WASTEWATER-BASED EPIDEMIOLOGY AS AN ALLY IN THE FIGHT AGAINST THE COVID-19: A BRIEF REVIEW

Abstract
Several researchers have detected the viral genome of SARS-CoV-2 in the stools of people (symptomatic or not) diagnosed with coronavirus disease-2019 (COVID-19). In domestic wastewater, this concentration can be between 19 and 5.5 x 10^6 viral genomes per liter of wastewater. Wastewater-based epidemiology (WBE) starts from the premise that it is possible to estimate the number of people infected by the new coronavirus within a given population group, by analyzing the amount of viral genome present in the wastewater produced by these people. This paper presents the procedures to estimate the number of COVID-19 cases from raw wastewater samples, discusses the main aspects that involve the WBE assay and presents a discussion based on the case studies applied to COVID-19 in some countries. Finally, there is a final discussion about the current situation and the main challenges for WBE to be applied efficiently for COVID-19, as an alternative to massive human testing, especially for developing and underdeveloped countries.

Keywords: novel coronavirus, SARS-CoV-2, sewage, viral genome, wastewater sampling.
Introduction

From December 2019 to November 2020, the outbreak of coronavirus disease-2019 (COVID-19) was responsible for the deaths of more than 1,400,000 people worldwide (WHO - World Health Organization, 2020). The disease is caused by SARS-CoV-2, a new coronavirus probably originated from the bat, which initially spread in China, the first epicenter of COVID-19 in the world (Zhou et al., 2020).

It is known that the viral genome (VG) of SARS-CoV-2 is present in the stool of infected people in concentrations ranging from $6.3 \times 10^5$ to $3.2 \times 10^7$ VG per mL of stool (Wölfel et al., 2020; Zhang et al., 2020a), whether the patients are symptomatic or not (Tang et al., 2020). In domestic wastewater, this concentration can be between 19 and $5.5 \times 10^6$ VG per liter of wastewater, as shown in Table 1. This elastic variation occurs both due to factors that act in the formation of wastewater (flow rate, per capita load of viral RNA excreted by infected people) and aspects related to sampling and analytical procedures for determining viral RNA. The effect of each of these variables on SARS-CoV-2 VG quantification in wastewater will be clarified throughout the text.

<table>
<thead>
<tr>
<th>Reference</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentration in wastewater samples (VG/L)</td>
<td>19 to 120</td>
<td>30 to 1700</td>
<td>$1.0 \times 10^4$ to $2.4 \times 10^5$</td>
<td>$1.5 \times 10^4$ to $1.5 \times 10^6$</td>
<td>$3.3 \times 10^4$ to $5.5 \times 10^6$</td>
<td>$6.3 \times 10^4$ to $1.3 \times 10^6$</td>
<td>$3.2 \times 10^5$ to $5.0 \times 10^6$</td>
</tr>
</tbody>
</table>

Notes: 1 Ahmed et al. (2020); 2 Nemudryi et al. (2020); 3 Wu et al. (2020a); 4 Wu et al. (2020b); 5 Wurtzer et al. (2020); 6 Randazzo et al. (2020); 7 Fongaro et al. (2020)

Although there are no evidences of the COVID-19 transmission via the fecal-oral route, the presence of new coronavirus RNA in the wastewater can be useful in the fight against this disease through wastewater-based epidemiology (WBE). The objective of this tool is to correlate the concentration of genetic material of an etiological agent in the wastewater, with cases of a certain disease that plagues a certain population group (Ahmed et al., 2020), without having to conduct mass clinical tests (Randazzo et al., 2020), which is currently unfeasible in most countries affected by the pandemic. Therefore, wastewater offers an aggregated sample of an entire community, providing information on the various moments of the pandemic, such as the onset, gradual reduction or resurgence of the disease (Bivins et al., 2020), as well as the possibility of mapping hotspots and non hotspots (Manoj, 2020).
The aim of this review paper is to discuss the main aspects that involve the WBE assay, showcase some case studies of WBE conducted around the world, and approach the main challenges for WBE to be applied to COVID-19 as an alternative to massive human testing.

*How estimate the number of infected people*

To estimate the number of people carrying the SARS-CoV-2 virus, it must first calculate the total VG load of the new coronavirus in the wastewater (in VG/day) using the product between the concentration of genetic material in the wastewater generated by the same population \((\text{VG/L}_{\text{wastewater}})\) and the flow rate \((\text{L}_{\text{wastewater}}/\text{day})\), as shown in Equation 1. Then, simply divide the total load of the VG of the new coronavirus in the wastewater (in VG/day) by the VG load produced by each carrier of COVID-19 (VG/inhabitant/day), whether symptomatic or not, according to Equation 2.

\[
\text{Total VG load in the wastewater} = \text{VG concentration in wastewater} \times \text{wastewater flow rate} \\
\text{Estimated number of infected people} = \frac{\text{Total VG load in the wastewater}}{\text{Individual VG load}}
\]

Equation (1)  
Equation (2)

To make the calculations of Equations 1 and 2 possible the interested party must have data entries. The wastewater flow rate can be obtained by the wastewater company in the region to be studied. However, it must be considered how many people contributed to that amount of wastewater generated, as not all households in the region may be connected to the wastewater network. Therefore, the total VG load in the wastewater will be relative only to the number of people who effectively contributed to the daily volume of wastewater that reached the collection point of the sample, although it is possible to estimate the actual number of people with SARS-CoV-2 in the territory studied through mathematical extrapolation. The number of people who effectively contribute to wastewater in a given territory is also very useful when you want to know the proportion of people infected in that same territory. This proportion can be obtained by dividing the estimated number of infected people (obtained from Equation 1) and the total population that contributes to wastewater in this same area of influence.

The individual VG load can vary widely among infected people (Wölfel et al., 2020), which ends up causing inaccuracy in the estimated number of people with COVID-19. While Hart and Halden (2020) recommend values between \(5.7 \times 10^7\) and \(1.1 \times 10^{10}\) viral genomes per infected person per day, considering the reality of the USA and European countries, Brazilian researchers (developing country) considered the individual VG load equal to \(10.0\) million viral genomes per infected person per day in a WBE study (INCT ETEs Sustentáveis/ANA/COPASA/SES/IGAM, 2020).

In addition to estimating the number of people infected in a given population, WBE has other benefits, such as: i) the possibility of knowing which locations in a given region have a higher incidence of coronavirus, which will serve as a basis for strategic measures to control disease to
be taken in time by public health authorities (this can be done by analyzing the wastewater from different collection systems); ii) possibility of predicting an outbreak of the disease, through periodic analysis of the presence of SARS-CoV-2 in wastewater (Mallapaty, 2020; Orive et al., 2020; Randazzo et al., 2020). Although there is no consensus on the time interval between the detection of SARS-CoV-2 in the wastewater and the emergence of new cases, some researchers were able to detect the new coronavirus in wastewater samples between 41 and 91 days before the first cases of COVID-19 in large populations were confirmed (Chavarria-Miró et al., 2020; Fongaro et al., 2020). WBE can also be very useful for wastewater samples collected in places with high flow rate and density of people and prevent an epidemic outbreak in time, such as in airports, shopping malls, large condominiums and housing estates, and also hospitals.

**Wastewater sampling for the WBE assay**

Some aspects related to sampling can influence the analysis, namely:

- **Collection point:** it is very common for samples to be collected on WWTP (Ahmed et al., 2020; Bar-or et al., 2020; Wurtzer et al., 2020), exactly where the wastewater is assembled, making it possible to evaluate a larger number of people. However, if the hydraulic retention time (HRT) in the sewer pipes is too high, the viral RNA can deteriorate along the way, which can considerably underestimate the number of people infected with the new coronavirus (Nghiem et al., 2020; Wu et al., 2020a). It is recommended that this HRT does not exceed 7 hours, which is the estimated survival time of the new coronavirus in sewage (Hart and Halden, 2020).

- **Type of sampling:** several researchers have chosen to obtain a wastewater composite sample, with collections made over 24 hours, with the justification that it provides less variability and greater representation of the results (Bar-or et al., 2020; Nemudryi et al., 2020; Randazzo et al., 2020). However, this type of sampling provides a large dilution of viral RNA, which can make it undetectable to the conventional molecular biology techniques. There is no study proving which hours in a day are the most suitable for collecting wastewater samples with detectable concentrations of SARS-CoV-2. However, grab sampling, collected at peak times of using the toilet to evacuate feces (between 6:00 and 9:00 am, according to Heaton et al. (1992)), make the sample more concentrated with a viral genome and more likely to be detected (Chavarria-Miró et al., 2020), this being the most recommended alternative when there is no possibility to determine the most suitable time for collect samples from sanitary wastewater. Some authors also choose to collect grab samples (Haramoto et al., 2020; Hata et al., 2020).

- **Interval between collections:** the objective is that the sampling takes place in the shortest interval possible (≤ 5 days), seeking to have greater precision in the analysis and to ensure that eventual changes in the concentration of the viral genome are detected in time. Wu et al. (2020b) proved the time lag between the wastewater signal and clinically reported cases varies between 4 and 5 days. Wurtzer et al. (2020) applied the WBE to an area of the city of Paris, sampling every 2-3 days in the final stage of the analysis, and it was
evident a greater precision of the correlation, compared to the initial period of the experiment, when the sampling was done in an interval between 7 to 20 days.

- Underestimation of the number of infected patients from hospital samples: as mentioned earlier, monitoring of wastewater generated in hospitals can prevent an eventual outbreak of COVID-19 in these institutions. However, many patients with COVID-19 (especially intubated patients) need to use diapers, preventing this viral load from being detected in samples of the generated wastewater, which would negatively affect the efficiency of this forecast.

- Wastewater dilution: the infiltration water into the wastewater network and intrusion of water from rain (Chavarria-Miró et al., 2020) can substantially reduce the concentration of the viral genome and, thus, underestimate the number of infected individuals. In the case of infiltration waters, for example, a 50 km long drainage network that conducts 200 L/s of wastewater and which has an infiltration rate of 0.5 L/s.km, can have the concentration of VG reduced by at least 11% until arrival at the WWTP. Still on this issue, Hata et al. (2020) suggest that smaller populations tend to have less dilution of wastewater (probably due to the lower leakage rate), making detection more sensitive.

- Sample volume and concentration: the volume of the sample is another critical factor as it could lead to false negative results along with low recovery. Nevertheless, increased volumes can also affect the following steps of analysis. Alygizakis et al. (2020) have reviewed the analysis volume used in several studies and showed that it varied from 45 mL to 5,000 mL depending on the sampling and following concentration procedures. The concentration protocols are another important issue, which influences recovery efficiencies of SARS-CoV-2 and can lead to large magnitude errors in the estimation of virus load in wastewater samples (Ahmed et al., 2020). Those authors evaluated the adsorption-extraction method with acidification and polyethylene glyco (PEG) precipitation as having the worst recovery efficiency. Particularly, PEG is related to the co-concentration of PCR inhibitors (e.g. bile salts and lipids found in fecal material) (Schrader et al., 2012). Adsorption-extraction with MgCl₂ pre-treatment or without pre-treatment is the most efficient method for concentration (Ahmed et al., 2020).

- Sample storage: most WBE approaches use -20°C as the preferred storage temperature. However, this temperature may not be suitable for SARS-CoV-2 genetic material conservation. Also, freezing and thawing of the samples can lead to degradation of the genetic material by RNases present in wastewater samples and result in loss of viral load and errors in the results. Independently of the methods used, it is preferable to store the concentrated samples, not raw wastewater, at -80°C (Alygizakis et al., 2020). Another procedure that can affect the recovery of SARS-CoV-2 from wastewater is the pre-centrifugation to remove particles and debris as the generation of the pellet can absorb part of the genetic material (Ahmed et al., 2020).
Processing and detection of SARS-CoV-2 in wastewaters

The choice of appropriate analytical protocols is a critical step that can drastically affect the results. The RNA isolation is usually performed by using commercially available kits that can result in differences concerning RNA quality. For this reason, it is recommended the use of external controls, such as *Pseudomonas* bacteriophage φ6 and Hep G Armored RNA, to check on RNA extraction efficiency and PCR inhibitors (Rimoldi *et al.*, 2020). In addition, a non-target RNA sequence needs to be added in all test samples as a spiked-in control after the concentration or addition of lysis solution to estimate sample to sample variation in the viral RNA extraction (Medema *et al.*, 2020).

Concerning detection, several specific assays are in use, with sensitivity and specificity quite diverse due to the use of different gene targets and probes. The protocol suggested by the Center of Disease Control and Prevention (CDC) is the most used. However, the low viral loads contained in wastewater combined with the potential RNA degradation may limit its performance. Besides the Envelope protein gene (E-gene) and nucleocapsid-gene (N-gene) firstly used, RdRP and OFR1ab are also used as targets. Such variations of targeting sequences may explain the range of results (Alygizakis *et al.*, 2020). In addition, as SARS-CoV-2 genetic material is degradable in wastewater samples, it is recommended that more than one gene target is used to limit false positive and negative results. It is also important to highlight that there are many RT-qPCR kits routinely used for COVID-19 diagnosis. Therefore, it is strongly recommended an evaluation of their performances with wastewater (Vogels *et al.*, 2020). Some ultrasensitive alternatives, such as Nested PCR and droplet digital PCR, were proven to be applicable methods for WBE, but can be more expensive. Other options in study are sequencing, loop-mediated isothermal amplification (LAMP) and biosensors, with LAMP and biosensors being the most promising ones due to sensitivity (LAMP’s limit of detection is 12 copies per reaction) and rapidity (Alygizakis *et al.*, 2020).

Case studies of WBE applied to COVID-19 in some countries

A WBE has already been efficiently applied to diseases such as polio, hepatitis A and norovirus (Asghar *et al.*, 2014; Hellmér *et al.*, 2014). More recently, WBE was applied to COVID-19 by several researches (Bar-or *et al.*, 2020; Randazzo *et al.*, 2020; Wu *et al.*, 2020b; Wurtzer *et al.*, 2020). Bar-or *et al.* (2020) correlated the cycle threshold (Ct) with the number of COVID-19 cases in Bnei Brak City, Israel. Although these authors were able to find a direct relationship between these two variables, the number of samples analyzed was low (3 samples in just 2 days of collection), making it impossible to analyze the data in question more precisely. Wurtzer *et al.* (2020) evaluated the VG concentration of SARS-CoV-2 in samples of raw wastewater from the three largest WWTPs in the city of Paris between March 5 and April 23, and observed a proportionality profile between the concentration of VG in the samples and the estimated number of the virus excreted by the population.
Wurtzer et al. (2020) also compared the concentration of VG with epidemiological indicators of COVID-19 for the same population that generated the wastewater that reached these WWTPs. These indicators were the daily number of consultations of people with symptoms of COVID-19, the daily growth of patients hospitalized with the disease, and the total number of patients hospitalized and deceased due to the new coronavirus. These researchers (Wurtzer et al., 2020) found out that when the number of registered cases of the disease decreased, the reduction in the concentration of VG in the wastewater only began to occur 13 days after, suggesting that patients continued to excrete the virus, even a few days after being healed. Some researches have showed that patients can continue to excrete the virus through their stools for 10 - 33 days (Wu et al., 2020c; Zhang et al., 2020b), even after the results of clinical respiratory tract tests have been negative. This comparison (number of registered cases x concentration of VG in wastewater) was only possible because there was a prolonged monitoring of the concentration of viral RNA content in the wastewater. In contrast, Randazzo et al. (2020) interrupted the analysis of the less than 15 days after the beginning of the stabilization of the number of cases of the disease, and were not able to observe the change in the concentration of viral RNA in the wastewater samples produced by the main cities of the Region of Murcia (Spain).

In Brazil, a project coordinated and executed by the National Institute of Science and Technology - Sustainable Wastewater Treatment Plants of the Federal University of Minas Gerais (UFMG), undertook actions to detect and quantify the new coronavirus in wastewater samples in Belo Horizonte and Contagem, cities that make up the third largest metropolitan area in the country (Chernicharo et al., 2020). With wastewater samples collected over 5 consecutive weeks (between May 11 and June 12, 2020), the WBE study allowed the generation of regionalized trend curves for COVID-19 circulation. As part of the results, it is observed that the estimate of the infected population based on the monitoring of wastewater allowed greater detail in the trend of increasing cases in the period of the study, in contrast to the trend obtained from the number of confirmed cases. The authors also concluded that the estimated number of infected people ranged from 2 to 20 times more than the number of cases registered by the state's public health agency. This overestimation of infected people can be attributed to the low number of clinical tests performed in the studied area (only 0.5% of the total population in the region was tested) and the viral load from asymptomatic cases of COVID-19.

Also in Brazil, the Oswaldo Cruz Foundation (Fiocruz) in partnership with the city of Niterói, located in the metropolitan area of the capital city Rio de Janeiro, developed an online platform that allows monitoring the dissemination of the new coronavirus from the analysis of samples from the sewerage system. With collections started in April 2020 and carried out weekly, 29 georeferenced points were selected for the survey of material, in a staggered way and distributed among wastewater treatment stations, points for hospital wastewater disposal and wastewater collection network.
Data made available at the platform indicate that the average of positive samples of SARS-CoV-2 was 81% at the end of the epidemiological week 32 and that fragments of the virus were detected in all monitored regions in the city. Bearing in mind that Niterói has more than 95% of its population with access to the wastewater network, with an exception in Brazil (which has an average of only 53%), the current research challenge, which is expected to last 12 months, it is to establish a model that allows to contribute in the promotion of public actions in the face of the pandemic. (https://arcg.is/0HXfXX).

To our knowledge, the most accurate and mathematically adjusted WBE study to date has been conducted in Massachusetts (USA) by Wu et al. (2020b). These authors were able to observe a very high correlation between the concentration of VG in the wastewater and the number of new registered cases of COVID-19, reaching Pearson's correlation coefficient (r) equal to 0.92 between these two variables.

In view of the complexity in establishing WBE validation methods for monitoring SARS-CoV-2, scientific collaboration networks have been built since the beginning of the pandemic. The "COVID-19 WBE Collaborative" was one of the first to be developed, seeking to coordinate methodologies and data sharing to maximize WBE studies by meeting working groups on the website www.covid19wbec.org, in addition to creating a platform sharing research protocols through the Protocols.io system. Other networks at the regional level also stand out, such as "The ColoSSoS Project", which aims to undertake a research network in Australia to understand the impacts of COVID-19 on the wastewater system, as well as the "COVID-19 Wastewater Coalition", created by the Canadian Water Network (CWN) with the aim of gathering studies to assess the potential of the WBE as a national surveillance program.

In economic terms, the use of WBE proves to be extremely viable. In a simulation made for the reality of Germany, Hart and Halden (2020) estimated that the costs with inputs needed for health surveillance of COVID-19 through wastewater, would be equivalent to only 0.014% of the total amount spent with mass testing for the entire population of the country, which represents a significant economy, which would significantly reduce the costs of the government to face this pandemic.

**Final discussion**

The researches that aim to quantify the new coronavirus in the wastewater is still in the first stages, and face some limitations of analytical nature, namely: lack of standardization of sampling and viral genome concentration techniques, that is necessary in cases where there is a low concentration of viral genome in the sample, and high impurity content, as is the case with wastewater samples (Ahmed et al., 2020; La Rosa et al., 2020; Medema et al., 2020); and standardization of the viral structure region (envelope protein [E], nucleocapsid [N] gene or RNA-
dependent RNA polymerase [RdRp] gene) or primer/probe mixes (N1, N2 or N3, for example) most suitable for extraction of genetic material (Ahmed et al., 2020; Medema et al., 2020; Randazzo et al., 2020; Rimoldi et al., 2020; Wu et al., 2020a). These factors can influence the analytical sensitivity and the limit of quantification of VG in the sample (Randazzo et al., 2020).

Regarding WBE, the studies applied to monitor COVID-19 and other diseases so far make it clear that this is a promising tool from when it comes to controlling viral epidemics. However, based on the results presented by Wu et al. (2020b), Wurtzer et al. (2020) and Randazzo et al. (2020), the monitoring of VG concentration in the wastewater of a population needs to be done periodically, and for enough time so that a reduction in the VG concentration in the wastewater can be observed at values similar to those found at the beginning of the disease outbreak.

Although WBE is seen as a tool aimed at reducing costs in monitoring the spread of the new coronavirus, it is evident that most of the work related to the detection of COVID-19 in wastewater is carried out in higher-income regions, such as Europe, the United States, Asia and Oceania, highlighting the inequality in R&D in the area (Daughton, 2020). Even in those countries with potential for investment in the area, it is noted that despite the potential of the WBE in viral monitoring, it has not yet been widely adopted, partly due to problems related to the disarticulation of public health management with the management of wastewater, which is usually found in different administrative spheres and without articulation (Hrudey et al., 2020). Another factor pointed out is the fact that, given that the data coming from the WBE occur on a community and regional scale, governments fear losses in the economic field due to measures of social isolation.

Acknowledgements

The authors thank the Institute of Marine Science of the Federal University of Ceará, Brazil.

References


