



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.



Review

Correlation between SARS-CoV-2 RNA concentration in wastewater and COVID-19 cases in community: A systematic review and meta-analysis

Xuan Li^{a,b}, Shuxin Zhang^a, Samendra Sherchan^c, Gorka Orive^{d,e}, Unax Lertxundi^f, Eiji Haramoto^g, Ryo Honda^h, Manish Kumarⁱ, Sudipti Arora^j, Masaaki Kitajima^k, Guangming Jiang^{a,l,*}

^a School of Civil, Mining and Environmental Engineering, University of Wollongong, Wollongong, Australia

^b Centre for Technology in Water and Wastewater, School of Civil and Environmental Engineering, University of Technology Sydney, Ultimo, NSW 2007, Australia

^c Department of Environmental Health Sciences, Tulane University, New Orleans, LA 70112, USA

^d NanoBioCel Group, Laboratory of Pharmaceuticals, School of Pharmacy, University of the Basque Country UPV/EHU, Paseo de la Universidad 7, Vitoria-Gasteiz 01006, Spain

^e Biomedical Research Networking Centre in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Vitoria-Gasteiz, Spain

^f Bioaraba Health Research Institute; Osakidetza Basque Health Service, Araba Mental Health Network, Araba Psychiatric Hospital, Pharmacy Service, Vitoria-Gasteiz, Spain

^g Interdisciplinary Center for River Basin Environment, University of Yamanashi, Kofu, Japan

^h Faculty of Geosciences and Civil Engineering, Kanazawa University, Kanazawa, Japan

ⁱ Sustainability Cluster, School of Engineering, University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India

^j Dr. B. Lal Institute of Biotechnology, Jaipur, India

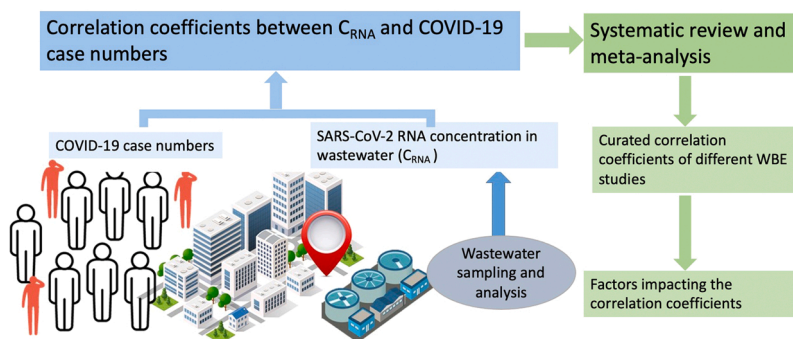
^k Division of Environmental Engineering, Faculty of Engineering, Hokkaido University, Hokkaido, Japan

^l Illawarra Health and Medical Research Institute (IHMRI), University of Wollongong, Wollongong, Australia

HIGHLIGHTS

- A systematic review of WBE performance in estimating COVID prevalence or incidence.
- Differentiating the outbreak stage can improve the correlation between COVID-19 cases and C_{RNA} .
- Large temperature fluctuation and catchment size correlated with reduced WBE performance.
- Sampling frequency and epidemiological factors are essential to improve WBE estimation.
- WBE studies should report a minimum set of data for cross-study comparability.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Dr. Danmeng Shuai

Keywords:
COVID-19

ABSTRACT

Wastewater-based epidemiology (WBE) has been considered as a promising approach for population-wide surveillance of coronavirus disease 2019 (COVID-19). Many studies have successfully quantified severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA concentration in wastewater (C_{RNA}). However, the correlation between the C_{RNA} and the COVID-19 clinically confirmed cases in the corresponding wastewater

* Corresponding author at: School of Civil, Mining and Environmental Engineering, University of Wollongong, Wollongong, Australia.

E-mail addresses: gjiang@uow.edu.au, guangming.jiang@gmail.com (G. Jiang).

<https://doi.org/10.1016/j.jhazmat.2022.129848>

Received 26 April 2022; Received in revised form 22 August 2022; Accepted 23 August 2022

Available online 27 August 2022

0304-3894/© 2022 Elsevier B.V. All rights reserved.

SARS-CoV-2
Wastewater-based epidemiology
Prevalence
Incidence, clinical testing
Viral shedding

catchments varies and the impacts of environmental and other factors remain unclear. A systematic review and meta-analysis were conducted to identify the correlation between C_{RNA} and various types of clinically confirmed case numbers, including prevalence and incidence rates. The impacts of environmental factors, WBE sampling design, and epidemiological conditions on the correlation were assessed for the same datasets. The systematic review identified 133 correlation coefficients, ranging from -0.38 to 0.99. The correlation between C_{RNA} and new cases (either daily new, weekly new, or future cases) was stronger than that of active cases and cumulative cases. These correlation coefficients were potentially affected by environmental and epidemiological conditions and WBE sampling design. Larger variations of air temperature and clinical testing coverage, and the increase of catchment size showed strong negative impacts on the correlation between C_{RNA} and COVID-19 case numbers. Interestingly, the sampling technique had negligible impact although increasing the sampling frequency improved the correlation. These findings highlight the importance of viral shedding dynamics, in-sewer decay, WBE sampling design and clinical testing on the accurate back-estimation of COVID-19 case numbers through the WBE approach.

1. Introduction

During the coronavirus disease 2019 (COVID-19) pandemic, a significant amount of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causing agent of COVID-19, is shed into sewers from feces, urine, sputum and other potential sources (i.e. blood, saliva, etc.) (Li et al., 2021b; Pan et al., 2020; van Doorn et al., 2020). The viral shedding allows the application of wastewater-based epidemiology (WBE) in estimating the community prevalence or incidence of COVID-19 as a complementary approach to clinical testing (Huang et al., 2021; Nemudryi et al., 2020; Róka et al., 2021). By systematic collection and analysis of wastewater samples for the virus, either at the inlet of wastewater treatment plants (WWTPs), sewer pumping stations, or manholes, WBE provides an estimation of COVID-19 case numbers in the connected population (Betancourt et al., 2021; Rusiñol et al., 2021; Wong et al., 2021).

The concept of WBE for the estimation of COVID-19 case number is appealing because it is believed that regardless of the status of the infection (symptomatic or asymptomatic) nor clinical testing (tested or not), as long as they shed viruses, the entire infected sub-population can be captured by WBE. However, to date, the shedding sources and their loads from patients remain unclear especially for different infectious statuses within a large population (Jones et al., 2020). Thus, the SARS-CoV-2 shedders (patients) cannot be estimated using only SARS-CoV-2 RNA concentration in wastewater (C_{RNA}) but the changes in the C_{RNA} are expected to reflect the changes in the number of patients in the catchment area (Jafferli et al., 2020; Wu et al., 2021). Thus, the performance of WBE can be assessed based on the correlation between SARS-CoV-2 RNA concentration in wastewater (C_{RNA}) and clinically confirmed cases. To date, various studies have observed the correlation between C_{RNA} and the clinically confirmed COVID-19 cases within a community (Nemudryi et al., 2020; Róka et al., 2021). However, the various COVID-19 prevalence and incidence rates, i.e. active cases, daily new cases, and future cases etc., that C_{RNA} correlated with, and the strength of the correlation varied greatly in different studies. For instance, Huang et al. (2021) found a strong correlation (Pearson's correlation coefficient, $R = 0.98$) between the C_{RNA} and active cases. D'Aoust et al. (2021b) observed that C_{RNA} correlated better with daily new cases ($R = 0.37$) than active cases ($R = 0.21$), although both correlations are weak. These conflicting observations in the literature raise two questions: 1) which epidemiological parameter can C_{RNA} correlate with better, and thus provide a good back-estimation through the WBE approach? 2) what are the potential factors that affect the correlation between C_{RNA} and COVID-19 case number in the community?

To date, as the 'true' infection cannot be quantified in the community, all the studies correlated C_{RNA} with the clinically confirmed cases. However, clinical testing is known to only capture a part of the infections (Reese et al., 2020), which makes the clinical testing coverage critical for assessing the correlations between C_{RNA} and case numbers. Furthermore, SARS-CoV-2 RNA decay during in-sewer transportation, wastewater sampling technique, and the sampling frequency have also

led to variations of C_{RNA} detected in wastewater samples (Ahmed et al., 2020a; Jafferli et al., 2020; Li et al., 2021a; Rusiñol et al., 2020; Wu et al., 2021). To date, the understanding of the impact of these factors on the correlations between C_{RNA} and COVID-19 case numbers is limited.

This paper summarizes WBE studies of COVID-19 by collating literature data of the correlation between C_{RNA} and COVID-19 case numbers with different environmental factors, wastewater sampling design, and epidemiological parameters through a systematic literature review. The contributions of environmental factors (i.e. variation of temperature and catchment size), WBE sampling design (sampling technique and frequency), and epidemiological conditions (i.e. prevalence levels and clinical testing coverage) in different countries on the correlations between C_{RNA} and clinically confirmed COVID-19 case numbers were further assessed through a meta-level analysis. The results provide a comprehensive assessment of the accuracy of WBE in estimating infection numbers in the current or future pandemics, and identify potential contributing factors to enhance the conventional WBE back-estimation.

2. Methods

2.1. Systematic review

The systematic literature search was conducted on November 30th, 2021 following PRISMA guidelines (Silverman and Boehm, 2020). The literature search collected a comprehensive set of WBE data regarding the correlation (Pearson or Spearman correlation coefficient, R) between the C_{RNA} detected in wastewater and the clinically confirmed COVID-19 cases. The clinically confirmed case numbers include active cases, daily new cases, weekly new cases (or seven-day rolling averages), cumulative cases, and future cases (upcoming new cases from the following days of the sampling date). Databases (i.e., Web of Science core collection, Scopus, and PubMed) were searched using the term "SARS-CoV-2 AND wastewater AND prevalence OR incidence". The incidence or prevalence of infection refers to the number of persons in a population who become ill (incidence) or are ill at a given time (prevalence) (Dicker et al., 2006). This study uses incidence or prevalence rates to indicate the number of new or active cases per unit of population (cases/100,000 people).

A total of 749 unique papers were identified after removing duplicates using the EndNote Reference Manager software. Titles and abstracts of the retained articles were screened and assessed for eligibility using these criteria: 1) reported data of both C_{RNA} and clinically confirmed COVID-19 prevalence or incidence in the catchment area or reported the correlation between C_{RNA} and COVID-19 prevalence or incidence; 2) the article is in English and is peer-reviewed. Articles that passed the initial screening were further assessed by full-text reading and finally, 27 articles were included in this study. If data were available in graphs, then GetData Graph Digitizer was used to digitize the C_{RNA} with corresponding prevalence in the catchment area. Details of the systematic review process are provided in the [supplementary information](#) (SI).

2.2. Collection and meta-analysis of wastewater surveillance data

Previous studies found that the decay of SARS-CoV-2 RNA in wastewater was dependent on the wastewater temperature (T_w) and time (Bivins et al., 2020a). The wastewater temperature and hydraulic retention time (HRT) of the catchment were not reported in all these articles. However, Hart and Halden (2020a) observed a close relation between T_w and the air temperature (T_a) and reached a good agreement with empirical observations. To reflect the impacts of T_w , the average T_a of each sampling day was collected from Google weather data. A previous study modeled the in-sewer transportation time (HRT) of different catchment areas and found that HRT strongly correlated to the catchment size of a WWTP, ranging from several minutes to 6–10 h in small and large scale WWTPs, respectively (McCall et al., 2017). Thus, the population size that the catchment (either the treatment plant or the sewer shed) served was collected from each study and used to reflect the catchment size and its HRT.

The sampling technique and frequency also played an important role in the prevalence estimation of COVID-19 (Gerrity et al., 2021a; Li et al., 2021a). Thus, the sampling technique applied in each study was summarized as a category variable to indicate grab sampling and composite sampling as – 1 and 1, respectively. Sampling frequency was determined as the total number of samples divided by total weeks of wastewater surveillance, i.e. weekly sample numbers, in each study. The clinical test rates (tests/1000 people over 30 days) and test coverages (tests/confirmed cases) are commonly used to evaluate whether the clinical testing is adequate (Oranization, 2020). The clinical testing rates and testing practice in the catchment area were not reported in all these articles, but they are publicly available for each country, which can approximately reflect the testing resource availability in the study areas. Thus, the clinical test rates (tests/1000 people over 30 days) and test coverages (tests/confirmed cases) of the specific country in each study during the wastewater surveillance period were sourced from the database established by Hasell et al. (2020).

The theoretically ideal condition for evaluating the correlation between C_{RNA} and COVID-19 infection cases is that all other factors (i.e. temperature, clinical testing conditions, etc.) remain the same during the study period. Considering the durations (i.e. months to years) of WBE studies and dynamic environmental and clinical testing conditions, the variations of temperature clinical test rates, and testing coverage were further calculated based on the maximum value minus the minimum value for each factor to evaluate their impacts on correlation between C_{RNA} and COVID-19 cases. To assess the impacts of catchment size, sampling frequency, and variations of environmental and clinical testing factors (i.e. temperature, clinical testing rates, and testing coverage) and range of C_{RNA} and prevalence on the correlation coefficients between C_{RNA} and COVID-19 case numbers, the normality of these data was firstly evaluated using Shapiro-Wilk test through R (ver. 3.31, <http://www.R-project.org/>), with results provided in Table S2. Although most of the data were not in normal distribution ($p < 0.05$ in the Shapiro-Wilk test), previous studies revealed Pearson's correlation coefficient was insensitive to the violations of the normality and showed significantly better performance over Spearman's and Kendall's correlation for non-normal data (Chok, 2010; Havlicek and Peterson, 1976). Thus, the impacts of these factors on the correlation coefficients between C_{RNA} and COVID-19 case numbers were assessed through Pearson's correlation analysis using R. The results from Pearson's correlation were comparable ($p > 0.05$) to that of Spearman's correlation (Fig. S2) in this study. As the sampling technique was summarized as a category variable, its impact on the correlation coefficients between C_{RNA} and COVID-19 case numbers were assessed through Point-biserial correlation using R.

3. Results and discussions

3.1. Correlations between C_{RNA} and clinical confirmed case numbers

The systematic review identified 133 correlation coefficients between the C_{RNA} and clinically confirmed COVID-19 case numbers from 27 publications (Table 1). Samples were collected from different WWTPs with a population of 10,000 - 2 million, or manholes of hospitals or university dorms inhabited by thousands of people. Within the 27 studies, 9 studies correlated C_{RNA} (concentration of SARS-CoV-2 RNA in wastewater or Ct or Cq values obtained in RT-qPCR analysis) with active cases (P_A); 4 studies with cumulative cases (P_C); 8 studies with daily new cases (P_{DN}); 8 studies with weekly new cases (P_{WN}), or seven-day rolling average cases (P_{7D}); and 7 studies with various upcoming new daily cases in the future (2–24 days after the wastewater sampling, details in Table 1) (P_{FN}). Twelve studies observed significant correlations ($p < 0.05$) between C_{RNA} with clinically confirmed case numbers. Five studies correlated the normalized $C_{RNA(N)}$ (using human fecal indicators including pepper mild mottle virus (PMMoV), crAssphage, and Bacteroides HF183 as a process control) with case numbers, while 26 studies applied raw WBE data for the correlation analysis. In these 5 studies, the correlation coefficients between COVID-19 cases and both C_{RNA} and normalized $C_{RNA(N)}$ were collected. This review aims to evaluate the importance of normalization using fecal indicators on the correlation between C_{RNA} and COVID-19 cases, thus, the correlation coefficients collected from different fecal indicators were not separated.

The correlation coefficient between C_{RNA} and P_A , P_C , P_{DN} , and P_{FN} was labeled as R_A , R_C , R_{DN} , and R_{FN} , respectively. Due to the limited data size of correlation coefficients between C_{RNA} and P_{7D} , and the similarity of P_{7D} and P_{WN} , the correlation coefficients between C_{RNA} and P_{7D} and between C_{RNA} and P_{WN} were both classified as R_{WN} . A recent study revealed that 36 standard operating procedures (SOPs) commonly applied in WBE studies, including both wastewater sampling and analytical approaches (i.e. waster concentration, RNA extraction, reverse transcription, and qPCR quantification), have a high reproducibility with a standard deviation of 0.13 \log_{10} (Pecson et al., 2021). Thus, for each study using the same sampling and analytical procedures, the impact of wastewater sampling and subsequent analytical protocols on the correlation coefficients between C_{RNA} and COVID-19 cases might be limited. The literature data available does not support a specific analysis towards SOPs in this study.

Variations of C_{RNA} in the same sample due to the choice of primer and probes sets were commonly observed (Feng et al., 2021; Wu et al., 2020). The 27 studies included in this review measured the C_{RNA} using either one or both of N1 and N2 primer and probe sets. Since the correlation coefficients between C_{RNA} and COVID-19 case numbers were calculated based on the results from the same primer and probe sets, and a strong correlation ($R > 0.8$) between the results from N1 and N2 was achieved in these studies (Ai et al., 2021; Kitamura et al., 2021; Nagarkar et al., 2021), the correlation coefficients between C_{RNA} and COVID-19 case numbers included in this study were not differentiated based on the primer and probe sets.

All the WBE studies included in the meta-analysis did not specifically mention their outbreak stages. Considering the potential changes in the shedding dynamic during the infection (Miura et al., 2021), the correlation coefficients in these studies were further summarized based on the outbreak stages. The studies with a clear increasing trend of pandemic activity (active cases or daily new cases) were categorized as pre-peak stage, while the studies with a clear decreasing trend were categorized as post-peak stage (WHO, 2009). The correlation coefficients from the studies where data from both the outbreak stages (i.e. pre-peak and post-peak stage) were involved or the stage was not clear and specific, were summarized as both/non-specified stages. The correlation coefficients were reported at the pre-peak stage in 9 studies, the post-peak stage in 4 studies, and both/non-specified (include both pre- and post-peak stage or not clearly specified stage) in 16 studies. With the raw

Table 1Summary of correlations between SARS-CoV-2 RNA concentration (raw (C_{RNA}) or normalized ($C_{RNA}(N)$)) in wastewater and clinically confirmed COVID-19 case numbers.

Location	WBE sampling				Range of cases ^a or prevalence (cases/100,000 people) ^b	Concentration range (C_{RNA} : copies/mL; $C_{RNA}(N)$: copies/copies HFT ^c)	Correlation			Outbreak stage	Reference
	Site, population	Type, sampling mode, and frequency	Sampling date	Number			Type	R	p		
Calgary, Canada	Hospitals, > 2100	Hospital wastewater, 24 h composite, weekly	August to December 2020	23–40	4–45 ^a	Cq:28.5–41.1 $C_{RNA}(N)$: 20–100	C_q vs P_A C_{RNA} vs P_A $C_{RNA}(N)$ vs P_A	0.48–0.86 0.29–0.54 0.19–0.53	< 0.05 0.0008–0.0858 0.0009–0.2737	Pre-peak	(Acosta et al., 2021)
Riyadh, Saudi Arabia	WWTPs, n.m. ^d	Influent from WWTPs, n.m., monthly	June to August 2020	3	57–160 ^b	Ct:20–35, C_{RNA} :60–90	Ct vs P_A	0.37–0.42	Not significant	n.m.	(Alahdal et al., 2021)
Santiago de Queretaro, Mexico	WWTPs, 0.01–0.35 M	Influent from WWTPs, grab, monthly	April to July 2020	3–6	10 ² –10 ^{3.5a}	C_{RNA} : 10–10 ⁶	C_{RNA} vs P_C	0.63–0.97	0.13–0.18	Pre-peak	(Carrillo-Reyes et al., 2021)
Ottawa, Canada	WWTPs, 1 M	Primary clarified sludge, 24-h composite, bi-monthly	June to August 2020	19	0.3–3 ^b	Ct:33–41 $C_{RNA}(N)$: 0.5×10^{-4} – 4×10^{-4}	$C_{RNA}(N)$ vs P_{DN}	0.65–0.67	< 0.001	Pre- and post- peak	(D'Aoust et al., 2021a)
Quebec, Canada	WWTP, 1.1 M	Primary clarified sludge, 24-hour composite, fortnightly	April to June 2020	8–13	10–60 ^b	C_{RNA} : 25–750	C_{RNA} vs P_A $C_{RNA}(N)$ vs P_A C_{RNA} vs P_{DN}	-0.29–0.16 0.24–0.35 -0.14–0.21	< 0.001–0.02 < 0.05 < 0.001–0.02	Pre- and post- peak	(D'Aoust et al., 2021b)
	WWTP, 0.3 M				15–60 ^b	C_{RNA} : 200–400	$C_{RNA}(N)$ vs P_{DN} C_{RNA} vs P_A $C_{RNA}(N)$ vs P_A C_{RNA} vs P_{DN} $C_{RNA}(N)$ vs P_{DN}	0.14–0.37 0.92–0.95 0.26–0.48 0.14–0.40 -0.14–0.40	< 0.05 < 0.05 < 0.05 < 0.05 < 0.05		
University of North Carolina at Charlotte, USA	Plumbing cleanouts, or manholes	Campus wastewater, 24 h composite, n.m.	October to November 2020	19 sites, 6 sampling events	0–700 ^a	Sample positive ratio: 0–3%	Total number of positive wastewater samples vs P_{DN}	0.77	n.m.	Pre-peak	(Gibas et al., 2021)
Mendoza, Argentina	WWTP, 0.5 M WWTP, 0.5 M	Influent from WWTPs, grab, weekly, n.m.	July to November 2020	10–13	50–390 ^b	C_{RNA} : 16–220 C_{RNA} : 40–310	C_{RNA} vs P_{WN}	0.32–0.39 0.55–0.63	0.14–0.15 0.0069–0.3946	Pre- and post- peak	(Giraud-Billoud et al., 2021)
Jeddah, Saudi Arabia	Hospital, ca. 2884 persons	Hospital wastewater, grab, 3–5 samples per week	April to July 2020	22–30	73–236 ^a	C_{RNA} : 0.2–6	C_{RNA} vs P_C	0.21–0.24	n.m.	Pre- and post- peak	(Hong et al., 2021)
Halifax, Canada	WWTP, 0.1 M	Influent from WWTPs, 24 h composite, n.m.	October 2020 to March 2021	5–7	25–150 ^b	C_{RNA} : 2.5–25	C_{RNA} vs P_A	0.98–0.99	< 0.001	Pre- and post- peak	(Huang et al., 2021)
The Netherlands	WWTPs, 0.3–1.0 M	Influent from WWTPs, 24 h composite, n.m.	February to March 2020	14–16	0.1–100 ^b	C_{RNA} : $10-2.2 \times 10^3$	C_{RNA} vs P_C Ct vs P_C	0.77–0.81 -0.88	< 0.05 < 0.001	Pre-peak	(Medema et al., 2020)
Montana, USA	WWTP, 0.05 M	Influent from WWTPs, 24 h composite, n.m.	March to June 2020	18	10–120 ^b	C_{RNA} : 0.1–10	C_{RNA} vs P_{FN} 2 days later C_{RNA} vs P_{FN} 4 days later	0.91–0.99 0.90–0.98	n.m. n.m.	Post-peak Pre-peak	(Nemudryi et al., 2020)
Budapest, Hungary	WWTPs, 0.05 M	Influent from WWTPs, 24 h composite and grab, n.m.	June to October 2020	22	28.88–769.76 ^b	C_{RNA} : 5–10 ³	C_{RNA} vs P_A C_{RNA} vs P_{WN} Viral load vs P_A Viral load vs P_{DN}	0.65 0.76 0.75 0.82 0.84	< 0.01 < 0.001	Pre-peak	(Róka et al., 2021)

(continued on next page)

Table 1 (continued)

Location	WBE sampling		Sampling date	Number	Range of cases ^a or prevalence (cases/100,000 people) ^b	Concentration range (C _{RNA} : copies/mL; C _{RNA} (N): copies/copies HFI ^c)	Correlation			Outbreak stage	Reference
	Site, population	Type, sampling mode, and frequency					Type	R	p		
Catalonia, Spain	WWTPs, 0.2–1.5 M	Influent from WWTPs, 24 h composite, n.m.	March to November 2020	27–46	10–10 ^{3.5a}	C _{RNA} : 10–10 ⁵	Viral load vs P _{FN} 11 days later				
							Viral load vs P _{FN} 3 days later	0.88			
							C _{RNA} vs P _{FN} 7 days	0.62	n.m.	Post-peak	(Rusiñol et al., 2021)
							Viral load vs P _{FN} 7 days	0.86			
							C _{RNA} vs P _{FN} 7 days	0.53		Pre-peak	
Tulane University, USA	Manholes in a university	Campus wastewater, grab, n.m.	August–December 2020	81	269–526 ^a	C _{RNA} : 0.2–802	Viral load vs P _{FN} 7 days	0.87			
							C _{RNA} vs P _A	0.48–0.51	< 0.001	Pre- and post- peak	(Scott et al., 2021)
Porto, Portugal	WWTPs, 0.2 M	Influent from WWTPs, 24 h composite, n.m.	May 2020 to March 2021	25–40	0.6–42.1 ^b	Solids: C _{RNA} : 0–0.15 copies/ng-RNA Liquid: C _{RNA} : 0–0.15 copies/ng-RNA	C _{RNA} vs P _{7DA}	0.43	< 0.001	Pre- and post- peak	(Tomasino et al., 2021)
							C _{RNA} vs P _{7DA}	0.54–0.65	< 0.001		
Utah, USA	7 WWTPs, 1.26 M WWTP, 9095	Influent from WWTPs, 24 h composite, n.m.	April to May 2020	77	0–80 ^b	C _{RNA} : 0.1–1000 copies/day-person	C _{RNA} vs P _{DN}	0.54	< 0.001	Pre- and post- peak	(Weidhaas et al., 2021)
				6	0–250 ^b	C _{RNA} : 0.1–600 copies/day-person	C _{RNA} vs P _{WN}	0.96	< 0.01	Pre-peak	
				6	0–150 ^b	C _{RNA} : 0.1–250 copies/d-person	C _{RNA} vs P _{WN}	0.82	< 0.05	Pre-peak	
				8	0–250 ^b	C _{RNA} : 0.1–500 copies/day-person	C _{RNA} vs P _{FN} 7 days later	0.8	< 0.01	Post-peak	
Germany	WWTPs, 0.1–2 M	Influent from WWTPs, 24 h composite, n.m.	April 2020	9	50–1000 ^b	C _{RNA} : 10–100 copies/mL	Daily viral load vs P _C	0.99	n.m.	n.m.	(Westhaus et al., 2021)
France	Sewer network, 0.4 M	Influent from WWTPs, 24 h composite, n.m.	July to December 2020	117	5–180 ^a	C _{RNA} : 300–8000	Viral load vs P _A	0.99			
Frankfurt, Germany	WWTPs, 0.5–1.4 M	Influent from WWTPs, 24 h composite, twice/week	March to September, 2020	13	2–40 ^a	C _{RNA} : 1–5000	C _{RNA} vs P _A	0.65	< 0.01	Pre- and post- peak	(Wurtz et al., 2021)
Ohio, the USA	WWTPs, 0.01–0.05 M	Influent from WWTPs, 24 h composite, twice/week	July 2020 to January, 2021	250	500–1000 ^b	C _{RNA} : 0.1–100	C _{RNA} vs P _{DN}	0.75	< 0.01	Pre- and post- peak	(Ai et al., 2021)
							C _{RNA} vs P _{FN} 3–5 days	0.48–0.79	< 0.01		
							C _{RNA} vs P _{WN}	0.76–0.85	< 0.001		
							C _{RNA} (N) vs P _{FN} 5 days	0.78–0.84	< 0.001		
Buenos Aires, Argentina	WWTPs, 0.01–0.02 M	Influent from WWTPs, grab, weekly	June 2020 to April 2021	174	0–1100 ^a	C _{RNA} : 0–1300	C _{RNA} vs P _{FN} 10 days	0.79	< 0.001	Pre- and post- peak	(Barrios et al., 2021)
							C _{RNA} vs P _{FN} 15 days	0.80	< 0.001		
							C _{RNA} vs P _{FN} 20 days	0.81			
								0.81			

(continued on next page)

Table 1 (continued)

Location	WBE sampling		Sampling date	Number	Range of cases ^a or prevalence (cases/100,000 people) ^b	Concentration range (C _{RNA} : copies/mL; C _{RNA} (N): copies/copies HFI ^c)	Correlation			Outbreak stage	Reference
	Site, population	Type, sampling mode, and frequency					Type	R	p		
Scotland	WWTPs, 0.01–0.6 M	Influent from WWTPs, 24 h composite, weekly	April 2020 to February 2021	12–112	0–15,000 ^a	C _{RNA} : 0–150	C _{RNA} vs P _{7D} Viral load vs P _{7D}	0.79 0.91	n.m.	Pre- and post- peak	(Fitzgerald et al., 2021)
Japan	Manhole and WWTPs	Influent from WWTPs, grab, weekly	June to August 2020	10–11	0–500 ^a	C _{RNA} : 0.1–15	C _{RNA} vs P _{WN}	0.71	< 0.01	Pre- and post- peak	(Kitamura et al., 2021)
France	WWTPs, 0.05–0.5 M	Influent from WWTPs, 24 h composite, weekly to monthly	July to December 2020	138	1–1000 ^b	C _{RNA} : 1–1000	C _{RNA} vs P _A	0.66	n.m.	Pre- and post- peak	(Lazuka et al., 2021)
USA	WWTPs, 0.03–0.5 M	Influent from WWTPs, 24 h composite, weekly to monthly	May to December 2020	20–29	10–800 ^a	C _{RNA} : 1–1000	C _{RNA} vs P _{DN} C _{RNA} (N) vs P _{DN}	0.22–0.71 0.25–0.67	n.m.	Pre- and post- peak	(Nagarkar et al., 2021)
Bangkok, Thailand	WWTPs, 0.01–0.6 M	Influent from WWTPs, 24 h composite and grab, weekly to monthly	January to April 2021	25	0–20 ^b	C _{RNA} : 1–3100	C _{RNA} vs P _{FN} 22–24 days later	0.85–1	n.m.	Pre-peak	(Sangsanont et al., 2021)
USA	WWTPs	Influent from WWTPs, grab, weekly	July to August 2020	138	5–25 ^b	C _{RNA} : 1–9000	C _{RNA} vs P _{7D}	0.83	0.04	Post- peak	(Street et al., 2021)

Note:

Ct: cycle threshold.

Cq: quantification cycle.

^a Range of cases.

^b Range of prevalence (cases/100,000 people).

^c HFI: human fecal indicators including PMMoV, HF183, and crAssphage.

^d n.m. Not mentioned.

C_{RNA} data, at the pre-peak stage, R_A , R_{DN} , R_{FN} , and R_C were 0.66 ± 0.17 , 0.82 ± 0.04 , 0.81 ± 0.13 , and 0.80 ± 0.24 , respectively (Fig. 1). At the post-peak stage, the R_{WN} was 0.81 ± 0.10 and R_{FN} was 0.85 ± 0.16 using the raw C_{RNA} . At both/non-specified stages, the R_A , R_{DN} , R_{WN} , R_{FN} , and R_C were 0.62 ± 0.46 , 0.43 ± 0.33 , 0.46 ± 0.36 , 0.80 ± 0.02 , 0.48 ± 0.44 , respectively, based on raw C_{RNA} . Separation of the outbreak stages improved the correlations between C_{RNA} and case numbers, and reduced the variations and scatteredness of the correlation coefficients.

Using the raw C_{RNA} , with a clear indication of the stage (either pre- or post-peak stage), reasonable estimations ($R \geq 0.7$) for all the case numbers can be achieved through the WBE approach, although WBE surveillance performs better on the estimation of new cases (either weekly new, daily new, or future new), rather than the active cases or the cumulative cases. With the normalized concentration $C_{RNA}(N)$, R_{DN} , R_{WN} , and R_{FN} was 0.53 ± 0.21 , 0.54 ± 0.23 , 0.81 ± 0.05 , respectively at both/non-specified stages. With the raw C_{RNA} data, R_{DN} , R_{WN} , and R_{FN} were 0.43 ± 0.33 , 0.46 ± 0.36 , and 0.80 ± 0.02 , respectively at both/non-specified stages. The normalization of C_{RNA} through fecal indicators slightly improved the correlations in comparison to raw C_{RNA} for R_{DN} , R_{WN} , and R_{FN} . The correlation coefficient between WBE data and clinically identified cases was higher for new cases using $C_{RNA}(N)$ than that of active cases. These observations suggest that the viral shedding from new cases might play a more important role on WBE estimations for COVID-19 prevalence or incidence, and the correlations between C_{RNA} and case numbers are likely different at the pre-peak and post-peak stages.

The better performance of WBE surveillance on capturing the new cases is likely related to the shedding dynamics. To date, sputum and feces are regarded as the major shedding sources for SARS-CoV-2 RNA in wastewater, due to their higher shedding probability and magnitude, and the possibility of entering sewers than other shedding sources (i.e. saliva, blood, and urine, etc.) (Crank et al., 2022; Li et al., 2022). However, depending on the severity of the symptom and personal hygiene behavior, the contribution of sputum and feces to the overall C_{RNA} varies greatly (Crank et al., 2022; Li et al., 2022). Thus, the shedding dynamic of both sources affects the C_{RNA} under a certain infection status. A meta-level study analyzed the clinical findings of fecal shedding and estimated that the fecal shedding peaked at 0.34 day after the symptom onset, and the peak fecal shedding concentration was about 10^3 times higher than the median concentration over the whole shedding period (Miura et al., 2021). The sputum shedding loads peaked in the first week after the symptom onset, which was about 10 times higher than the average concentration over the whole sputum shedding period (Jones et al., 2020). Therefore, the daily new and weekly new cases likely play a dominant role in the total SARS-CoV-2 RNA loads from patients in the wastewater during the pre-peak stage when new cases make up a significant portion of the total active cases.

Using both raw concentration and normalized concentration, R_{FN} reached 0.8–0.9, which was the highest among other correlation

coefficients, suggesting a strong correlation between the C_{RNA} and P_{FN} (Fig. 2). The clinical testing majorly relies on the motivation of individuals, in addition to some mandatory testing required for cross-border travelers or close contacts of infected patients (Table S1). The availability of the testing facility, awareness of the testing and contact tracing policy, and own motivation and symptom severity could lead to potential delays between the SARS-CoV-2 RNA shedding from the patients and their clinical confirmations (Mackey et al., 2021; Yearby and Mohapatra, 2020). Delays between the appearance of symptoms, testing and the reporting of test results were commonly observed in different countries (Peccia et al., 2020). Moreover, a study stated the potential lag in viral detection in deep nasal and oropharyngeal samples (current major clinical testing approaches) than fecal shedding by monitoring the wastewater collected from international aircraft. Although 48 h prior to the flight, all repatriation flight passengers (age greater than 5 years) were tested negative for COVID-19 via both deep nasal and oropharyngeal testing, wastewater collected from aircraft were tested positive, while 112 COVID-19 cases were detected during the mandatory 14-day quarantine after landing (Ahmed et al., 2022). Thus, the future cases are likely at the early stage of the infection, who already shed SARS-CoV-2 RNA into sewers at a considerable amount, although have not been tested or confirmed clinically. This is also consistent with the varied leading time (2–24 days) between WBE estimations and the clinical testing for future cases in different catchments (Table 1), where the clinical testing policy, coverage and population demographics are likely different. In addition, some previous studies observed higher than expected C_{RNA} in wastewater and hypothesized a potential surge of shedding (either from feces or other sources) before the symptom onset at several orders of magnitude greater than typical values (Wu et al., 2020). This could be another reason for the better performance of WBE

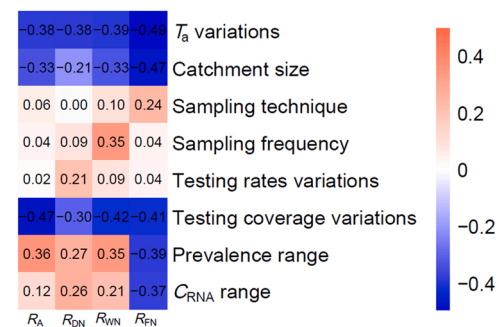


Fig. 2. Pairwise correlation plot between WBE correlation coefficients (R_A , R_{DN} , R_{WN} , P_{FN}) and environmental, sampling design, and epidemiological parameters. The Point-Biserial correlation was determined between sampling technique and WBE correlation coefficients, while Pearson correlations were determined for other parameters.

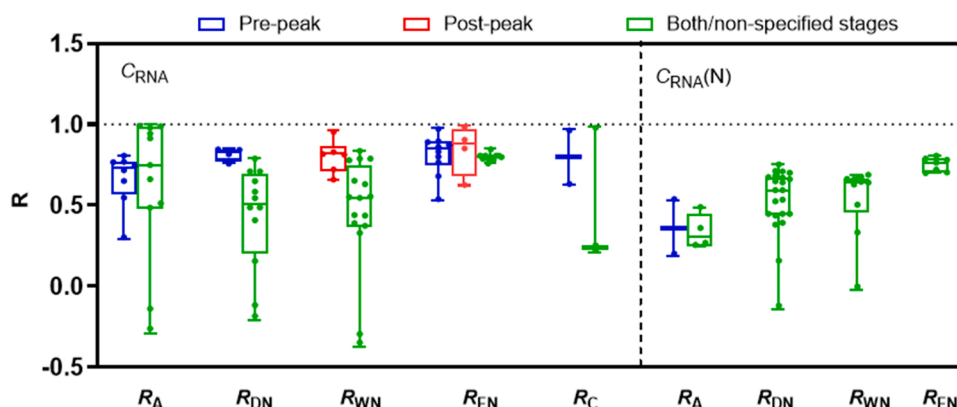


Fig. 1. Correlation coefficients (Pearson or Spearman correlation, R) that were reported in different studies between the raw SARS-CoV-2 RNA concentrations in wastewater (C_{RNA}) or normalized SARS-CoV-2 RNA concentrations in wastewater ($C_{RNA}(N)$) and P_A (R_A), P_{DN} (R_{DN}), P_{WN} (R_{WN}), P_{FN} (R_{FN}) and P_C (R_C). Blue, red and green indicates data associated with the pre-peak, post-peak and both/non-specified stages of the COVID-19 outbreak, respectively. The middle line of the box represents the median; the upper and lower lines represent the 25th and 75th percentiles; the whiskers extending the box and the outliers represent the data outside the interquartile range.

surveillance on capturing the new cases, especially the future cases as observed by the systematic literature review. However, since most of the virological assessments were carried out on patients with clear clinical evidence (such as oropharyngeal swabs, nasopharyngeal swabs, and symptoms), this hypothesis lacks support by clinical data and requires further investigations.

Higher correlation coefficients were observed for studies with a clear indication of the outbreak stage (i.e. pre-peak or post-peak) than both/non-specified stages (Fig. 1). It suggests that the correlation between C_{RNA} and the clinically confirmed cases is potentially different at pre- and post-peak stages of the same outbreak in the same catchment. As discussed above, WBE showed better capability in capturing new cases than active cases due to the higher shedding loads from new patients. Thus, the different correlations at pre-peak and post-peak stages are likely related to the changed fractions of new cases among the total infections. At the pre-peak stage, the new cases among the active cases (shedders of SARS-CoV-2 RNA into sewers) are higher than those at the post-peak stage; while at the post-peak stage, the majority of patients (shedders) are recovering from the infection with lower shedding loads (i.e. about 0.3%–6.2% of the shedding loads at the early stage of the infection) (Jones et al., 2020; Miura et al., 2021). Therefore, it is recommended to consider the outbreak stage for the prevalence estimation through the WBE approach.

With the current dataset, most of the correlation coefficients observed between $C_{RNA}(N)$ and case numbers were at both/non-specified stages. Under such conditions, the normalization of C_{RNA} using human fecal indicators including PMMoV, HF183, or crAssphage concentrations slightly increased the correlations between C_{RNA} and P_{DN} , P_{WN} , and P_{FN} , but reduced the correlations between C_{RNA} and P_A (Fig. 1). Normalization through fecal indicators was applied in the WBE studies for COVID-19 in the last 1.5 years, where the awareness and the experience towards SARS-CoV-2 RNA detection in wastewater have been greatly improved than the early stage of the COVID-19 outbreaks (i.e. 2019–2020). Previously, many studies have attributed the superior performance of normalized C_{RNA} on COVID-19 estimations to the reduced noise inherent to the sampling, transport, and processing of the samples, as the concentration of human fecal indicators represents the fecal loads in wastewater (D'Aoust et al., 2021b). However, normalization reduced the correlation coefficients in 8 of 12 WWTPs using PMMoV (Feng et al., 2021) and in another catchment area using crAssphage (Ai et al., 2021). The reduced correlation coefficients after the normalization were attributed to that the influence of variability across measurements (SARS-CoV-2 and fecal indicators like PMMoV) is stronger than that of differences in the fecal loads in the samples (Feng et al., 2021). With the current limited data from the systematic review, the normalization of C_{RNA} improved the accuracy of WBE on COVID-19 estimations for new cases (daily new, weekly new or future new). The performance of each fecal indicator on COVID-19 prevalence or incidence estimations varied in different catchments. For instance, PMMOV showed better performance than HF183 and crAssphage in 2/3 of the catchments but HF183 had better performance in the other 1/3 of the catchments (Nagarkar et al., 2021). The selection and performance of fecal indicators in improving WBE estimation by normalization remain unclear, but important for further investigations.

3.2. Impact of environmental factors on the WBE performance

Environmental factors including the variations of air temperature during the study period (maximum temperature minus minimum temperature) and the catchment size were included to assess their impacts on each correlation coefficients (i.e. R_A , R_{DN} , R_{WN} , and R_{FN}). Considering the recovery or death of patients, cumulative cases (P_C) do not reflect the actual shedders of SARS-CoV-2 RNA into sewers thus R_C was not included for the impacting parameter analysis. The catchment size and variations of air temperature during the study period negatively correlated ($R = -0.2$ to -0.5) with R_A , R_{DN} , R_{WN} , and R_{FN} . Higher variations of

air temperature and catchments with larger sizes were associated with lower R_A , R_{DN} , R_{WN} , and R_{FN} , suggesting their negative impact on the accuracy of COVID-19 estimation. This is likely related to the SARS-CoV-2 RNA decay during the in-sewer transportation.

In bulk wastewater, the decay of SARS-CoV-2 RNA was reported to follow first-order decay kinetics as Eq. 1, where the decay was facilitated by higher wastewater temperature and longer HRT (Ahmed et al., 2020b; Bivins et al., 2020a).

$$C_t = C_0 \times e^{-kt} \quad (1)$$

where C_t and C_0 are the concentrations of SARS-CoV-2 RNA in wastewater at time t and time 0, respectively, and k is the decay rate constant, which increases with the increase of temperature (Ahmed et al., 2020b). Before reaching the WWTP or sewer sampling points, wastewater are transported by sewers for several minutes to hours (i.e. HRT) (Kulandaivelu, 2020; Shi, 2022). Theoretically, under the same temperature in the same catchment (fixed HRT), the decay ratio $\frac{C_t}{C_0}$ is constant, which makes the correlation between C_t and COVID-19 cases comparable to that between C_0 and COVID-19 cases. However, during the WBE surveillance, the wastewater temperature is generally dynamic. Although the wastewater temperature and HRT of the catchment were not reported in these studies included in the meta-analysis, a higher air temperature (T_a) and larger catchment size were generally associated with higher wastewater temperature and longer HRT of the catchment, respectively (Hart and Halden, 2020b; Jiang et al., 2022; McCall et al., 2017). A higher fluctuation of air temperature (maximum temperature minus minimum temperature) during the study period could lead to higher variations of RNA loss (i.e. $\frac{C_t}{C_0}$ in Eq. 1) during the in-sewer transportation, in particular for the larger catchments with longer HRTs. Furthermore, a previous study using WBE for estimating illicit drug consumption found that the location and distribution of drug consumers also greatly affected the in-sewer traveling times of chemicals, especially for larger catchments, leading to higher uncertainties for the consumer estimations (McCall et al., 2017). Although the impact of the distribution of COVID-19 patients in the catchment area on the SARS-CoV-2 RNA decay during in-sewer transportation remains unclear, the distribution of patients in a larger catchment is likely to be more diverse than in a small catchment. Therefore, a higher fluctuating range of air temperature and a larger catchment size introduced more uncertainties in SARS-CoV-2 RNA decay and thus the measured C_{RNA} , leading to lower R_A , R_{DN} , R_{WN} , and R_{FN} .

3.3. Impact of sampling design on WBE performance

The impacts of sampling technique and frequency on the correlation coefficient between C_{RNA} and clinically confirmed COVID-19 cases were also evaluated. In the reviewed studies, both grab and composite sampling have been applied and the correlation coefficients were determined with 3–117 samples collected over 1–11 months (Table 1). Composite sampling showed negligible positive correlations ($R = 0.00$ – 0.10) with R_A , R_{DN} , and R_{WN} , and weak positive correlation ($R = 0.24$) with R_{FN} (Fig. 2). This implies that based on currently limited data from these studies, grab and composite sampling provided comparable correlation coefficients between C_{RNA} and clinically confirmed COVID-19 cases, while the composite sampling technique slightly improved the performance of WBE on COVID-19 prevalence estimation.

Composite sampling is generally recommended as it captures all the contributions of virus shedders in the catchment within 24 h. However, 45% of WBE studies still use grab sampling due to the high cost of autosamplers and subsequent requirements for the operation (Bertels et al., 2022). In these studies for the meta-analysis, except that Barrios et al. (2021) and Street et al. (2021) did not provide the detailed sampling time, all other studies using grab sampling collected wastewater between 8 and 11:30 am, where wastewater level or fecal loading was high based on their previous observations. While, composite sampling

covered 24 h, which potentially captured more shedders. The comparable R_A , R_{DN} , and R_{WN} between grab and composite sampling are thus likely related to the diurnal pattern of C_{RNA} in the wastewater. Clear diurnal patterns of C_{RNA} have been observed on consecutive days in WWTPs of different capacities (i.e. thousands to millions equivalent population) (Augusto et al., 2022; Bivins et al., 2021; Curtis et al., 2020; George et al., 2022). The diurnal pattern of C_{RNA} is likely related to the water usage pattern, where toilet use (a major source for SARS-CoV-2 entering sewers) more commonly occurs in the morning than other times (Heaton et al., 1992). Furthermore, Augusto et al. (2022) revealed that the C_{RNA} from grab sampling at the peak shedding time (i.e. 8–10 am), reached a good agreement with the C_{RNA} from 24 composite sampling, so did the correlation between C_{RNA} and COVID-19 cases from both from grab and composite sampling results, over a 17 weeks monitoring.

However, it is worthy to mention that the impact of sampling technique (i.e. grab or composite sampling) on the variability of C_{RNA} remain controversial to date, with conflicting reports from the literature. Gerrity et al., 2021a observed a 10-fold increase in C_{RNA} from composite sampling than that of corresponding grab sampling of primary effluent samples, presumably highlighting diurnal variability in the SARS-CoV-2 signal. In contrast, another study found that the sampling technique showed negligible impacts on C_{RNA} , where a good agreement between most grab samples and their respective composite was observed (Curtis et al., 2021). Thus, the variability and design of the grab sampling on the correlation between C_{RNA} and clinically confirmed COVID-19 cases in comparison to the corresponding composite sampling still require future investigations.

The sampling frequency had slightly positive correlations ($R=0.04$ – 0.35) with R_A , R_{DN} , R_{WN} , and R_{FN} (Fig. 2). The sampling frequency in these studies in the meta-analysis ranged from 0.5 to 2, 1–4, 0.5–2, and 0.25–2 times per week for R_A , R_{DN} , R_{WN} , and R_{FN} , respectively (Table 1). This is consistent with observations from long-term wastewater surveillance studies, where the sensitivity of WBE increased with the increase of sampling frequency over a week (Feng et al., 2021; Graham et al., 2021). However, in the studies involved in the meta-analysis, most of the sampling frequency was around or below 2 times per week. For future research, detailed evaluations of the sampling technique and the impact of higher sampling frequency (>2 times per week) on the performance of WBE for COVID-19 prevalence estimation are highly recommended. It is worth noting that Pearson's correlation between the sampling frequency and R_A , R_{DN} , and R_{WN} (Fig. 2) was comparable to Spearman's correlation (Fig. S2). However, a negative Spearman's correlation ($R=-0.3$) was observed for R_{FN} , in contrast to a weak positive Pearson's correlation ($R=0.04$). This difference might be related to the methodology of these two correlation analyses, where Spearman's correlation coefficient is based on the ranked values for each variable, while Pearson's correlation is based on the raw data (Chok, 2010). Considering the limited data points of the R_{FN} available in WBE studies included in this study and the conflicting correlation between Spearman's correlation and Pearson's correlation, a more detailed analysis evaluating the impact of sampling frequency on the R_{FN} is highly recommended in the future.

3.4. Impact of epidemiological factors on WBE performance

The epidemiological factors such as the clinical testing rate, testing coverage, and the range of C_{RNA} and prevalence rate (reflecting the level of COVID-19 infection in the population) are directly associated with the pandemic monitoring and outbreak status. The variation of clinical testing rate (tests/1000 people over 30 days) had negligible correlations with R_A , R_{DN} , R_{WN} , and R_{FN} ($R=0.02$ – 0.21), while a larger variation of the clinical testing coverage (tests/confirmed cases) showed stronger negative correlations ($R=-0.47$ to -0.30), with R_A , R_{DN} , R_{WN} , and R_{FN} (Fig. 2). These observations imply that the correlation coefficient between C_{RNA} and clinically confirmed COVID-19 cases is more closely

associated with the variations of clinical testing coverage rather than the variations of clinical testing rates. The variations of clinical testing coverage during the study period reduced R_A , R_{DN} , R_{WN} , and R_{FN} .

For WBE studies, the C_{RNA} detected in the wastewater relies on the contributions of SARS-CoV-2 RNA shedders (patients) regardless of their clinical conditions (tested or not tested), which can largely reflect the 'true' infection rate in the population (Ahmed et al., 2020a). However, due to the unknown 'true' infection status, most WBE studies compared the C_{RNA} with the clinically confirmed COVID-19 cases. The correlation between C_{RNA} and clinically confirmed COVID-19 cases is thus biased to the capability of clinical testing in capturing the true infections. In most countries, symptom-onset can be a major trigger for the motivation of testing, in addition to some mandatory testing required for cross-border travelers or close contacts of infected patients (Li et al., 2022). Thus, the presence of asymptomatic patients likely led to the under-reporting of COVID-19 patients through clinical testing (He et al., 2021). Furthermore, the clinical testing is potentially biased due to social-economic factors and individual motivations, especially for resource-poor countries or population groups. For instance, racial and ethnic minorities tend to have a lower testing ratio but higher infections and mortality rates due to being unable to work from home or taking sick leaves and or limited understanding and resources of the public health system and testing policy, etc. (Mackey et al., 2021; Yearby and Mohapatra, 2020). Thus, clinical testing usually captures part of the 'true' infection. The Centre for Disease Control and Prevention (CDC) estimated that only 1 in 4.3 (95% UI 3.7–5.0) of total COVID-19 infections were reported through clinical testing (Reese et al., 2020). Therefore, the ratio of clinically captured cases among the total infection is critical when assessing the correlation between C_{RNA} and clinically confirmed COVID-19 cases.

Theoretically, if a fixed ratio of cases among the 'true infection' can be captured by the clinical testing throughout the study period, the relationship between C_{RNA} and clinically confirmed COVID-19 cases will be comparable to that between C_{RNA} and 'true' COVID-19 infections. However, the clinical testing coverage and rate are generally dynamic, which would affect the ratio of clinically captured cases among the total 'true' infection, and the subsequent correlation between C_{RNA} and clinically confirmed COVID-19 cases. The variation (change) of the clinical testing (i.e. testing rate and coverage) in each study was thus expected to reflect the changes in the ratio between clinically captured cases and the total 'true' infection. A larger variation of clinical testing coverage is thus linked to a more fluctuating ratio of patients captured by the clinical testing among the total infections during the wastewater monitoring, which leads to the reduced R_A , R_{DN} , R_{WN} , and R_{FN} . Although the clinical testing rate is also affected by the disease prevalence in the community, it is not directly associated with the infection ratios that can be captured by the clinical testing. This is consistent with its negligible correlations observed with R_A , R_{DN} , R_{WN} , and R_{FN} .

The infection status was also reflected by the range of C_{RNA} and corresponding P_A , P_{DN} , P_{WN} , and P_{FN} . The range of C_{RNA} and corresponding prevalence varied from 0.1 to 10^6 copies/mL and from 0.3 to 1000 cases per 100,000 people, respectively (Table 1). A wider range of prevalence and C_{RNA} was associated with higher R_A , R_{DN} , and R_{WN} ($R=0.12$ – 0.36) but lower R_{FN} ($R=-0.39$ to -0.41) (Fig. 2). It suggests that an inherently higher range of P_A , P_{DN} , and P_{WN} during the wastewater monitoring period would improve the correlation coefficients between C_{RNA} and active cases, daily new cases, and weekly new cases, but not the future cases. The increased R_A , R_{DN} , and R_{WN} due to the increase of prevalence range and C_{RNA} range is likely related to the analytical uncertainties. Although a higher C_{RNA} is generally associated with a higher prevalence, varied reproducibility has been observed using the same wastewater sample due to analytical uncertainties (Pecson et al., 2021). As the recovery efficiency of the analytical method was not reported in all the studies (Table S2), it was not included as a factor for the meta-analysis. The change of C_{RNA} under a narrower range of prevalence or C_{RNA} would potentially stress the analytical

uncertainties rather than the actual changes of prevalence, resulting in the negative correlations between the range of prevalence or C_{RNA} and R_A , R_{DN} , and R_{WN} .

The negative impacts of the increasing range of C_{RNA} and P_{FN} on R_{FN} are potentially related to the lag time between the viral shedding and clinical testing of patients. As discussed in Section 3.1, the future cases are likely at the early stage of the infection, who are already shedding SARS-CoV-2 RNA at a considerable load but not being tested or confirmed by the clinical testing (Mackey et al., 2021; Yearby and Mohapatra, 2020). In WBE studies, the correlation coefficients between C_{RNA} and upcoming cases confirmed after the sampling day with different ranges (i.e. several days) are compared, and the range of the future cases with the best correlation coefficient is used for the study period. Thus, the increasing range of C_{RNA} is associated with a higher range of COVID-19 infection, which is likely related to a more diverse lag time between the viral shedding and the clinical testing of patients and the subsequent R_{FN} . This is also consistent with the positive correlation observed between WBE leading time and the ranges of P_{FN} and corresponding C_{RNA} (Fig. S1). The longer and more diverse leading time under a higher range of P_{FN} and corresponding C_{RNA} would thereby lead to a lower R_{FN} . However, it is worthy to mention that the minimum R_{FN} was above 0.7, which implies good estimations of P_{FN} using WBE.

In addition, the prevalence range showed stronger correlations than C_{RNA} range with R_A , R_{DN} , R_{WN} , and R_{FN} (Fig. 2), which is also likely related to the analytical uncertainty of C_{RNA} and in-sewer RNA decay as discussed in the above sections. A recent study stated that small WWTPs were more likely to have the C_{RNA} level lower than the LOD, where the minimum prevalence that can be captured by WBE increased from 0.1 cases per 1000 inhabitants in large catchments to 0.59 cases per 1000 inhabitants in small catchments (Rusiñol et al., 2021). Thus, the range of C_{RNA} detected is not only related to the prevalence range, but also the contribution of multiple parameters including the temperature, analytical approach, and catchment size. The detailed impact of each parameter on the measurement of C_{RNA} and subsequent WBE back-estimation remains unclear, and requires future investigations. In addition, although comparable results were observed between Pearson's correlation and Spearman's correlation for R_A , R_{DN} , and R_{WN} ($R=0.01-0.19$, Spearman's correlation, Fig. S2), a negative correlation ($R=-0.42$) between the variation of clinical testing rate and R_{FN} was observed in Spearman's correlation. As discussed above, this difference is likely related to the methodological difference between Spearman's and Pearson's correlation, and relatively limited R_{FN} data available for a substantive analysis. Thus, further analysis for evaluating the impact of testing rates on the R_{FN} is essential when more R_{FN} data is available in the future.

4. Limitations and implications for WBE application in monitoring COVID-19

The systematic review summarized and compared the correlation coefficients determined between SARS-CoV-2 RNA concentrations in wastewater and clinically confirmed COVID-19 case numbers from WBE studies in different countries. Through the meta-analysis, this study found that WBE surveillance showed higher performance in estimating new cases than active cases or cumulative cases. This is consistent with the current clinical observations of the shedding dynamic from COVID-19 patients. As most of the studies were carried in 2020 to early 2021, the clinical shedding dynamics were majorly based on the Alpha variant (original variant) of SARS-CoV-2. Recent clinical studies revealed that higher viral titers in the Delta and Epsilon variants (recent variants) than that in the Alpha variant of SARS-CoV-2 (Despres et al., 2021). However, the shedding dynamics of these variants in feces, sputum, or other bodily fluids have not been reported yet, which might affect the estimation of case numbers using the WBE approach.

The catchment size, prevalence, sampling frequency, and variations of air temperature and testing coverage showed strong correlations ($|$

$R|=0.3-0.5$) with R_A , R_{DN} , R_{WN} , and R_{FN} . It implies that the WBE performance relies on multiple parameters rather than SARS-CoV-2 RNA concentrations in wastewater samples alone. This is consistent with previous observations for predicting the COVID-19 prevalence using modeling approaches, where incorporating other parameters such as air temperature, catchment size, and testing coverage improved the prediction accuracy (Li et al., 2021a). Due to the data limitation, the average air temperature was collected for the region of the study, while the HRT of the catchment was reflected by the catchment size. However, the actual HRT of the catchment is also affected by the population density, resource availability, etc. Furthermore, under a certain prevalence in a catchment, the C_{RNA} also varied due to the variability of wastewater flow rate, daily per capita water use, and inflow/infiltration, leading to lower correlation coefficients between C_{RNA} and COVID-19 case numbers (Feng et al., 2021; Westhaus et al., 2021). Recent studies stated that reporting daily viral load per capita, which is calculated based on the C_{RNA} , catchment population and wastewater flow rate increased the correlations coefficients (Feng et al., 2021; Rusiñol et al., 2021). Due to the unavailability of such information, these factors and viral loads per capita were not included in the meta-analysis. For future WBE studies, it is highly recommended to record environmental epidemiology parameters, such as temperature, catchment size, HRT, and wastewater flow rate, and optimize the sampling technique and frequency to achieve better accuracy.

In addition, due to the data unavailability of the clinical test rates (tests/1000 people over 30 days) and test coverages (tests/confirmed cases) in the study area, the county-wide information was applied in this study. Although the country-wide information can approximately reflect the testing resource availability in the study area, the socioeconomic development and the outbreak stage might also contribute to the regional differences between the study area and the country. The variation of clinical testing rates and coverage was calculated as the maximum value minus the minimum in this study, which cannot fully reflect the complexity of the clinical testing, especially under different temporal and geological testing capabilities. Thus, for future studies, a time-weighted average clinical testing rate and coverage can be tested together with their variances. Also, special attention shall be paid when the testing capability reaches saturation, while assessing the correlation between C_{RNA} and COVID-19 cases.

Furthermore, in these studies involved in the meta-analysis, the case numbers were either counted based on the residence postcode of the patient or estimated based on the population-weighted average of the city ((infection cases/total population of the city) \times catchment population). The commuting or uneven regional distribution of patients can lead to the inaccuracy of patient numbers related to the SARS-CoV-2 in the wastewater. Thus, exact clinical testing information of studied catchments and consideration of the commuting or regional patient distribution are highly recommended for future study about their impact on WBE accuracy. In addition, for all the WBE studies, the C_{RNA} and COVID-19 cases were collected as time series data. However, information about data whitening for correlation analysis were not provided in these WBE studies, which potentially affects the correlation levels between C_{RNA} and COVID-19 cases (Yue et al., 2002). Thus, for future research, the evaluations of the impact of pre-whitening on such correlations are highly recommended.

Two potential measures may be employed to evaluate the impact of environmental factors (i.e. variation of temperature and catchment size), WBE sampling design (sampling technique and frequency), and epidemiological conditions (i.e. prevalence levels and clinical testing coverage) on the relationship between C_{RNA} and COVID-19 cases: the ratio or correlation coefficients between SARS-CoV-2 copies/mL (or the normalized measure) and corresponding clinical cases. As discussed above, among different countries or catchments with different population distribution, educational levels, socioeconomic development, and access to testing facilities, the ratio of clinically captured cases among the 'true infection' is likely different and cannot be quantified. In

addition, as stated above, the standard operation protocol (SOP) (i.e. sampling and analytical) applied in the same study intend to have high reproducibility but inter-laboratory measurement of C_{RNA} using different SOPs could lead to up to 7 \log_{10} difference (Pecson et al., 2021). Thus, this study directly collected the correlation coefficients between the C_{RNA} and corresponding COVID-19 cases in each study, where the impact of socioeconomic and population conditions and SOPs can be minimized as a consistent SOP was applied in the same catchment in each study. However, the ratio of SARS-CoV-2 copies/mL (or a normalized measure) vs clinical cases would be very useful and straightforward in assessing a single WBE study. For instance, this ratio may change due to the evolution of new SARS-CoV-2 variants, the rollout of vaccination, etc., which requires further investigations in the future.

Due to the lack of information in most WBE studies, 27 out of 749 papers were involved in the meta-analysis. The relevant weather, clinical testing, and catchment information were also collected to the best of the authors' capability. This meta-analysis aims to shed light on the potentially influential factors on the correlation coefficient between C_{RNA} and COVID-19 cases and provide suggestions for future WBE studies to report or consider such factors. When more accurate information about such factors is available in the future, a more detailed assessment is highly recommended. In addition to those intrinsic variations, such as weather or those associated with different WWTPs or sewer catchments, the variation of sampling and analytical methods should be avoided. It is recommended that umbrella organizations like UNEP, WHO and others should gather the published data and provide a standardized method for achieving a more accurate and applicable correlation between C_{RNA} and COVID-19 case numbers. It is possible to achieve a robust, reliable and accurate WBE based environmental surveillance of COVID-19.

5. Conclusions

This study systematically summarized the correlation coefficients determined between C_{RNA} and the corresponding clinically confirmed COVID-19 prevalence or incidence, and provided meta-level analysis assessing impacts of environmental factors, WBE sampling designs and epidemiological parameters on the correlation coefficients. This leads to the following conclusions:

- The C_{RNA} exhibited better correlations with new cases (either daily new, weekly new, or future cases) than that of active or cumulative cases. This is consistent with clinical observations pertaining to shedding dynamics of much higher shedding loads by the patients at the early infection stage.
- Differentiation according to pre- or post-peak of the outbreak can improve the performance of WBE back-estimation of COVID-19 cases significantly. Normalization of C_{RNA} based on the fecal indicators also slightly improved the WBE performance.
- Variations of environmental conditions, epidemiological conditions, and WBE sampling design were found critical for the WBE estimation of COVID-19. Larger variations of air temperature and clinical test coverage, and the increase of catchment size illustrated strong negative impacts on the correlation between C_{RNA} and COVID-19 case numbers.

Environmental implication

The estimation of COVID-19 infections through wastewater-based epidemiology (WBE) approach largely relies on correlation between SARS-CoV-2 RNA concentration in wastewater (C_{RNA}) and relevant COVID-19 epidemiological parameters (cases). Through systematic review and meta-analysis, this study for the first time revealed that C_{RNA} exhibited better correlations with new cases (either daily new, weekly new, or future cases) than that of active or cumulative cases across WBE

studies and variations of environmental conditions, epidemiological conditions, and WBE sampling design could further impact such correlations and infection estimations. These findings provide suggestions and improvements for the precise estimation of COVID-19 case numbers through WBE.

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.

Data availability

Data will be made available on request.

Acknowledgements

This research was supported by the Australian Research Council Discovery Project (DP190100385) and a COVID-19 Digital Grant funded by Australian Academy of Science and the Department of Industry, Science, Energy and Resources through the Regional Collaborations Programme. Shuxin Zhang receives the support from a University of Wollongong Ph.D. scholarship. Samendra Sherchan was partially supported from the NIH grant R21AI157434.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jhazmat.2022.129848.

References

- Acosta, N., Bautista, M.A., Hollman, J., McCalder, J., Beaudet, A.B., Man, L., Waddell, B. J., Chen, J., Li, C., Kuzma, D., Bhatnagar, S., Leal, J., Meddings, J., Hu, J., Cabaj, J.L., Ruecker, N.J., Naugler, C., Pillai, D.R., Achari, G., Ryan, M.C., Conly, J.M., Frankowski, K., Hubert, C.R., Parkins, M.D., 2021. A multicenter study investigating SARS-CoV-2 in tertiary-care hospital wastewater: viral burden correlates with increasing hospitalized cases as well as hospital-associated transmissions and outbreaks. *Water Res.* 201.
- Agrawal, S., Orschler, L., Lackner, S., 2021. Long-term monitoring of SARS-CoV-2 RNA in wastewater of the Frankfurt metropolitan area in Southern Germany. *Sci. Rep.* 11 (1), 5372.
- Ahmed, W., Angel, N., Edson, J., Bibby, K., Bivins, A., O'Brien, J.W., Choi, P.M., Kitajima, M., Simpson, S.L., Li, J., 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.*, 138764.
- Ahmed, W., Bertsch, P.M., Bibby, K., Haramoto, E., Hewitt, J., Huygens, F., Gyawali, P., Korajkic, A., Riddell, S., Sherchan, S.P., Simpson, S.L., Sirikanchana, K., Symonds, E. M., Verhagen, R., Vasan, S.S., Kitajima, M., Bivins, A., 2020b. Decay of SARS-CoV-2 and surrogate murine hepatitis virus RNA in untreated wastewater to inform application in wastewater-based epidemiology. *Environ. Res.* 191.
- Ahmed, W., Bivins, A., Simpson, S.L., Bertsch, P.M., Ehret, J., Hosegood, L., Metcalfe, S. S., Smith, W.J.M., Thomas, K.V., Tynan, J., Mueller, J.F., 2022. Wastewater surveillance demonstrates high predictive value for COVID-19 infection on board repatriation flights to Australia. *Environ. Int.* 158, 106938.
- Ai, Y., Davis, A., Jones, D., Lemeshow, S., Tu, H., He, F., Ru, P., Pan, X., Bohrerova, Z., Lee, J., 2021. Wastewater SARS-CoV-2 monitoring as a community-level COVID-19 trend tracker and variants in Ohio, United States. *Sci. Total Environ.* 801, 149757.
- Alahdal, H.M., Ameen, F., AlYahya, S., Sonbol, H., Khan, A., Alsafayyan, Y., Alahmari, A., 2021. Municipal wastewater viral pollution in Saudi Arabia: effect of hot climate on COVID-19 disease spreading. *Environ. Sci. Pollut. Res.*
- Augusto, M.R., Claro, I.C.M., Siqueira, A.K., Sousa, G.S., Caldereiro, C.R., Duran, A.F.A., de Miranda, T.B., Bomediano Camillo, Ld.M., Cabral, A.D., de Freitas Bueno, R., 2022. Sampling strategies for wastewater surveillance: evaluating the variability of SARS-CoV-2 RNA concentration in composite and grab samples. *J. Environ. Chem. Eng.* 10 (3), 107478.
- Barrios, M.E., Díaz, S.M., Torres, C., Costamagna, D.M., Blanco Fernández, M.D., Mbayed, V.A., 2021. Dynamics of SARS-CoV-2 in wastewater in three districts of the Buenos Aires metropolitan region, Argentina, throughout nine months of surveillance: a pilot study. *Sci. Total Environ.* 800, 149578.
- Bertels, X., Demeyer, P., Van den Bogaert, S., Boogaerts, T., van Nuijs, A.L., Delputte, P., Lahousse, L., 2022. Factors influencing SARS-CoV-2 RNA concentrations in wastewater up to the sampling stage: a systematic review. *Sci. Total Environ.*, 153290.

- Betancourt, W.Q., Schmitz, B.W., Innes, G.K., Prasek, S.M., Pogreba Brown, K.M., Stark, E.R., Foster, A.R., Sprissler, R.S., Harris, D.T., Sherchan, S.P., Gerba, C.P., Pepper, I.L., 2021. COVID-19 containment on a college campus via wastewater-based epidemiology, targeted clinical testing and an intervention. *Sci. Total Environ.* 779.
- Bivins, A., Greaves, J., Fischer, R., Yinda, K.C., Ahmed, W., Kitajima, M., Munster, V.J., Bibby, K., 2020a. Persistence of SARS-CoV-2 in water and wastewater. *Environ. Sci. Technol. Lett.* 7 (12), 937–942.
- Bivins, A., North, D., Wu, Z., Shaffer, M., Ahmed, W., Bibby, K., 2021. Within- and between-day variability of SARS-CoV-2 RNA in municipal wastewater during periods of varying COVID-19 prevalence and positivity. *ACS EST Water* 1 (9), 2097–2108.
- Carrillo-Reyes, J., Barragán-Trinidad, M., Buitrón, G., 2021. Surveillance of SARS-CoV-2 in sewage and wastewater treatment plants in Mexico. *J. Water Process Eng.* 40.
- Chok, N.S., 2010. Pearson's versus Spearman's and Kendall's Correlation Coefficients for Continuous Data. University of Pittsburgh.
- Crank, K., Chen, W., Bivins, A., Lowry, S., Bibby, K., 2022. Contribution of SARS-CoV-2 RNA shedding routes to RNA loads in wastewater. *Sci. Total Environ.* 806, 150376.
- Curtis, K., Keeling, D., Yetka, K., Larson, A., Gonzalez, R., 2020. Wastewater SARS-CoV-2 concentration and loading variability from grab and 24-hour composite samples. *medRxiv*.
- Curtis, K., Keeling, D., Yetka, K., Larson, A., Gonzalez, R., 2021. Wastewater SARS-CoV-2 RNA Concentration and Loading Variability from Grab and 24-Hour Composite Samples. *medRxiv: the preprint server for health sciences*, 2020.2007.2010.20150607.
- D'Aoust, P.M., Graber, T.E., Mercier, E., Montpetit, D., Alexandrov, I., Neault, N., Baig, A.T., Mayne, J., Zhang, X., Alain, T., Servos, M.R., Srikanthan, N., MacKenzie, M., Figeys, D., Manuel, D., Jüni, P., MacKenzie, A.E., Delatolla, R., 2021a. Catching a resurgence: increase in SARS-CoV-2 viral RNA identified in wastewater 48h before COVID-19 clinical tests and 96h before hospitalizations. *Sci. Total Environ.* 770, 145319.
- D'Aoust, P.M., Mercier, E., Montpetit, D., Jia, J.J., Alexandrov, I., Neault, N., Baig, A.T., Mayne, J., Zhang, X., Alain, T., Langlois, M.A., Servos, M.R., MacKenzie, M., Figeys, D., MacKenzie, A.E., Graber, T.E., Delatolla, R., 2021b. Quantitative analysis of SARS-CoV-2 RNA from wastewater solids in communities with low COVID-19 incidence and prevalence. *Water Res.* 188, 116560.
- Despres, H.W., Mills, M.G., Shirley, D.J., Schmidt, M.M., Huang, M.-L., Jerome, K.R., Greninger, A.L., Bruce, E.A., 2021. Quantitative measurement of infectious virus in SARS-CoV-2 Alpha, Delta and Epsilon variants reveals higher infectivity (viral titer: RNA ratio) in clinical samples containing the Delta and Epsilon variants. *medRxiv: the preprint server for health sciences*, 2021.2009.2007.21263229.
- Dicker, R.C., Coronado, F., Koo, D., Parrish, R.G., 2006. Principles of epidemiology in public health practice; an introduction to applied epidemiology and biostatistics.
- van Doorn, A.S., Meijer, B., Frampton, C.M., Barclay, M.L., de Boer, N.K., 2020. Systematic review with meta-analysis: SARS-CoV-2 stool testing and the potential for faecal-oral transmission. *Alimentary pharmacology & therapeutics*.
- Feng, S., Roguet, A., McClary-Gutierrez, J.S., Newton, R.J., Kloczko, N., Meiman, J.G., McLellan, S.L., 2021. Evaluation of sampling, analysis, and normalization methods for SARS-CoV-2 concentrations in wastewater to assess COVID-19 Burdens in Wisconsin Communities. *ACS EST Water* 1 (8), 1955–1965.
- Fitzgerald, S.F., Rossi, G., Low, A.S., McAteer, S.P., O'Keefe, B., Findlay, D., Cameron, G. J., Pollard, P., Singleton, P.T.R., Ponton, G., Singer, A.C., Farkas, K., Jones, D., Graham, D.W., Quintela-Baluja, M., Tait-Burkard, C., Gally, D.L., Kao, R., Corbushley, A., 2021. Site specific relationships between COVID-19 cases and SARS-CoV-2 viral load in wastewater treatment plant influent. *Environ. Sci. Technol.* 55 (22), 15276–15286.
- George, A.D., Kaya, D., Layton, B.A., Bailey, K., Mansell, S., Kelly, C., Williamson, K.J., Radnicki, T.S., 2022. Impact of sampling type, frequency, and scale of the collection system on SARS-CoV-2 quantification fidelity. *Environ. Sci. Technol. Lett.* 9 (2), 160–165.
- Gerrity, D., Papp, K., Stoker, M., Sims, A., Fehner, W., 2021a. Early-pandemic wastewater surveillance of SARS-CoV-2 in Southern Nevada: methodology, occurrence, and incidence/prevalence considerations. *Water Res.* 10.
- Gibas, C., Lambirth, K., Mittal, N., Juel, M.A.I., Barua, V.B., Roppolo Brazell, L., Hinton, K., Lontai, J., Stark, N., Young, I., Quach, C., Russ, M., Kauer, J., Nicolosi, B., Chen, D., Akella, S., Tang, W., Schlueter, J., Munir, M., 2021. Implementing building-level SARS-CoV-2 wastewater surveillance on a university campus. *Sci. Total Environ.* 782.
- Giraud-Billoud, M., Cuervo, P., Altamirano, J.C., Pizarro, M., Aranibar, J.N., Catapano, A., Cuello, H., Maschessi, G., Vega, I.A., 2021. Monitoring of SARS-CoV-2 RNA in wastewater as an epidemiological surveillance tool in Mendoza, Argentina. *Sci. Total Environ.* 796.
- Graham, K.E., Loeb, S.K., Wolfe, M.K., Catoe, D., Sinnott-Armstrong, N., Kim, S., Yamahara, K.M., Sassoubre, L.M., Mendoza Grijalva, L.M., Roldan-Hernandez, L., Langenfeld, K., Wigginton, K.R., Boehm, A.B., 2021. SARS-CoV-2 RNA in wastewater settled solids is associated with COVID-19 cases in a large urban sewershed. *Environ. Sci. Technol.* 55 (1), 488–498.
- Hart, O.E., Halden, R.U., 2020a. 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.
- Hart, O.E., Halden, R.U., 2020b. Modeling wastewater temperature and attenuation of sewage-borne biomarkers globally. *Water Res.* 172, 115473.
- Hasell, J., Mathieu, E., Beltekian, D., Macdonald, B., Giattino, C., Ortiz-Ospina, E., Roser, M., Ritchie, H., 2020. A cross-country database of COVID-19 testing. *Sci. Data* 7 (1), 1–7.
- Havlicek, L.L., Peterson, N.L., 1976. Robustness of the Pearson correlation against violations of assumptions. *Percept. Mot. Skills* 43 (3 suppl), 1319–1334.
- He, J., Guo, Y., Mao, R., Zhang, J., 2021. Proportion of asymptomatic coronavirus disease 2019: a systematic review and meta-analysis. *J. Med. Virol.* 93 (2), 820–830.
- Heaton, K., Radvan, J., Cripps, H., Mountford, R., Braddon, F., Hughes, A., 1992. Defecation frequency and timing, and stool form in the general population: a prospective study. *Gut* 33 (6), 818–824.
- Hong, P.Y., Rachmadi, A.T., Mantilla-Calderon, D., Alkahtani, M., Bashawri, Y.M., Al Qarni, H., O'Reilly, K.M., Zhou, J., 2021. Estimating the minimum number of SARS-CoV-2 infected cases needed to detect viral RNA in wastewater: to what extent of the outbreak can surveillance of wastewater tell us? *Environ. Res.* 195.
- Huang, Y., Johnston, L., Parra, A., Sweeney, C., Hayes, E., Hansen, L.T., Gagnon, G., Stoddart, A., Jamieson, R., 2021. Detection of SARS-CoV-2 in wastewater in Halifax, Nova Scotia, Canada, using four RT-qPCR assays. *Facets* 6, 959–965.
- Jafferali, M.H., Khatami, K., Atasoy, M., Birgersson, M., Williams, C., Cetecioglu, Z., 2020. Benchmarking virus concentration methods for quantification of SARS-CoV-2 in raw wastewater. *Sci. Total Environ.*, 142939.
- Jiang, G., Wu, J., Weidhaas, J., Li, X., Chen, Y., Mueller, J., Li, J., Kumar, M., Zhou, X., Arora, S., Haramoto, E., Sherchan, S., Orive, G., Lertxundi, U., Honda, R., Kitajima, M., Jackson, G., 2022. Artificial neural network-based estimation of COVID-19 case numbers and effective reproduction rate using wastewater-based epidemiology. *Water Res.* 218, 118451.
- Jones, D.L., Baluja, M.Q., Graham, D.W., Corbushley, A., McDonald, J.E., Malham, S.K., Hillary, L.S., Connor, T.R., Gaze, W.H., Moura, I.B., Wilcox, M.H., Farkas, K., 2020. Shedding of SARS-CoV-2 in feces and urine and its potential role in person-to-person transmission and the environment-based spread of COVID-19. *Sci. Total Environ.* 749.
- Kitamura, K., Sadamasu, K., Muramatsu, M., Yoshida, H., 2021. Efficient detection of SARS-CoV-2 RNA in the solid fraction of wastewater. *Sci. Total Environ.* 763, 144587.
- Kulandaivelu, Jagadeeshkumar, Choi, P.M., Shrestha, S., Li, X., Song, Y., Li, J., Sharma, K., Yuan, Z., Mueller, J.F., Wang, C. and Jiang, G., 2020. Assessing the removal of organic micropollutants from wastewater by discharging drinking water sludge to sewers. *Water Research* 181 (115945). <https://doi.org/10.1016/j.watres.2020.115945>.
- Lazuka, A., Arnal, C., Soyeux, E., Sampson, M., Lepeuple, A.S., Deleuze, Y., Pouradier Duteil, S., Lacroix, S., 2021. COVID-19 wastewater based epidemiology: long-term monitoring of 10 WWTP in France reveals the importance of the sampling context. *Water Sci. Technol. J. Int. Assoc. Water Pollut. Res.* 84 (8), 1997–2013.
- Li, X., Kulandaivelu, J., Zhang, S., Shi, J., Sivakumar, M., Mueller, J., Luby, S., Ahmed, W., Coin, L., Jiang, G., 2021a. Data-driven estimation of COVID-19 community prevalence through wastewater-based epidemiology. *Sci. Total Environ.* 147947.
- Li, X., Zhang, S., Shi, J., Luby, S.P., Jiang, G., 2021b. Uncertainties in estimating SARS-CoV-2 prevalence by wastewater-based epidemiology. *Chem. Eng. J.* 415, 129039. Lausanne, Switzerland: 1996.
- Li, X., Kulandaivelu, J., Guo, Y., Zhang, S., Shi, J., O'Brien, J., Arora, S., Kumar, M., Sherchan, S.P., Honda, R., Jackson, G., Luby, S.P., Jiang, G., 2022. SARS-CoV-2 shedding sources in wastewater and implications for wastewater-based epidemiology. *J. Hazard. Mater.* 432, 128667.
- Mackey, K., Ayers, C.K., Kondo, K.K., Saha, S., Advani, S.M., Young, S., Spencer, H., Rusek, M., Anderson, J., Veazie, S., 2021. Racial and ethnic disparities in COVID-19-related infections, hospitalizations, and deaths: a systematic review. *Ann. Intern. Med.* 174 (3), 362–373.
- McCall, A.-K., Palmitezza, R., Blumensaat, F., Morgenroth, E., Ort, C., 2017. Modeling sewer transformations at catchment scale—implications on drug consumption estimates in wastewater-based epidemiology. *Water Res.* 122, 655–668.
- Medema, G., Heijnen, L., Elsinga, G., Italiaander, R., 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. *Environ. Sci. Technol. Lett.* 7 (7), 511–516.
- Miura, F., Kitajima, M., Omori, R., 2021. Duration of SARS-CoV-2 viral shedding in faeces as a parameter for wastewater-based epidemiology: Re-analysis of patient data using a shedding dynamics model. *Sci. Total Environ.* 769.
- Nagarkar, M., Keely, S.P., Jahne, M., Wheaton, E., Hart, C., Smith, B., Garland, J., Varughese, E.A., Braam, A., Wiechman, B., Morris, B., Brinkman, N.E., 2021. SARS-CoV-2 monitoring at three sewersheds of different scales and complexity demonstrates distinctive relationships between wastewater measurements and COVID-19 case data. *Sci. Total Environ.*, 151534.
- Nemudryi, A., Nemudraia, A., Wiegand, T., Surya, K., Buyukyoruk, M., Cicha, C., Vanderwood, K.K., Wilkinson, R., Wiedenheft, B., 2020. Temporal detection and phylogenetic assessment of SARS-CoV-2 in municipal wastewater. *Cell Rep. Med.* 1 (6), 100098.
- Oranization, W.H., 2020. COVID-19 - virtual press conference - 30 March 2020.
- Pan, Y., Zhang, D., Yang, P., Poon, L.L., Wang, Q., 2020. Viral load of SARS-CoV-2 in clinical samples. *Lancet Infect. Dis.* 20 (4), 411–412.
- Peccia, J., Zulli, A., Brackney, D.E., Grubaugh, N.D., Kaplan, E.H., Casanovas-Massana, A., Ko, A.I., Malik, A.A., Wang, D., Wang, M., Warren, J.L., Weinberger, D. M., Arnold, W., Omer, S.B., 2020. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nat. Biotechnol.* 38 (10), 1164–1167.
- Pecson, B.M., Darby, E., Haas, C., Amha, Y., Bartolo, M., Danielson, R., Dearborn, Y., Di Giovanni, G., Ferguson, C., Fevig, S., 2021. Reproducibility and sensitivity of 36 methods to quantify the SARS-CoV-2 genetic signal in raw wastewater: findings from an interlaboratory methods evaluation in the US. *Environ. Sci. Water Res. Technol.*
- Reese, H., Iuliano, A.D., Patel, N.N., Garg, S., Kim, L., Silk, B.J., Hall, A.J., Fry, A., Reed, C., 2020. Estimated Incidence of Coronavirus Disease 2019 (COVID-19) Illness and Hospitalization—United States, February–September 2020. *Clinical Infectious Diseases*.

- Róka, E., Khayer, B., Kis, Z., Kovács, L.B., Schuler, E., Magyar, N., Málnási, T., Oravecz, O., Pályi, B., Pándics, T., Vargha, M., 2021. Ahead of the second wave: early warning for COVID-19 by wastewater surveillance in Hungary. *Sci. Total Environ.* 786.
- Rusiñol, M., Martínez-Puchol, S., Forés, E., Itarte, M., Girones, R., Bofill-Mas, S., 2020. Concentration methods for the quantification of coronavirus and other potentially pandemic enveloped virus from wastewater. *Curr. Opin. Environ. Sci. Health* 17, 21–28.
- Rusiñol, M., Zammit, I., Itarte, M., Forés, E., Martínez-Puchol, S., Girones, R., Borrego, C., Corominas, L., Bofill-Mas, S., 2021. Monitoring waves of the COVID-19 pandemic: Inferences from WWTPs of different sizes. *Sci. Total Environ.* 787.
- Sangsanont, J., Rattanakul, S., Kongprajug, A., Chyerochana, N., Sresung, M., Sriporatana, N., Wanlapakorn, N., Poovorawan, Y., Mongkolsuk, S., Sirikanchana, K., 2021. SARS-CoV-2 RNA surveillance in large to small centralized wastewater treatment plants preceding the third COVID-19 resurgence in Bangkok, Thailand. *Sci. Total Environ.*, 151169.
- Scott, L.C., Aube, A., Babahaji, L., Vigil, K., Tims, S., Aw, T.G., 2021. Targeted wastewater surveillance of SARS-CoV-2 on a university campus for COVID-19 outbreak detection and mitigation. *Environ. Res.* 200.
- Shi, Jiahua, Li, X., Zhang, S., Sharma, E., Sivakumar, M., Sherchan, S. P., & Jiang, G., 2022. Enhanced decay of coronaviruses in sewers with domestic wastewater. *Science of The Total Environment* 813 (151919). <https://doi.org/10.1016/j.scitotenv.2021.151919>.
- Silverman, A.I., Boehm, A.B., 2020. Systematic review and meta-analysis of the persistence and disinfection of human coronaviruses and their viral surrogates in water and wastewater. *Environ. Sci. Technol. Lett.* 7 (8), 544–553.
- Street, R., Mathee, A., Mangwana, N., Dias, S., Sharma, J.R., Ramharack, P., Louw, J., Reddy, T., Brocker, L., Surujlal-Naicker, S., Berkowitz, N., Malema, M.S., Nkambule, S., Webster, C., Mahlangeni, N., Gelderblom, H., Mdhluli, M., Gray, G., Muller, C., Johnson, R., 2021. Spatial and temporal trends of SARS-CoV-2 RNA from wastewater treatment plants over 6 weeks in Cape Town, South Africa. *Int. J. Environ. Res. Public Health* 18, 22.
- Tomasino, M.P., Smedo, M., Vieira e Moreira, P., Ferraz, E., Rocha, A., Carvalho, M.F., Magalhães, C., Mucha, A.P., 2021. SARS-CoV-2 RNA detected in urban wastewater from Porto, Portugal: method optimization and continuous 25-week monitoring. *Sci. Total Environ.* 792.
- Weidhaas, J., Aanderud, Z.T., Roper, D.K., VanDerslice, J., Gaddis, E.B., Ostermiller, J., Hoffman, K., Jamal, R., Heck, P., Zhang, Y., Torgersen, K., Laan, J.V., LaCross, N., 2021. Correlation of SARS-CoV-2 RNA in wastewater with COVID-19 disease burden in sewersheds. *Sci. Total Environ.* 775.
- Westhaus, S., Weber, F.A., Schiwy, S., Linnemann, V., Brinkmann, M., Widera, M., Greve, C., Janke, A., Hollert, H., Wintgens, T., Ciesek, S., 2021. Detection of SARS-CoV-2 in raw and treated wastewater in Germany – suitability for COVID-19 surveillance and potential transmission risks. *Sci. Total Environ.* 751.
- WHO, 2009. Pandemic Influenza Preparedness and Response: A WHO Guidance Document. World Health Organization.
- Wong, J.C.C., Tan, J., Lim, Y.X., Arivalan, S., Hapuarachchi, H.C., Mailepessov, D., Griffiths, J., Jayarajah, P., Setoh, Y.X., Tien, W.P., Low, S.L., Koo, C., Yenamandra, S. P., Kong, M., Lee, V.J.M., Ng, L.C., 2021. Non-intrusive wastewater surveillance for monitoring of a residential building for COVID-19 cases. *Sci. Total Environ.* 786, 147419.
- Wu, F., Zhang, J., Xiao, A., Gu, X., Lee, W.L., Armas, F., Kauffman, K., Hanage, W., Matus, M., Ghaeli, N., Endo, N., Duvallet, C., Poyet, M., Moniz, K., Washburne, A.D., Erickson, T.B., Chai, P.R., Thompson, J., Alm, E.J., 2020. SARS-CoV-2 titers in wastewater are higher than expected from clinically confirmed cases. *mSystems* 5 (4), e00614–e00620.
- Wu, F., Xiao, A., Zhang, J., Moniz, K., Endo, N., Armas, F., Bushman, M., Chai, P.R., Duvallet, C., Erickson, T.B., Foppe, K., Ghaeli, N., Gu, X., Hanage, W.P., Huang, K.H., Lee, W.L., McElroy, K.A., Rhode, S.F., Matus, M., Wuertz, S., Thompson, J., Alm, E.J., 2021. Wastewater surveillance of SARS-CoV-2 across 40 U.S. states from February to June 2020. *Water Res.* 202.
- Wurtz, N., Lacoste, A., Jardot, P., Delache, A., Fontaine, X., Verlande, M., Annessi, A., Giraud-Gatineau, A., Chaudet, H., Fournier, P.E., Augier, P., La Scola, B., 2021. Viral RNA in city wastewater as a key indicator of COVID-19 recrudescence and containment measures effectiveness. *Front. Microbiol.* 12.
- Yearby, R., Mohapatra, S., 2020. Law, structural racism, and the COVID-19 pandemic. *J. Law Biosci.* 7, 1.
- Yue, S., Pilon, P., Phinney, B., Cavadias, G., 2002. The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrol. Process.* 16 (9), 1807–1829.