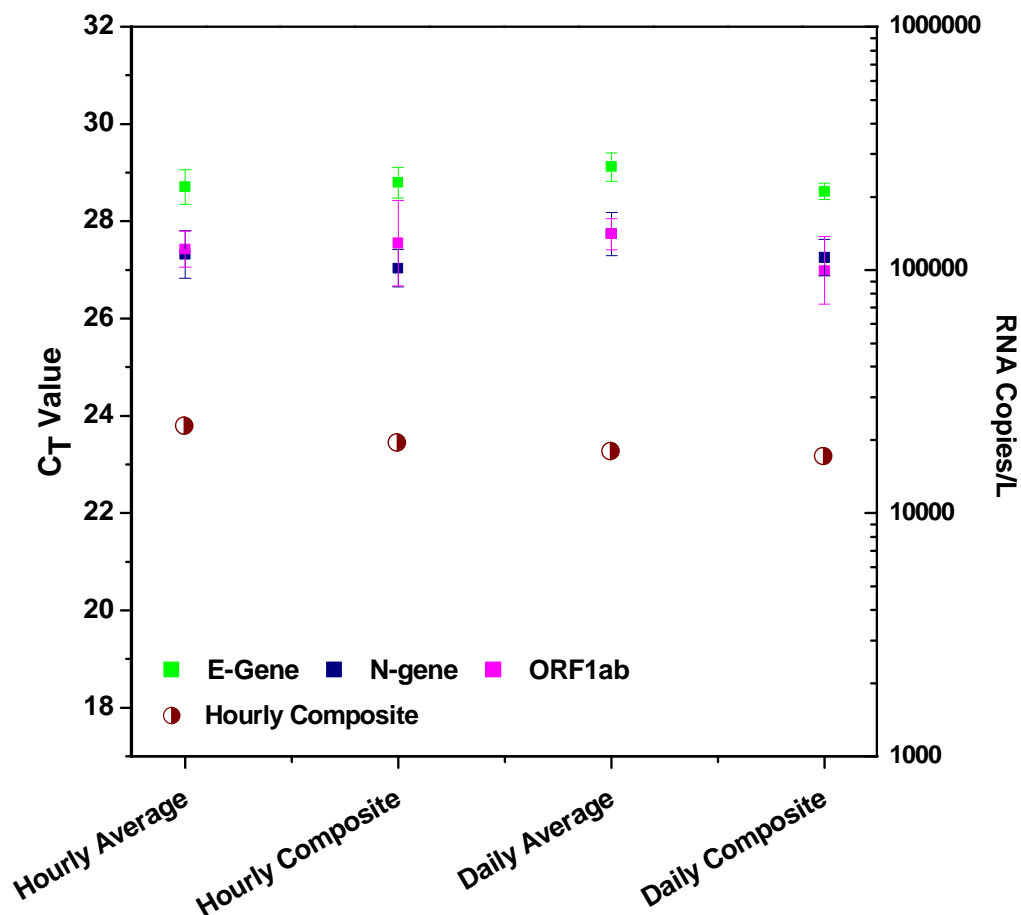


Figure 4: a) C_T values of E gene, N gene, and ORF1ab and RNA copies of hourly, daily average and composite samples (All values represent, $\bar{x} \pm SD$).

Table 3: SARS-CoV-2 RNA load with hourly and daily composite domestic sewage samples

| Sample | E gene* | N gene* | ORF1ab* | RNA copies/L *** (using Eq. 1) |
|------------------|--------------------|--------------------|--------------------|-----------------------------------|
| Hourly Average | 28.71±1.25% | 27.32±1.79% | 27.43±1.35% | 22871 |
| Hourly Composite | 28.80±1.10% | 27.04±1.43% | 27.56±3.20% | 19503 |
| Daily average | 29.12±1.00% | 27.74±1.62% | 27.74±1.15% | 17982 |
| Daily Composite | 28.62±0.60% | 27.25±1.38% | 26.99±2.58% | 17191 |
| Average | 28.81±0.99% | 27.34±1.56% | 27.43±2.07% | 19387 |

*Represent $\bar{x} \pm RSD$; ***RNA copies (based on E gene) were calculated based on the linear fit equation.

3.4. Epidemiological status- Estimation of Community Infection

To estimate the community spread of the virus, the average RNA copies of hourly average, daily average, and composites (hourly and daily) representing 7 days window of sampling period was taken into consideration along with the capacity of sewage generated in the selected community (Table 5). Both hourly and daily samples showed variability in the viral load with the time scale. The estimated number of infected individuals would include those in the early as well as later stages of infection and shedding viral particles in their faecal matter. This period was chosen based on the reports on persistence of SARS-CoV-2 RNA material for up to 47 days after the infection (Wu et al., 2020). Independent reports highlighted the replication of SARS-CoV-2 in GI tract and the prolonged shedding viral material through faeces during and after active infectious phases (Ahmed et al., 2020a; Kitajima et al., 2020; Holshue et al., 2020; Wu et al., 2020; Wurtzer et al., 2020; Cai et al., 2020; Ling et al., 2020; Woelfel et al., 2020; La Rosa et al., 2020; Xiao et al., 2020a,b). The number of infected individuals were estimated based on shedding range of 10^6 and 10^7 RNA copies/mL faeces (Foladori et al., 2020) employing the average RNA copies obtained from this study to avoid the ambiguity/discrepancy in the number of viral particles excreted by infected individuals. With reference to the studied community, the estimated infected individuals during the sampled window period is between 282 (10^7 RNA copies/mL faeces) and 2817 (10^6 RNA copies/mL faeces) and the number of individuals in the active phase of infection might be between 113 to 1127 with a wastewater flow rate of 18 MLD. Based on this number the infected individuals for the metropolitan Hyderabad city were also calculated with total sewage generation of 1800 MLD (<https://numerical.co.in/numerons/collection/5e8fd4f6f3c42b5803c09a3b>).

The number of infected individuals during the sampling window was estimated between 28,200 and 2,81,700 with a active phase of infection between 11,280 and 1,12,680. The loss of 0.02 to 3000 RNA copies/mL was reported during the transit of faeces from point of excretion to the drain (Foladori et al., 2020) which could influence the overall estimation of infected individuals. The obtained figures might include pre-symptomatic, post-symptomatic, asymptomatic, and symptomatic patients of which asymptomatic individuals might be the major contributors based on the community behaviour and serological data. WBE can quantify the scale of infection prevailing in the selected community with a benefit of detection for the individuals who have not been tested, asymptomatic, potentially symptomatic, pre-symptomatic, or only have mild symptoms (Osuolale et al., 2017; Medema et al., 2020; Lodder and de Roda Husman, 2020; Hata and Honda, 2020; Mallapaty, 2020; Naddeo and Liu, 2020; Qu et al., 2020; Venkata Mohan et al., 2021; Hemalatha et al., 2021). Asymptomatic and symptomatic cases also result in significant uncertainty in the estimated extent of SARS-CoV-2 infection (Li et al., 2020). However, this kind of estimation using the C_T value and RNA copies could help to predict the near precise number of infected individuals in a selected community/area.

Based on the outcome from this study, the following points can be considered for planning WBE (a) Sample/sampling station(s) should be a true representative of the selected community to be studied; (b) Sampling station(s) should be selected at the downstream converging point of discharge line if flowing water is being analysed. Inlet discharge point to STP can be considered for sampling which comprehensively indicates the infection of the community/area under the STP coverage; (c) Combined grab and composite sampling strategies can be adopted with a define sampling frequency; (d) Sampling frequency should be extended for not less than 24 h for hourly sampling and for not less than 7 days for daily

sampling to get a true average representation of the community. For daily sampling, the time of sample should be carefully chosen to represent average values and (e) Estimation of the viral infection of the study area can be calculated based on the population and its wastewater discharge data.

Table 4: Infection estimating through the number of people infected (symptomatic, asymptomatic, and recovered) during the sampling window on average basis of sampling

| Sample Name | Capacity of the Sewage (MLD) | 10 ⁷ RNA copies/mL faeces | | | 10 ⁶ RNA copies/mL faeces | | |
|--|------------------------------|--------------------------------------|----------|---------------|--------------------------------------|----------|-----------------|
| | | Per person contribution to Sewage | Method 1 | Method 2 | Per person contribution to Sewage | Method 1 | Method 2 |
| Hourly Average | 18 | 67 | 342 | 365 | 6.70 | 3418 | 3646 |
| Daily Average | | | 253 | 270 | | 2529 | 2697 |
| Hourly Composite | | | 274 | 293 | | 2743 | 2925 |
| Daily Composite | | | 242 | 258 | | 2417 | 2579 |
| Estimate of infected individuals (for Study Area with 18 MLD) | | | 273 | 291 | | 2726 | 2908 |
| Average estimate of infected individuals | | | | 282 | | | 2817 |
| Estimate of the population in active phase of the infection during the window period of 35 days | | | | 113 | | | 1127 |
| For 1800 MLD (on total Sewage Generation of Hyderabad city) | | 27,300 | 29,100 | | 2,72,600 | 2,90,800 | |
| Average estimate of infected individuals | | | | 28,200 | | | 2,81,700 |
| Estimate of the population in active phase of the infection during the window period of 35 days | | | | 11,280 | | | 1,12,680 |

Given the wide-spread transmission of SARS-CoV-2 it is almost impossible to test every individual. Sewage based surveillance is a holistic and effective approach to study the infection dynamics, which helps in the efficient management of the SARS-CoV-2 spread. A clear connection between the RT-PCR quantitative data and the recorded clinical data can be

correlated to reduce ambiguity viral-wastewater surveillance (O'Reilly et al., 2020; Xiao et al., 2020a).

Longitudinal sampling from the same location along with metagenomics can provide an additional illustration on the status (Alygizakis et al., 2020). Associated with clinical data, WBE could be able to provide information on SARS-CoV-2 transmission within a community including the beginning, tapering, or reemergence of an epidemic (Ahmed et al., 2020b). During the current pandemic, various mutations of SARS-CoV-2 may be evolved and WBE can also facilitate to detect the mutants by employing metagenomic analysis.

4. Conclusions

Overall, this study provides a methodological framework for WBE studies towards viral surveillance in wastewater/sewage infrastructure to precisely represent a selected community with a defined window period in terms of identifying/selection of surveillance sites, standardizing sampling policy, designing sampling protocols to improve sensitivity, adopting safety protocol, and interpreting the data. Wastewater/sewage-based surveillance can help to understand the occurrence and spread of the pandemic in the selected population or area in a more comprehensive approach by adopting a well-defined sampling protocol. Past experience with viral outbreaks there is more chances of epidemic/pandemic outbreaks in future due to the persistent anthropogenic induced ecological disturbances. In the environmental monitoring system, induction of viral/pathogen surveillance as a regular parameter along with routine monitoring of wastewater quality parameter will form an important basis to understand the early warning of the outbreaks. It is strongly recommended that healthcare/environmental agencies include WBE studies periodically to fight present and future pandemic outbreaks. Apart from an early warning signal for predicting the outbreaks, WBE also

supports clinical scrutiny along with the disease detection and management systems. The current study also offers a methodological approach to monitor other pathogens. A policy in the framework of public health by appending to the environmental systems will eventually help to safeguard the community with future outbreak.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The work was supported by Council of Scientific and Industrial Research (CSIR), New Delhi, India in the form of project entitled ‘Testing for COVID-19 in wastewater as a community surveillance measure (6/1/COVID-19/2020/IMD)’. UK thanks UGC, CGG and MH thank CSIR for the financial support received. KH, AT, MH and SVM acknowledge the Director, CSIR-IICT for the support.

References

1. Ahmed, W., Angel, N., Edson, J., Bibby, K., Bivins, A., O'Brien, J.W., Choi, P.M., Kitajima, M., Simpson, S.L., Li, J. and Tschärke, B., 2020a. First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: a proof of concept for the wastewater surveillance of COVID-19 in the community. *Science of the Total Environment*, 728, 138764.
2. Ahmed, W., Bertsch, P.M., Bivins, A., Bibby, K., Farkas, K., Gathercole, A., Haramoto, E., Gyawali, P., Korajkic, A., McMinn, B.R. and Mueller, J.F., 2020b. Comparison of virus concentration methods for the RT-qPCR-based recovery of

- murine hepatitis virus, a surrogate for SARS-CoV-2 from untreated wastewater. *Science of The Total Environment*, 739, 139960.
3. Alygizakis, N., Markou, A.N., Rousis, N.I., Galani, A., Avgeris, M., Adamopoulos, P.G., Scorilas, A., Lianidou, E.S., Paraskevis, D., Tsiodras, S. and Tsakris, A., 2020. Analytical methodologies for the detection of SARS-CoV-2 in wastewater: Protocols and future perspectives. *TrAC Trends in Analytical Chemistry*, 116125.
 4. APHA. 2017. Standard methods for the examination of water and wastewater. American Public Health Association (APHA), Washington, DC, USA.
 5. Asghar, H., Diop, O.M., Weldegebriel, G., Malik, F., Shetty, S., El Bassioni, L., Akande, A.O., Al Maamoun, E., Zaidi, S., Adeniji, A.J. and Burns, C.C., 2014. Environmental surveillance for polioviruses in the Global Polio Eradication Initiative. *The Journal of infectious diseases*, 210, S294-S303.
 6. Bibby, K. and Peccia, J., 2013a. Identification of viral pathogen diversity in sewage sludge by metagenome analysis. *Environmental science & technology*, 47(4), 1945-1951.
 7. Bibby, K., Fischer, R.J., Casson, L.W., Stachler, E., Haas, C.N. and Munster, V.J., 2015b. Persistence of Ebola virus in sterilized wastewater. *Environmental science & technology letters*, 2(9), 245-249.
 8. Brainard, J., Pond, K. and Hunter, P.R., 2017. Censored regression modeling to predict virus inactivation in wastewaters. *Environmental science & technology*, 51(3), 1795-1801.
 9. C.G. Daughton., 2018. Monitoring wastewater for assessing community health: sewage chemical-information mining (SCIM), *Sci. Total Environ.* 619–620, 748–764.
 10. Casanova, L.M. and Weaver, S.R., 2015. Inactivation of an enveloped surrogate virus in human sewage. *Environmental Science & Technology Letters*, 2(3), 76-78.

11. Foladori, P., Cutrupi, F., Segata, N., Manara, S., Pinto, F., Malpei, F., Bruni, L. and La Rosa, G., 2020. SARS-CoV-2 from faeces to wastewater treatment: what do we know? A review. *Science of the Total Environment*, 743, 140444.
12. Hamisu, A.W., Blake, I.M., Sume, G., Braka, F., Jimoh, A., Dahiru, H., Bonos, M., Dankoli, R., Bello, M., Yusuf, K.M. and Lawal, N.M., 2020. Characterizing environmental surveillance sites in Nigeria and their sensitivity to detect poliovirus and other enteroviruses. *The Journal of infectious diseases*.
13. Hart, O.E. and Halden, R.U., 2020. Computational analysis of SARS-CoV-2/COVID-19 surveillance by wastewater-based epidemiology locally and globally: Feasibility, economy, opportunities and challenges. *Science of the Total Environment*, 730, 138875.
14. Hata, A., Honda, R., 2020. Potential sensitivity of wastewater monitoring for SARS-CoV-2: comparison with norovirus cases. *Environ. Sci. Technol.* 54, 6451–6452.
15. Hellmér, M., Paxéus, N., Magnius, L., Enache, L., Arnholm, B., Johansson, A., Bergström, T., Norder, H., 2014. Detection of pathogenic viruses in sewage provided early warnings of hepatitis A virus and norovirus outbreaks. *Appl. Environ. Microbiol.* 80 (21), 6771–6781.
16. Hemalatha, M., Kiran, U., Kuncha, S.K., Kopperi, H., Gokulan, C.G., Mohan, S.V. and Mishra, R.K., 2021. Surveillance of SARS-CoV-2 spread using wastewater-based epidemiology: Comprehensive study. *Science of The Total Environment*, 768, 144704.
17. Holshue, M.L., DeBolt, C., Lindquist, S., Lofy, K.H., Wiesman, J., Bruce, H., Spitters, C., Ericson, K., Wilkerson, S., Tural, A. and Diaz, G., 2020. First case of 2019 novel coronavirus in the United States. *New England Journal of Medicine*.

18. Jiehao, C., Jin, X., Daojiong, L., Zhi, Y., Lei, X., Zhenghai, Q., Yuehua, Z., Hua, Z., Ran, J., Pengcheng, L. and Xiangshi, W., 2020. A case series of children with 2019 novel coronavirus infection: clinical and epidemiological features. *Clinical Infectious Diseases*, 71(6), 1547-1551.
19. Kitajima, M., Ahmed, W., Bibby, K., Carducci, A., Gerba, C.P., Hamilton, K.A., Haramoto, E. and Rose, J.B., 2020. SARS-CoV-2 in wastewater: State of the knowledge and research needs. *Science of The Total Environment*, 139076.
20. Kumar, M., Patel, A.K., Shah, A.V., Raval, J., Rajpara, N., Joshi, M. and Joshi, C.G., 2020. First proof of the capability of wastewater surveillance for COVID-19 in India through detection of genetic material of SARS-CoV-2. *Science of The Total Environment*, 746, 141326.
21. La Rosa, G., Iaconelli, M., Mancini, P., Ferraro, G.B., Veneri, C., Bonadonna, L., Lucentini, L., Suffredini, E. 2020. First detection of SARS-CoV-2 in untreated wastewaters in Italy. *Sci. Total Environ.* 139652.
22. Lednicky, J.A., Shankar, S.N., Elbadry, M.A., Gibson, J.C., Alam, M.M., Stephenson, C.J., Eiguren-Fernandez, A., Morris, J.G., Mavian, C.N., Salemi, M. and Clugston, J.R., 2020. Collection of SARS-CoV-2 virus from the air of a clinic within a university student health care center and analyses of the viral genomic sequence. *Aerosol and air quality research*, 20(6), 1167.
23. Ling, Y., Xu, S.B., Lin, Y.X., Tian, D., Zhu, Z.Q., Dai, F.H., Wu, F., Song, Z.G., Huang, W., Chen, J. and Hu, B.J., 2020. Persistence and clearance of viral RNA in 2019 novel coronavirus disease rehabilitation patients. *Chinese medical journal*.
24. Lodder, W.J., Buisman, A.M., Rutjes, S.A., Heijne, J.C., Teunis, P.F., de RodaHusman, A.M. 2012. Feasibility of quantitative environmental surveillance in poliovirus eradication strategies. *Appl. Environ. Microbiol.*, 78(11), 3800-3805.

25. Mallapaty, S., 2020. How sewage could reveal true scale of coronavirus outbreak. *Nature*, 580(7802), 176-177.
26. Mao, K., Zhang, H. and Yang, Z., 2020. Can a paper-based device trace COVID-19 sources with wastewater-based epidemiology? *Environ. Sci. Technol.* 2020, 54, 7, 3733–3735
27. McLellan, N., Pernitsky, D. and Umble, A., 2020. Coronavirus and the water cycle—here is what treatment professionals need to know. *Water Online* (<https://www.wateronline.com/doc/coronavirus-and-the-water-cycle-here-is-what-treatment-professionals-need-to-know-0001>).
28. Medema, G., Heijnen, L., Elsinga, G., Italiaander, R. and Brouwer, A., 2020. Presence of SARS-Coronavirus-2 RNA in sewage and correlation with reported COVID-19 prevalence in the early stage of the epidemic in the Netherlands. *Environmental Science & Technology Letters*, 7(7), .511-516.
29. Murakami, M., Hata, A., Honda, R. and Watanabe, T., 2020. Letter to the editor: wastewater-based epidemiology can overcome representativeness and stigma issues related to COVID-19. *Environmental science & technology*, 54(9), 5311-5311.
30. Nakamura, T., Hamasaki, M., Yoshitomi, H., Ishibashi, T., Yoshiyama, C., Maeda, E., Sera, N. and Yoshida, H., 2015. Environmental surveillance of poliovirus in sewage water around the introduction period for inactivated polio vaccine in Japan. *Applied and Environmental Microbiology*, 81(5), 1859-1864.
31. O'Reilly, K.M., Allen, D.J., Fine, P. and Asghar, H., 2020. The challenges of informative wastewater sampling for SARS-CoV-2 must be met: lessons from polio eradication. *The Lancet Microbe*, 1(5), 189-190.

32. Osuolale, O. and Okoh, A., 2017. Human enteric bacteria and viruses in five wastewater treatment plants in the Eastern Cape, South Africa. *Journal of infection and public health*, 10(5), 541-547.
33. Shaw, A.G., Majumdar, M., Troman, C., O'Toole, Á., Benny, B., Abraham, D., Praharaj, I., Kang, G., Sharif, S., Alam, M.M. and Shaukat, S., 2020. Rapid and sensitive direct detection and identification of poliovirus from stool and environmental surveillance samples by use of nanopore sequencing. *Journal of clinical microbiology*, 58(9).
34. Torrey, J., von Gunten, U. and Kohn, T., 2019. Differences in viral disinfection mechanisms as revealed by quantitative transfection of Echovirus 11 genomes. *Applied and environmental microbiology*, 85(14).
35. Usman, M., Farooq, M. and Hanna, K., 2020. Existence of SARS-CoV-2 in wastewater: implications for its environmental transmission in developing communities. *Environ. Sci. Technol.* 2020, 54, 13, 7758-7759.
36. Venkata Mohan, S., Hemalatha, M., Kopperi, H., Ranjith, I. and Kumar, A.K., 2021. SARS-CoV-2 in environmental perspective: Occurrence, persistence, surveillance, inactivation and challenges. *Chemical Engineering Journal*, 405, 126893.
37. Wang, J., Shen, J., Ye, D., Yan, X., Zhang, Y., Yang, W., Li, X., Wang, J., Zhang, L. and Pan, L., 2020. Disinfection technology of hospital wastes and wastewater: Suggestions for disinfection strategy during coronavirus Disease 2019 (COVID-19) pandemic in China. *Environmental pollution*, 114665.
38. Wang, W., Xu, Y., Gao, R., Lu, R., Han, K., Wu, G. and Tan, W., 2020. Detection of SARS-CoV-2 in different types of clinical specimens. *Jama*, 323(18), pp.1843-1844.
39. Woelfel, R., Corman, V.M., Guggemos, W., Seilmaier, M., Zange, S., Mueller, M.A., Niemeyer, D., Vollmar, P., Rothe, C., Hoelscher, M. and Bleicker, T., 2020. Clinical

- presentation and virological assessment of hospitalized cases of coronavirus disease 2019 in a travel-associated transmission cluster. MedRxiv.
40. Wu, F., Zhang, J., Xiao, A., Gu, X., Lee, W.L., Armas, F., Kauffman, K., Hanage, W., Matus, M., Ghaeli, N. and Endo, N., 2020. SARS-CoV-2 titers in wastewater are higher than expected from clinically confirmed cases. *Msystems*, 5(4).
41. Wu, Y., Guo, C., Tang, L., Hong, Z., Zhou, J., Dong, X., Yin, H., Xiao, Q., Tang, Y., Qu, X. and Kuang, L., 2020. Prolonged presence of SARS-CoV-2 viral RNA in faecal samples. *The lancet Gastroenterology & hepatology*, 5(5), 434-435.
42. Wurtzer, S., Marechal, V., Mouchel, J.M., Maday, Y., Teyssou, R., Richard, E., Almayrac, J.L., Moulin, L. 2020. Evaluation of lockdown impact on SARS-CoV-2 dynamics through viral genome quantification in Paris wastewaters. medRxiv. doi: [doi:https://doi.org/ 10.1101/2020.04.12.20062679](https://doi.org/10.1101/2020.04.12.20062679).
43. Xagorarakis, I. and O'Brien, E., 2020. Wastewater-based epidemiology for early detection of viral outbreaks. In *Women in water quality*. Springer, Cham. 75-97.
44. Xiao, F., Sun, J., Xu, Y., Li, F., Huang, X., Li, H., Zhao, J., Huang, J. and Zhao, J., 2020a. Infectious SARS-CoV-2 in feces of patient with severe COVID-19. *Emerging infectious diseases*, 26(8), 1920.
45. Xiao, F., Tang, M., Zheng, X., Liu, Y., Li, X. and Shan, H., 2020b. Evidence for gastrointestinal infection of SARS-CoV-2. *Gastroenterology*, 158(6), 1831-1833.
46. Young, B.E., Ong, S.W.X., Kalimuddin, S., Low, J.G., Tan, S.Y., Loh, J., Ng, O.T., Marimuthu, K., Ang, L.W., Mak, T.M. and Lau, S.K., 2020. Epidemiologic features and clinical course of patients infected with SARS-CoV-2 in Singapore. *Jama*, 323(15), 1488-1494.

47. Zahedi, A., Monis, P., Deere, D. and Ryan, U., Wastewater-based epidemiology—surveillance and early detection of waterborne pathogens with a focus on SARS-CoV-2, Cryptosporidium and Giardia. *Parasitology research*, 1-22.
48. Zhang, H., Kang, Z., Gong, H., Xu, D., Wang, J., Li, Z., Cui, X., Xiao, J., Meng, T., Zhou, W. and Liu, J., 2020. The digestive system is a potential route of 2019-nCov infection: a bioinformatics analysis based on single-cell transcriptomes. *BioRxiv*.
49. Zhou, N.A., Fagnant-Sperati, C.S., Komen, E., Mwangi, B., Mukubi, J., Nyangao, J., Hassan, J., Chepkurui, A., Maina, C., Van Zyl, W.B. and Matsapola, P.N., 2020. Feasibility of the bag-mediated filtration system for environmental surveillance of poliovirus in Kenya. *Food and environmental virology*, 12(1), 35-47.