Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Defining the methodological approach for wastewater-based epidemiological studies—Surveillance of SARS-CoV-2

Harishankar Kopperi a,c,1, Athmakuri Tharak a,1, Manupati Hemalatha a,c,1, Uday Kiran b,c,1, C.G. Gokulan b,1, Rakesh K. Mishra b,* S. Venkata Mohan a,c,**

a Bioengineering and Environmental Sciences Lab, Department of Energy and Environmental Engineering (DEEE), CSIR-Indian Institute of Chemical Technology (CSIR-IICT), Hyderabad 500007, India
b CSIR-Centre for Cellular and Molecular Biology (CSIR-CCMB), Hyderabad 500007, India
Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India

ARTICLE INFO

Article history:
Received 5 May 2021
Received in revised form 10 June 2021
Accepted 13 June 2021
Available online 17 June 2021

Keywords:
Urban area
Grab sample
Composite sample
Domestic sewage treatment plants (STP)
Sampling protocol

ABSTRACT

Since COVID-19 outbreak, wastewater-based epidemiology (WBE) studies as surveillance system is becoming an emerging interest due to its functional advantage as a tool for early warning signal and to catalyze effective disease management strategies based on the community diagnosis. An attempt was made in this study to define and establish 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. Data from hourly sampling indicated a peak in the viral RNA during the morning hours (6–9 am) when the all the domestic activities are maximum. The daily sampling and processing revealed the dynamic nature of infection spread among the population. The two sampling methods viz. grab, and composite showed a good correlation. Overall, this study establishes a structured protocol for performing WBE studies that could provide useful insights on the spread of the pandemic at a given point of time. Moreover, this framework could be extrapolated to monitor several other clinically relevant diseases. Following these guidelines, it is possible to achieve measurable and reliable SARS-CoV-2 RNA concentrations in wastewater infrastructure and therefore, provides a methodological basis for the establishment of a national surveillance system.

© 2021 Elsevier B.V. All rights reserved.

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., 2020a,b; Woelfe et al., 2020; Xiao et al., 2020; Young 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 (Venkata Mohan et al., 2021; Kumar et al., 2020; Hemalatha et al., 2021a,b). Previously, the wastewater functioned as a surveillance system for poliovirus and Aichi virus (Asghar et al., 2014; Lodder et al., 2012b). 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 (Shaw et al., 2020; Xagoraraki and O'Brien, 2020; Hemalatha et al., 2021a; Tharak et al., 2021). 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 (Medema et al., 2020; Shaw et al., 2020; Venkata Mohan et al., 2021; Xagoraraki and O'Brien, 2020; Hemalatha et al., 2021b; Tharak 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 (Torrey et al., 2019; Brainard et al., 2017; Bibby et al., 2015b; Casanova and Weaver, 2015; Bibby and Peccia, 2013a), which eventually minimize the outbreak and spread (Venkata Mohan et al., 2021; Lednicky et al., 2020; Daughton, 2018). Recent reports employed WBE-based approaches to detect SARS-CoV-2 in domestic/sewage wastewater (Ahmed et al., 2020a; Hemalatha et al., 2021a; La Rosa et al., 2020a; Medema et al., 2020; Usman et al., 2020; Wu et al., 2020a). 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 in WBE (Ahmed et al., 2020b; Hart and Halden, 2020; Tharak et al., 2021). The accuracy of the assessment of viral load among the community is an important pre-requisite of any WBE Studies (Hemalatha et al., 2021a; Nakamura et al., 2015). 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 (Xagoraraki and O'Brien, 2020; Wu et al., 2020b). 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 analyse 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).

2.2. Sampling protocol

2.2.1. Type of samples

The sampling method (grab/composite samples) play an important role in WBE surveillance studies along with population demographics and local epidemiological factors (Alygizakis et al., 2020). An aggregated sample representing an entire community is more 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 (APHA, 2014). 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 h or 7 days). The two sampling approaches employed will help mathematically to evaluate the heterogeneity of viral RNA.
Fig. 1. Sampling point selected for daily and hourly monitoring at the inlet of the STP in Nacharam, Hyderabad. Source: (Satellite map from Google Maps); (Picture Courtesy - Uday Kiran and Tharak Athmakuri)

### 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  |
|                  | 06-12-2020 | 8 pm  |
|                  |          | 11 pm |
| 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 |       |

#### 2.2.2. 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 and 9 am represent maximum sewage flow at the selected sampling station. No rainfall was recorded during the sampling window of 7 days. Two composite samples were prepared by pooling the hourly and daily grab samples.
2.2.3. Sampling procedure & safety considerations

The grab samples were collected in a clean plastic bottle (disposable; 1.2 litres) containing 20 mL of sodium hypochlorite (0.1%) to inactivate the pathogens (Hemalatha et al., 2021a,b). For sample collection, the sample container was placed slightly lower 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 noted in the field sheets (date and time) along with position (GPS readings), point codes and 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 of sampling for SARS-CoV-2 detection. Biosafety measures were taken for sample collection and processing, all the utilities (PPE kit, gloves, cover suite, eye safety glasses, N95 protective mask and shoes) (Hemalatha et al., 2021a) after use were disposed into bio safety bags followed by decontamination. The unused samples and materials were inactivated/disinfected before disposal.

2.3. Processing of samples

Sample processing and detection experiments were performed in a Biosafety level 2 (BSL-2). Collected samples were filtered using 1 mm blotting papers initially 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) (Hemalatha et al., 2021a). The concentrated samples (~600 µL) were aliquoted to 1.5 mL eppendorf vials and 150 µL of the sample was subjected for RT-PCR extraction.

2.4. RNA extraction and RT-PCR

The RNA extraction was performed 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 for RT-PCR. 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 frame 1ab (ORF1ab; FAM) of SARS-CoV-2. The RT-PCR (QuantaStudio™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 min) followed by 45 cycles at 95 °C for 5 s and 60 °C for 40 s. 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., 2021a).

2.5. Data analysis

The amplified E gene from SARS-CoV-2 RNA was cloned (KpnI and HindIII restriction sites) into pcDNA3.1 vector and quantified with dsDNA HS Assay Kit (Qubit™; Invitrogen, USA) and Qubit 4 Fluorometer (Qubit™; Invitrogen, USA). Based on the E gene and vector sequences, the number of copies per nanogram was calculated from https://www.ncbi.nlm.nih.gov/nuccore/NC_045512.2?report=fasta&from=26245&to=26472 and https://www.addgene.org/browse/sequence_vdb/2093/, respectively. The plasmid was serially diluted from a number of 9.01 log10 to 0.01 log10 copies and RT-PCR was performed. The obtained C_{T} of E gene were plotted against the log copy number from which a linear fit equation (Eq. (1)) was obtained (Hemalatha et al., 2021a) which is used to calculate the number of RNA copies in wastewater (Hemalatha et al., 2021a).

\[ \text{Log RNA copies for volume of RNA used for RT-PCR} = \frac{C_{T} \text{ of E gene} - 33.696}{-3.2839} \] (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 (Ahmed et al., 2020a; Hellmér et al., 2014).

Method 1 (Ahmed et al., 2020a)

\[ \text{No. of infected individuals} = \frac{\text{(RNA copies l water)} \times \text{(L water day)} \times \text{(g faeces day)} \times \text{(RNA copies g faeces)}}{\text{L water}} \] (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 (Hellmér et al., 2014).

\[ \text{No. of infected individuals} = \frac{\text{No. of RNA copies per liter of wastewater}}{\text{Contribution of RNA copies per person total sewage water}} \] (3)
Table 2
SARS-CoV-2 RNA load with hourly domestic sewage samples.

| Date       | Time   | E gene\(a\) | N gene\(a\) | ORF1ab\(a\) | RNA copies/L (using Eq. (1)) |
|------------|--------|-------------|-------------|-------------|-----------------------------|
| 05-12-2020 | 5 am   | 28.43 ± 0.78% | 26.63 ± 0.40% | 26.98 ± 3.92% | 22,282                      |
|            | 6 am   | 27.18 ± 3.95% | 25.96 ± 2.90% | 24.47 ± 1.14% | 53,531                      |
|            | 7 am   | 28.05 ± 0.39% | 26.88 ± 2.38% | 26.85 ± 3.26% | 29,086                      |
|            | 8 am   | 28.17 ± 1.20% | 26.10 ± 0.34% | 27.07 ± 0.75% | 26,738                      |
|            | 9 am   | 26.31 ± 5.74% | 25.95 ± 3.47% | 27.80 ± 0.85% | 98,522                      |
|            | 10 am  | 29.48 ± 0.98% | 28.34 ± 2.22% | 28.48 ± 2.06% | 10,671                      |
|            | 12 noon| 29.17 ± 0.15% | 27.18 ± 0.65% | 27.63 ± 0.10% | 13,262                      |
|            | 2 pm   | 29.07 ± 0.77% | 27.13 ± 0.55% | 27.24 ± 0.29% | 14,226                      |
|            | 4 pm   | 28.77 ± 1.08% | 27.55 ± 1.75% | 27.66 ± 0.66% | 17,556                      |
|            | 6 pm   | 28.79 ± 0.43% | 27.92 ± 0.82% | 27.38 ± 0.60% | 17,312                      |
|            | 8 pm   | 29.67 ± 1.26% | 27.88 ± 0.29% | 28.02 ± 1.25% | 9,340                       |
|            | 11 pm  | 29.26 ± 0.27% | 28.07 ± 0.88% | 27.80 ± 0.79% | 12,451                      |
| 06-12-2020 | 1 am   | 29.28 ± 1.09% | 27.62 ± 0.64% | 27.74 ± 1.38% | 12,278                      |
|            | 4 am   | 30.33 ± 0.65% | 29.28 ± 3.09% | 28.88 ± 1.83% | 5,880                       |

\(a\) Represent \(\bar{X}+\text{RSD}\).

\(b\) RNA copies (based on E gene) were calculated based on the linear fit equation.

Number of RNA copies excreted per mL of faeces \(= \times10^7\). Volume of faeces excreted \(= 120 \text{ mL} \) (calculated by considering the density of human faeces is \(1.07 \text{ g/mL}\) (Foladori et al., 2020).

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

\[
\text{RSD} = 100 \times \frac{\text{S}}{\text{X}} \tag{4}
\]

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.

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

3. Results and discussion

3.1. Hourly sampling — detection of viral load

SARS-CoV-2 RNA 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 indicating the spread of SARS-CoV-2 among the selected community. \(C_T\) value based on the average of each gene showed \(28.62 ± 1.34\%\) for E-Gene, \(27.32 ± 1.38\%\) for N-gene and \(27.43 ± 1.35\%\) for ORF1ab in the time frame of 24 h (Fig. 2a; Table 2). RNA copy number calculated based on E-gene (Hemalatha et al., 2021a) showed a 24 hour average of 22,871 RNA copies/L in the hourly monitored samples (Fig. 2b; Table 2). 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 h window (10 am to 1 am: 13,387 RNA copies/L).

Higher RNA copies were observed in the samples collected at 6 am (53,531 RNA copies/L) and 9 am (98,522 RNA copies/L). Between 10 am and 1 am (nearly 14 h) the viral loads got stabilized between 9340 and 17,556 RNA copies/L (Table 2). The lowest viral load of 5880 RNA copies/L was recorded with the sample collected at 4 am (Table 2). 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 \(26.31 ± 5.74\%\) and \(29.67 ± 1.26\%\). Similarly, the \(C_T\) values of N gene and ORF1 ab was observed between \(25.95 ± 3.47\%\) to \(28.07 ± 0.88\%\) and \(24.47 ± 1.14\%\) to \(28.48 ± 2.06\%,\) respectively (Table 2). Higher viral load was observed in early hours i.e., 6 am to 9 am, where the peak domestic activity will happen normally during the 24 h 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 factors like sampling frequency, variation in sampling method, the concentration of disinfectant, the flow rate of sewage, storage condition, and downstream process could affect the RNA copy numbers (Ahmed et al., 2020a, b).

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 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.
Fig. 2. (a) $C_T$ values of E gene, N gene, and ORF1ab; (b) RNA copies calculated based on linear fit equation of E-gene of samples collected hourly from selected point (All values represent, $\bar{X}\pm SD$).

E-gene $C_T$ values varied between 27.75 ± 0.26% and 31.37 ± 2.80% with an average of 29.12 ± 0.96%, N-gene between 26.03 ± 1.00% and 30.35 ± 2.26% with an average of 27.74 ± 1.60% and ORF1ab between 26.49 ± 0.21% and 29.60 ± 0.20% with an average of 27.74 ± 1.16% (Fig. 3a; Table 3). 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 (Fig. 3b; Table 3).

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 4). 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
Fig. 3. (a) C_T values of E gene, N gene, and ORF1ab; (b) RNA copies of hourly samples calculated based on linear fit equation of E-gene (All values represent, X±SD).

C_T values of the hourly sample (Fig. 4; Table 2). 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 (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 (Fig. 4; Table 4). Correlation between the composite and cumulative RNA copies represents the efficiency of sampling frequency. Variation in RNA copies might be due to several environmental factors at a given time such as temperature, humidity, presence of organic content, detergents, oxidizing chemicals affecting the persistence of the virus (Paul et al., 2021; Hemalatha et al., 2021a; La Rosa et al., 2020b).

3.4. Epidemiological status of the community

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 sampling period was taken into consideration along with the capacity
Table 3
SARS-CoV-2 RNA $C_T$ values detected for daily sample.

| Date      | Time | E gene | N gene | ORF1ab | RNA copies/L $^b$ (using Eq. (1)) |
|-----------|------|--------|--------|--------|----------------------------------|
| 05-12-2020 | 7am  | 28.05 ± 0.39% | 26.88 ± 2.38% | 26.85 ± 3.26% | 29,086 |
| 06-12-2020 | 7am  | 29.20 ± 1.28% | 27.25 ± 0.74% | 27.81 ± 1.13% | 12,986 |
| 07-12-2020 | 7am  | 31.37 ± 2.80% | 30.35 ± 2.26% | 29.60 ± 0.20% | 2,836  |
| 08-12-2020 | 7am  | 30.24 ± 1.33% | 29.29 ± 3.58% | 28.67 ± 2.51% | 6,263  |
| 09-12-2020 | 7am  | 28.72 ± 0.44% | 27.56 ± 0.53% | 27.59 ± 0.30% | 18,182 |
| 10-12-2020 | 7am  | 27.75 ± 0.26% | 26.03 ± 1.00% | 26.49 ± 0.21% | 35,895 |
| 11-12-2020 | 7am  | 26.54 ± 0.21% | 26.84 ± 0.73% | 27.19 ± 0.49% | 20,628 |

$^a$ Represent $\bar{X}$+RSD.

$^b$ RNA copies (based on E gene) were calculated based on the linear fit equation.

Fig. 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, $\bar{X}$+SD).

Table 4
SARS-CoV-2 RNA load with hourly and daily composite domestic sewage samples.

| Sample            | E gene | N gene | ORF1ab | RNA copies/L $^b$ (using Eq. (1)) |
|-------------------|--------|--------|--------|----------------------------------|
| Hourly composite  | 28.80 ± 1.10% | 27.04 ± 1.43% | 27.56 ± 3.20% | 19,503 |
| Daily composite   | 28.62 ± 0.60% | 27.25 ± 1.38% | 26.99 ± 2.58% | 17,191 |

$^a$ Represent $\bar{X}$+RSD.

$^b$ RNA copies (based on E gene) were calculated based on the linear fit equation.

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., 2020a,b). 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; Holshue et al., 2020; Kitajima et al., 2020; La Rosa et al., 2020a; Woelfel et al., 2020; Wurtzer et al., 2020; Wu et al., 2020b). The number of infected individuals were estimated based on
sheding 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 are between 267 ($10^6$ RNA copies/mL faeces) and 2,666 ($10^6$ RNA copies/mL faeces) and the number of individuals in the active phase of infection might be between 107 to 1,066 with a wastewater flow rate of 18 MLD (Table 5). 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 26,650 and 2,66,600 with a active phase of infection between 10,660 and 1,06,640 (Table 5). 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 (Hata and Honda, 2020; Hemalatha et al., 2021a; Medema et al., 2020; Mallapaty, 2020; Naddeo and Liu, 2020; Qu et al., 2020; Venkata Mohan et al., 2021; Lodder et al., 2012b). Asymptomatic and symptomatic cases also result in significant uncertainty in the estimated extent of SARS-CoV-2 infection. However, this kind of estimation using the Ct value and RNA copies could help to predict the near precise number of infected individuals in a selected community/area.

3.5. Benefits and challenges of WBE

WBE provides a cost-efficient way of surveillance due to representative sampling of a population that allows broad surveillance that which could detect crucial changes in the status of the infection without selection bias (Alygizakis et al., 2020). Surveillance for SARS-CoV-2 RNA in wastewater via WBE provides an information devoting to community-level screening and controlling measures, as these measurements have preceded disease cases in some cases. Detections of SARS-CoV-2 RNA employing different methods have been reported worldwide, illustrating the technical feasibility of routine monitoring (Table 6). WBE can be used as a complementary and possibly early tool to detect pathogens in a community, considering that specific and representative sampling points in areas with and without sewerage systems are chosen (Venkata Mohan et al., 2021; Hemalatha et al., 2021b). Similarly, WBE can be used as a complementary tool to help in the fight against the spread of COVID-19. WBE can be a key tool to track SARS-CoV-2 circulation in under-developed areas (Calabria de Araujo et al., 2021). It is possible to estimate the number of infected people in a specific sewage catchment area if the virus concentration in the wastewater per day and the virus concentration in the faeces of an infected person per day are known (Qu et al., 2020). WBE provides the categorization of the selected community into the zones based on the severity of viral load in the sewage (Tharak et al., 2021). WBE predicts infection dynamics including the symptomatic, asymptomatic and pre/post symptomatic characteristics (Bai et al., 2020). It is possible to predict the raise in infection rate before 2–5 weeks of actual manifestation in the population (Wu et al., 2020b).

One of the major challenges on WBE is the use of different protocols (Wu et al., 2020a). Strategies for WBE are unique and therefore is standard protocol for SARS-CoV-2 surveillance from wastewater (Calabria de Araujo et al., 2021) and to estimate the number of infected people in a community from the quantification of SARS-CoV-2 in wastewater (Wu et al., 2020a). The SARS-CoV-2 RNA concentration (copies/L) measured in the sewage may be dependent on excretion from the
Table 6
Various sampling methods adopted and detection assay for estimation of virus in sewage wastewater.

| Virus | Sampling | Detection assay and Primers | Reference |
|-------|----------|-----------------------------|-----------|
| SARS-CoV-2 | Grab and composite sampling (24 h; 1 L) Domestic wastewater (sewage) Hyderabad, India | RT-PCR ORF1ab, E and N | This study |
| | Grab Samples (1 L) STP inlet and outlet samples Hyderabad, India. | RT-PCR | Hemalatha et al. (2021a) |
| | Grab samples (100 mL) Domestic wastewater from Drains Hyderabad, India | RT-PCR ORF1ab, N and S genes | Kumar et al. (2020) |
| | Grab samples (100 mL) Domestic wastewater from Drains Ahmedabad, India. | RT-PCR | Ahmed et al. (2020a,b) |
| | Composite sample (10 h) Untreated sewage wastewater from Rio de Janeiro, Brazil | RT-PCR; CDC N2 | Prado et al. (2020) |
| | Grab samples Hospital septic tank influent Zhejiang University, China | RT-PCR; SARS-CoV-2 nucleic acid detection kit | Wang et al. (2020a);; |
| | Grab samples (2 L) Inlets and outlet of pre-processing disinfection pool of STP China | RT-PCR; CDC-CoV1, CCDC-N gene | Zhang et al. (2020) |
| | Composite sample (24 h; 500 mL) Untreated wastewater Czech Republic | RT-PCR EliGene® COVID19 BASIC A RT kit | Mlejnkova et al. (2020) |
| | Composite samples (24 h; 45 mL) Untreated and treated wastewater Federal State of North Rhine-Westphalia, Germany | RT-PCR M gene RdRP gene | Westhaus et al. (2021) |
| | Grab and composite samples (24 h; 250 mL) Untreated wastewater Italy | RT-PCR ORF1ab Spike protein RdRP | Loder and de Roda Huesman (2020a) |
| | Grab samples (200–500 mL) Influent and secondary treated wastewater Yamanashi Prefecture, Japan | RT-qPCR N_Sarbeco NIID_2019-nCOV_N CDC N1 CDC N2 | Haramoto et al. (2020) |
| | Grab sampling (200 mL) Influent, Secondary, Tertiary effluent, Region of Murcia, Spain | RT-qPCR CDC N1, N2, N3 CDC N1, N2, N3 CDC N1, N2 | Randazzo et al. (2020) |
| | Composite sampling (24 h; 250 mL) Untreated Wastewater Netherlands | RT-qPCR CDC N1, N2, N3 CDC N1, N2, N3 CDC N1, N2 | Medema et al. (2020) |
| | Composite sample (24 h; 40 mL) Untreated wastewater Massachusetts, USA | RT-qPCR CDC N1 CDC N2 CDC N3 E_Sarbeco | Wu et al. (2020b) |
| | Grab and composite samples (24 h, 100-750 mL) Untreated and Secondary treated wastewater Southern Louisiana, USA | RT-qPCR CDC N1 CDC N2 | Sherchan et al. (2020) |
| | Grab sampling (40–70 L) Untreated wastewater Southeast Michigan, USA | RT-qPCR CDC N1 | Miyani et al. (2020a,b) |
| Coronavirusidae (Virome): Alphacoronavirus; Betacoronavirus | Grab samples (1 L) Lake samples (1.0 m depth) Lake Balkhash, Kazakhstan | qPCR | Alexyuk et al. (2017) |
| Human Coronavirus 229E; Human Coronavirus HKU1 | Grab samples Sewage sludge and influent and effluent sludge U.S.A | PCR (In-silico) | Bibby and Peccia (2013a) |
| SARS-CoV-1 | Grab samples Sewage water from SARS patients Beijing, China | RT-qPCR | Wang et al. (2005c) |
infected person, immunity level, locality and age groups (Hemalatha et al., 2021a; Ahmed et al., 2020a; Foladori et al., 2020; Bivins et al., 2020; Hellmér et al., 2014). Advanced detection tools (RT-qPCR) for rapid, efficient identification may require standard laboratory equipment, being unfeasible for many countries (Medema et al., 2020; Wurtzer et al., 2020; Ahmed et al., 2020b). The persistence of virus in sewage influenced by factors such as temperature, population, drain flow rate, infected individuals, time and day of sampling, dilution effect, community socio-economic conditions, extent of sanitation and facility, humidity, rainfall event, etc (Calabria de Araujo et al., 2021). The supplemental quality control data and standardization in sampling design, sample processing, data interpretation, and reporting are mandated to reliably interpret data generated by these efforts for informing public health interventions (Calabria de Araujo et al., 2021). WBE also can detect different viral variants currently circulating in the population.

3.6. Way forward

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 defined 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. Along with this, in this study, the estimate of infected individuals was based on the Ct values of E-gene specific to SARS-CoV-2, as it was amplified and cloned in a vector as reported by Hemalatha et al. (2021a). In a similar way to get infected population one can also amplify and clone the respective gene(s) specific to SARS-CoV-2 as per the targets available in the testing kits.

WBE ability to detect low levels of viral titre especially at early stages of an outbreak or when infection levels are decreasing (following the interventions) is very important (Kitajima et al., 2020). Wastewater sampling in cities requires sewer network maps to understand the population being represented. Outside of dense populations, there are fewer sewer networks that restricts the representative wastewater sampling (Hamisu et al., 2020). Interpretation of positive wastewater samples along with clinical investigation will yield more information (Asghar et al., 2014; O’Reilly et al., 2020). WBE studies functioned as an important supplement to clinical surveillance in polio eradication (Xagoraraki and O’Brien, 2020; Shaw et al., 2020).

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 in viral-wastewater surveillance (Xiao et al., 2020; O’Reilly et al., 2020). Longitudinal sampling from the same location along with metagenomics can provide an additional illustration on the status (Torrey et al., 2019). Associated with clinical data, WBE could provide information on SARS-CoV-2 transmission within a community including the beginning, tapering, or reemergence of an epidemic (Lodder et al., 2012b; Naddeo and Liu, 2020). During the current pandemic, various mutations of SARS-CoV-2 have evolved and WBE can also facilitate to detect the mutants by employing metagenomic analysis.

4. Conclusion

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. The current study focuses on a methodological approach for sampling wastewater in order to detect SARS-CoV-2 RNA, as well as available information on sample concentration, extraction, and detection methods. Wastewater/sewage-based surveillance can help to understand the occurrence and spread of the pandemic in the selected population or area in a more compressive approach by adopting a well-defined sampling protocol. Past experience with viral outbreaks showed there is more chances of more such outbreaks in future due to the persistent anthropogenic induced ecological disturbances. In the environmental monitoring system, introduction 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. This work highlights potential areas for standardization, such as sampling time and frequency in relation to peak faecal loading times; the implementation of relevant information on sample collection type and points, sample volume collected, transport and storage conditions, extraction process and Real-Time PCR analysis. 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.
CRediT authorship contribution statement

Harishankar Kopperi: Methodology, Investigation, Writing - original draft. Athmakuri Tharak: Methodology, Investigation, Data curation, Writing - original draft. Manupati Hemalatha: Methodology, Investigation, Data curation, Writing - original draft. Uday Kiran: Methodology, Investigation, Data curation, Writing - original draft. C.G. Gokulan: Methodology, Investigation, Formal analysis, Writing - original draft. Rakesh K. Mishra: Conceptualization, Supervision, Funding acquisition, Validation, Writing - review & editing. S. Venkata Mohan: Conceptualization, Methodology, Supervision, Validation, Writing - original draft, Writing - review & editing.

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.

Acknowledgements

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, India, CGG, India and MH thank CSIR, India for the financial support received. KH, AT, MH and SVM acknowledge the Director, CSIR-IJCT, India (Manuscript No. IICT/Pubs./2019/292) for the support.

References

Ahmed, W., Angel, N., Edison, J., Bibby, K., Bivins, A., Brien, J., W.O., Choi, P.M., Kitajima, M., Simpson, S.L., Li, J., et al., 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. Sci Total Environ. 728, 138764.

Ahmed, W., Bertsch, P.M., Bivins, A., Bibby, K., Farkas, K., Gathercole, A., Haramoto, E., Gyawali, P., Korajkic, A., McMinn, B.R., 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. Sci. Total. Environ. 739, 139960.

Alexyuk, M.S., Turmagambetova, A.S., Alexyuk, P.G., Bogoyavlenskiy, A.P., Berezin, V.E., 2017. Comparative study of viromes from freshwater samples of the Ile-Balkhash region of Kazakhstan captured through metagenomic analysis. VirusDisease 28 (1), 18–25.

Alygizakis, N., Markou, A.N., Roussis, N.L., Galani, A., Avgiris, M., Adamopoulos, P.G., Sgorias, A., Lianidou, E.S., Paraskevis, D., Tsiodras, S., Tsakris, A., 2020. Analytical methodologies for the detection of SARS-CoV-2 in wastewater: Protocols and future perspectives. Trends Anal. Chem. 116125.

APHA, 2014. Standard Methods for the Examination of Water and Wastewater. American Public Health Association (APHA, Washington, DC, USA).

Calabria de Araujo, J., Gavazza, S., Leao, T.L., Florencio, L., da Silva, H.P., Albuquerque, J.D.O., de Lira Borges, M.A., de Oliveira Alves, R.B., Rodrigues, R.H.A., dos Santos, E.B., 2021. SARS-CoV-2 Sewage surveillance in low-income countries: potential and challenges. J. Water Health 19 (1), 1–19.

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., Burns, C.C., 2014. Environmental surveillance for polioviruses in the Global Polio Eradication Initiative. J. Infect. Dis. 210 (suppl. 1), 5294–5303.

Bai, Y., Yao, L., Wei, T., Tian, F., Jin, D.Y., Chen, L., Wang, M., 2020. Presumed asymptomatic carrier transmission of COVID-19. JAMA 323 (14), 1466–1470.

Bibby, K., Fischer, R.J., Casson, I.W., Stachler, E., Haas, C.N., Munster, V.J., 2015b. Persistence of Ebola virus in sterilized wastewater. Environ. Sci. Technol. Lett. 2 (9), 245–249.

Bibby, K., Peccia, J., 2013a. Identification of viral pathogen diversity in sewage sludge by metagenome analysis. Environ. Sci. Technol. 47 (4), 1945–1951.

Bivins, A., North, D., Ahmad, A., Lee, J., Been, F., Bhattacharya, P., Bijlsma, L., Boehm, A.B., Brown, J., Buttiglieri, G., 2020. Wastewater-based epidemiology: global collaborative to maximize contributions in the fight against COVID-19. Environ. Sci. Technol. 54 (13), 7754–7757.

Brainard, J., Hendricks, P.R., 2017. Censored regression modeling to predict virus inactivation in wastewaters. Environ. Sci. Technol. 51 (3), 1406–1407.

Casanova, L.M., Wever, S.R., 2015. Inactivation of an enveloped surrogate virus in human sewage. Environ. Sci. Technol. Lett. 2 (3), 76–78.

Daughton, C.G., 2018. Monitoring wastewater for assessing community health: Sewage chemical-information mining (SCIM). Sci. Total Environ. 619, 748–764.

Foladori, P., Cutrupi, F., Segata, N., Manara, S., Pinto, F., Malpei, F., Bruni, L., La Rosa, G., 2020. SARS-CoV-2 From faeces to wastewater treatment: what do we know? A review. Sci. Total Environ. 743, 140444.

Hamisu, A.W., Blake, I.M., Sume, G., Braka, F., Jimoh, A., Dahiro, H., Bonos, M., Dankoli, R., Bello, M., Yusuf, K.M., Lawal, N.M., 2020. Characterizing environmental surveillance sites in Nigeria and their sensitivity to detect poliovirus and other enteroviruses. J. Infect. Dis. http://dx.doi.org/10.1093/infdis/jiaa175.

Haramoto, E., Malla, B., Thakali, O., Kitajima, M., 2020. First environmental surveillance for the presence of SARS-CoV-2 RNA in wastewater and river water in Japan. Sci Total Environ. 731, 140405.

Hart, O.E., 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. Sci. Total Environ. 730, 138875.

Hata, A., Honda, R., 2020. Potential sensitivity of wastewater monitoring for SARS-CoV-2: comparison with norovirus cases. Environ. Sci. Technol. 54, 6451–6452.

Hellmér, M., Paxèus, 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.

Hemalatha, M., Kiran, U., Kuncha, S.K., Kopperi, H., Gokulan, C.G., Mohan, S.V., Mishra, R.K., 2021a. Surveillance of SARS-CoV-2 spread using wastewater-based epidemiology: Comprehensive study. Sci. Total Environ. 768, 144704.

Hemalatha, M., Tharak, A., Kopperi, H., Kiran, U., Gokulan, C.G., Mishra, R.K., Venkata Mohan, S., 2021b. Comprehensive and temporal surveillance of SARS-CoV-2 in urban water bodies: Early signal of second wave onset. medRxiv.

Holshue, M.L., DeBolt, C., Lindquist, S., Loft, K.H., Wieseman, J., Bruce, H., Spitters, C., Ericson, K., Wilkerson, S., Tural, A., Diaz, G., 2020. First case of 2019 novel coronavirus in the United States. New Eng. J. Med. 382, 929–936.
It is evident that wastewater-based epidemiology can provide invaluable insights into the dynamics of viral transmission, allowing for early detection and monitoring of emerging trends. This approach not only enhances our understanding of viral spread but also supports targeted interventions and public health strategies. As our technological capabilities advance, so too does our ability to harness wastewater as a surveillance tool, making it an indispensable component in the arsenal of tools available for tracking and managing viral outbreaks.
Xiao, F., Sun, J., Xu, Y., Li, F., Huang, X., Li, H., Zhao, J., Huang, J., Zhao, J., 2020. Infectious SARS-CoV-2 in feces of patient with severe COVID-19. Emerg. Infect. Dis 26 (8), 1920.

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., Lau, S.K., 2020. Epidemiologic features and clinical course of patients infected with SARS-CoV-2 in Singapore. JAMA 323 (15), 1488–1494.

Zahedi, A., Monis, P., Deere, D., Ryan, U., 2021. Wastewater-based epidemiology—surveillance and early detection of waterborne pathogens with a focus on SARS-CoV-2, Cryptosporidium and Giardia. J. Parasitol Res. 1–22.

Zhang, D., Ling, H., Huang, X., Li, J., Li, W., Yi, C., Zhang, T., Jiang, Y., He, Y., Deng, S., et al., 2020. Potential spreading risks and disinfection challenges of medical wastewater by the presence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) viral RNA in septic tanks of fangcang hospital. Sci Total Environ. 741, 140445.