Title	Evaluating survey techniques in wastewater-based epidemiology for accurate COVID-19 incidence estimation
Author(s)	Murakami, Michio; Ando, Hiroki; Yamaguchi, Ryo et al.
Citation	Science of the Total Environment. 2024, 954, p. 176702
Version Type	VoR
URL	https://hdl.handle.net/11094/98589
rights	This article is licensed under a Creative Commons Attribution 4.0 International License.
Note	

Science of the Total Environment 954 (2024) 176702

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Evaluating survey techniques in wastewater-based epidemiology for accurate COVID-19 incidence estimation[☆]

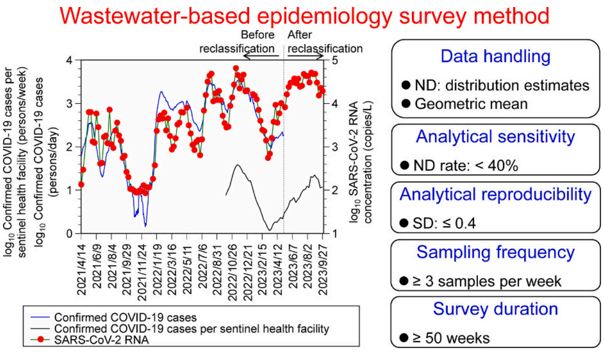
Michio Murakami [a,*], Hiroki Ando [b,c], Ryo Yamaguchi [d], Masaaki Kitajima [a,b,e]

[a] *Center for Infectious Disease Education and Research, Osaka University, 2-8 Yamadaoka, Suita-shi, Osaka 565-0871, Japan*
[b] *Division of Environmental Engineering, Faculty of Engineering, Hokkaido University, North 13 West 8, Kita-ku, Sapporo, Hokkaido 060-8628, Japan*
[c] *Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, AZ 85724, United States*
[d] *Public Health Office, City of Sapporo, West 19, Odori, Chuo-ku, Sapporo, Hokkaido 060-0042, Japan*
[e] *Research Center for Water Environment Technology, School of Engineering, The University of Tokyo, 2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan*

HIGHLIGHTS

- Wastewater-based epidemiology requires meeting good survey method criteria.
- Non-detect data should be replaced with estimates based on distribution.
- A non-detect rate is necessary to be <40 %.
- An analytical standard deviation needs to be ≤ 0.4.
- Three wastewater samples per week for fifty weeks are required.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Damià Barceló

Keywords:
Analytical sensitivity
Data handling
SARS-CoV-2
Sampling frequency
Wastewater surveillance
Wastewater-based epidemiological monitoring

ABSTRACT

Wastewater-based epidemiology (WBE) requires high-quality survey methods to determine the incidence of infections in wastewater catchment areas. In this study, the wastewater survey methods necessary for comprehending the incidence of infection by WBE are clarified. This clarification is based on the correlation with the number of confirmed coronavirus disease 2019 (COVID-19) cases, considering factors such as handling non-detect data, calculation method for representative values, analytical sensitivity, analytical reproducibility, sampling frequency, and survey duration. Data collected from 15 samples per week for two and a half years using a highly accurate analysis method were regarded as gold standard data, and the correlation between severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA concentrations in wastewater and confirmed COVID-19 cases was analyzed by Monte Carlo simulation under the hypothetical situation where the quality of the wastewater survey method was reduced. Regarding data handling, it was appropriate to replace non-detect data

---

[☆] This article has already been registered for Preprints on medRxiv. DOI is as follows: https://doi.org/10.1101/2024.06.09.24308677. Preliminary results were presented at the 58th Annual Conference of Japan Society on Water Environment 2024 (Murakami, M., Ando, H., Yamaguchi, R., Kitajima, M., 2024. Assessment of the investigation method of wastewater based epidemiology based on correlations with COVID 19 cases. In: The 58th Annual Conference of Japan Society on Water Environment 2024 (in Japanese)).

[*] Corresponding author.
E-mail address: michio@cider.osaka-u.ac.jp (M. Murakami).

https://doi.org/10.1016/j.scitotenv.2024.176702
Received 22 July 2024; Received in revised form 26 September 2024; Accepted 1 October 2024

with estimates based on distribution, and to use geometric means to calculate representative values. For the analysis of SARS-CoV-2 RNA in samples, using a highly sensitive and reproducible method (non-detect rates of <40 %; ≤0.4 standard deviation) and surveying at least three samples, preferably five samples, per week were considered desirable. Furthermore, conducting the survey over a period of time that included at least 50 weeks was necessary. A WBE that meets these survey criteria is sufficient for the determination of the COVID-19 infection incidence in the catchment. Furthermore, WBE can offer additional insights into infection rates in the catchment, such as the estimated 48 % decrease in confirmed COVID-19 cases visiting a clinic following a COVID-19 legal reclassification in Japan.

1. Introduction

Wastewater-based epidemiology (WBE), also known as wastewater surveillance, is an economical, representative, and early means of determining the incidence of infection in a target area without requiring personal information (Hart and Halden, 2020; Kitajima et al., 2020; Murakami et al., 2020; Shah et al., 2022). Since the beginning of the coronavirus disease 2019 (COVID-19) pandemic, numerous studies have been conducted on correlations between concentrations of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA in untreated wastewater and the infection incidence of wastewater catchment areas (Ahmed et al., 2020; La Rosa et al., 2020; Randazzo et al., 2020). Monitoring the infection incidence using WBE provides information that can contribute to public health, such as identifying the origin of infection, promoting infection control measures, including testing and vaccination campaigns in targeted areas, and preparing resource allocation in healthcare institutions (Betancourt et al., 2021; Karthikeyan et al., 2022; Kitajima et al., 2022; Klapsa et al., 2022; Li et al., 2023a).

In Japan, where notifiable disease surveillance (clinical surveillance that captures all reported cases of infection) of the number of COVID-19-infected individuals has been conducted, a high correlation (Pearson's $r = 0.94$) was reported between the number of newly confirmed COVID-19 cases and SARS-CoV-2 RNA concentrations in wastewater through a two-year sampling campaign with twice-weekly sample collection using an analytical method that is highly reproducible and capable of quantifying low concentrations of viral RNA (Ando et al., 2023). In contrast, several factors influence the correlation between the virus concentration in wastewater and the number of infected individuals in the catchment. These factors are categorized as clinical, environmental, and wastewater survey methods (Li et al., 2023b). The clinical factors include COVID-19 prevalence and testing coverage. Environmental factors include changes in air temperature and the catchment size of the wastewater treatment plants. Factors related to wastewater survey methods include the sampling frequency (i.e., the number of samples per week). Additionally, the survey duration and virus detection methods, such as analytical reproducibility and quantification at low virus concentrations (analytical sensitivity) in wastewater, are also relevant (Li et al., 2021; Medema et al., 2020). Li et al. (2023b) conducted a systematic review of the correlations between SARS-CoV-2 RNA concentrations in wastewater and infection incidence and reported that COVID-19 prevalence, testing coverage, air temperature variation, catchment size, and sampling frequency were more likely to be associated with the strength of the correlations, whereas the sampling method (i.e., grab or composite sampling) had a smaller effect.

The World Health Organization declared the end of the Public Health Emergency of International Concern on May 5, 2023. The number of infected individuals has shifted towards partial monitoring worldwide. Japan reclassified COVID-19 legal status into the fifth category of communicable diseases under the Communicable Diseases Law on May 8, 2023, and also shifted from notifiable disease surveillance to sentinel surveillance (clinical surveillance that partially captures reported cases of infection in a representative manner), in which only the designated health facilities report the number of COVID-19 cases seen there. As its clinical ability to ascertain the incidence of infection has declined, the utility of WBE has increased. Therefore, it is important to identify effective survey methods for the WBE that can adequately account for the number of infected individuals in the target areas. Among the aforementioned clinical factors, environmental factors, and wastewater survey methods, wastewater survey methods can be handled by the implementers of the WBE in a target area. However, no studies have comprehensively examined the appropriate wastewater survey methods in terms of sampling frequency, survey duration, analytical sensitivity, and reproducibility.

This study aimed to clarify the survey methods necessary for understanding the incidence of WBE infection based on the correlation with confirmed COVID-19 cases. First, the handling of non-detect data and the calculation of representative values were investigated. Second, the sampling frequency, survey duration, analytical sensitivity, and analytical reproducibility necessary to determine the infection incidence through WBE were analyzed. For the sampling frequency, the correlation between surveys conducted at two or three different catchments on the same day of the week and those conducted at the same catchment on two or three different days of the week was also examined. Additionally, changes in the relationship between virus concentrations and confirmed COVID-19 cases were analyzed, considering the behavior of individuals visiting health facilities before and after the legal reclassification of COVID-19 in Japan.

2. Methods

2.1. Data

The quality of wastewater surveys depends on factors related to sampling (e.g., sampling frequency, sampling duration) and virus analysis from wastewater samples (e.g., non-detect rate, standard deviation of the analysis). In this study, WBE measurements over two and a half years with high analytical accuracy and sampling frequency were used as "gold data," and the strength of the correlation coefficient with the confirmed COVID-19 cases was analyzed under the condition that the quality of wastewater survey methods would decline (e.g., reduced sampling frequency, short sampling durations, high non-detect rates, and high analytical standard deviations). WBE data obtained through a survey commissioned by the City of Sapporo and Hokkaido University were used in the analysis. Specifically, this study used data on SARS-CoV-2 RNA concentrations in untreated wastewater (24-hour composite samples) collected three times a week (Monday, Wednesday, and Friday in principle) at each of the five catchments covered by three adjacent wastewater treatment plants in the City of Sapporo from April 12, 2021, to September 29, 2023. Pepper mild mottle virus (PMMoV) RNA was also measured from September 27, 2021, to September 29, 2023. The population and area of the City of Sapporo were 1.96 million and 1121 km^2 in 2023, respectively. The population covered in each catchment, estimated from the overall facility's treated population and wastewater flow volume, ranged from approximately 164,700 to 246,300 individuals (Table S1), and the five catchments together covered 52 % of Sapporo's population. If fewer than 15 samples were collected per week, data from the corresponding week were excluded from the analysis.

Assuming a 100 % recovery efficiency throughout the process, the theoretical limit of detection (LOD) for SARS-CoV-2 RNA using the

Efficient and Practical virus Identification System with ENhanced Sensitivity for Solids (EPISENS-S)—the viral RNA detection method used in this study—was 93 copies/L (Ando et al., 2022). The EPISENS-S method has been employed by researchers other than the authors (Kuroiwa et al., 2023; Miyazawa et al., 2024; Okada and Nishiura, 2024). Briefly, EPISENS-S involves pelleting via low-speed centrifugation, direct RNA extraction, multiplex one-step RT-preamplification reaction, and quantitative real-time polymerase chain reaction (qPCR). By targeting all solid materials and performing pre-amplification, this method achieves an LOD approximately one-hundredth that of the common polyethylene glycol precipitation (PEG) method, although the SARS-CoV-2 RNA recovery efficiency, calculated from concentrations in wastewater seeded with heat-inactivated SARS-CoV-2 RNA, was slightly higher for the PEG method (5.9 %) than for EPISENS-S (1.1 %). The analytical reproducibility of SARS-CoV-2 RNA in the same wastewater samples showed standard deviations at $\log_{10}$ values of 0.03, 0.09, and 0.4 for concentrations of 4.11×10^6 copies/L, 4.11×10^5 copies/L, and 4.11×10^4 copies/L, respectively. Further details are provided in a previous study (Ando et al., 2022). Due to its high analytical reproducibility, one SARS-CoV-2 RNA analysis was performed per sample. A total of 15 weekly samples were analyzed over 122 weeks, with a total of 1830 samples. Of these, 157 samples were non-detect, while 44 were positive but below the LOD. In total, 201 samples were classified as below the LOD ("non-detect samples" hereinafter). Wastewater sample concentrations were converted to representative values using 1-week samples (Section 2.2). The standard deviation at $\log_{10}$ values for 15 samples from the same week was 0.40 (95 % confidence interval: 0.38–0.43), which closely matched the standard deviation for $\log_{10}$ values from 3 samples per catchment in the same week (Fig. S1).

Two types of clinical data were used for the confirmed COVID-19 cases: notifiable disease surveillance and sentinel surveillance. For the former, published data up to May 7, 2023, were used (City of Sapporo, 2024; Data-smart City Sapporo, 2023), and a moving average of the number of newly confirmed COVID-19 cases for a week was calculated using the three days before and after a representative date of wastewater sampling (arithmetic mean of sampling dates). During the notifiable disease surveillance period, from April 12, 2021 (the start of wastewater sampling) to May 7, 2023, 1515 wastewater samples were collected over 101 weeks. Regarding the sentinel surveillance data, the confirmed COVID-19 cases per sentinel health facility contained two types of data sources: estimation from the notifiable disease surveillance data from October 3, 2022, to May 7, 2023 and published data thereafter (City of Sapporo, 2024). Regarding the former, data representing the reported infection cases from designated health facilities were calculated as follows: the number of confirmed COVID-19 cases in notifiable disease surveillance during the epidemiological week was multiplied by 0.091891 and divided by 56 sentinel health facilities (Ministry of Health Labour and Welfare, 2023). From October 3, 2022, to September 29, 2023 (the end of wastewater sampling), 735 wastewater samples were collected over 49 weeks. Fig. 1 shows the temporal trends of SARS-CoV-2 RNA concentrations in wastewater and confirmed COVID-19 cases. The confirmed COVID-19 cases comprised five infection waves in two years for the notifiable disease surveillance data period and two infection waves in one year for the sentinel surveillance data period.

2.2. Treatment of non-detect data and calculation method of representative values (Preliminary analysis 1)

Our study confirmed that the SARS-CoV-2 RNA concentrations in the 1830 wastewater samples used were closer to a normal or log-normal distribution. Because the Shapiro–Wilk or Kolmogorov–Smirnov tests for normality are unsuitable for large sample sizes, Q-Q plots were used to assess the normality of SARS-CoV-2 RNA concentrations and their $\log_{10}$ values (Mishra et al., 2019), indicating that the $\log_{10}$ values are closer to normality (Fig. S2). In this study, Pearson's correlation coefficient r was calculated between the $\log_{10}$ values of representative SARS-CoV-2 RNA concentrations in wastewater for one week and the $\log_{10}$ values of the confirmed COVID-19 cases for one week (hereafter, a correlation represents that between SARS-CoV-2 RNA concentration in wastewater and the confirmed COVID-19 cases unless otherwise noted). To examine the treatment of non-detect data and the method for calculating representative values, the correlation coefficients were analyzed as follows (Table S2): First, the non-detect data were replaced with four types of values: (1) LOD, (2) LOD/2, (3) LOD/$\sqrt{2}$, and (4) the value corresponding to half of the non-detect rate using the estimated distribution (hereinafter, distribution estimates). The replacement values for (1)–(3) are often used, as previously reported (Croghan and Egeghy, 2003). The replacement value for (4) was calculated using the maximum likelihood method, assuming a lognormal distribution for all 1830 left-censored data points. The R package "NADA" was used to estimate the distribution (Lee, 2022; R Development Core Team, 2021). The estimated distribution using $\log_{10}$ SARS-CoV-2 RNA concentrations

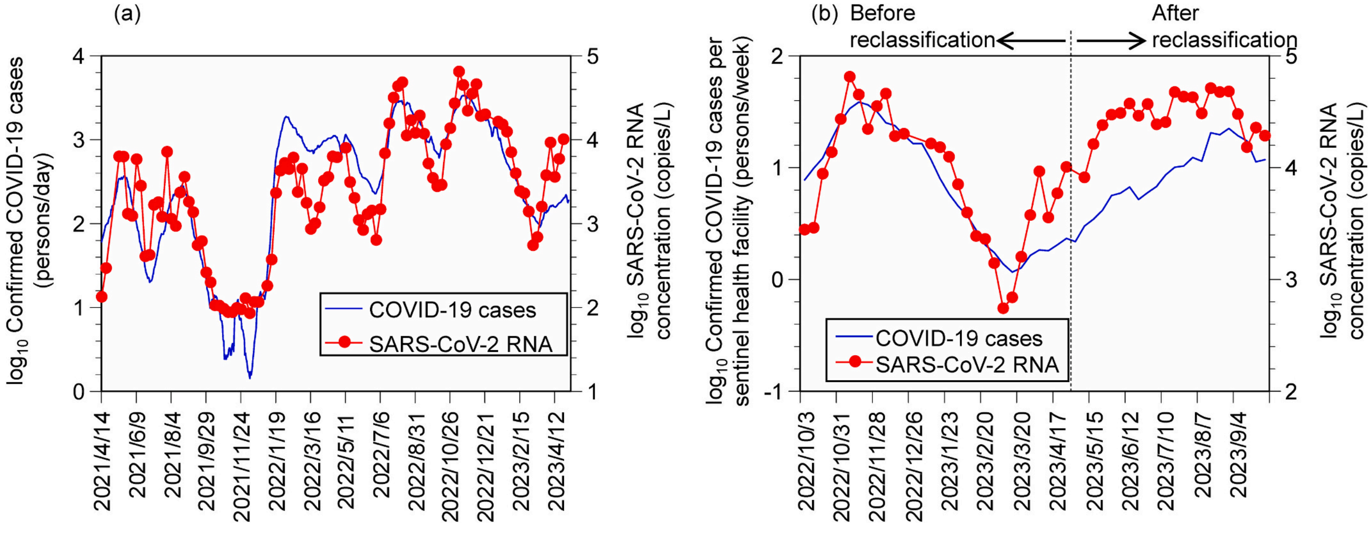


Fig. 1. Temporal changes of SARS-CoV-2 concentrations in wastewater and confirmed COVID-19 cases in the City of Sapporo. (a) Notifiable disease surveillance data; (b) sentinel surveillance data. The SARS-CoV-2 RNA concentration in wastewater was calculated based on the geometric mean values for one week (15 samples). The non-detect data were replaced using the distribution estimates. Notifiable disease surveillance: clinical surveillance that captures all reported cases of infection. Sentinel surveillance: clinical surveillance that captures a representative portion of reported infection cases.

was confirmed to align with the observed distribution (Fig. S3). The arithmetic mean, geometric mean, and median values were calculated for one week using the dataset with non-detect data replaced in this manner. Conditions with high correlation coefficients were extracted using a notifiable disease surveillance data set ($n = 1515$) with a range of −7 to +14 days (time lag) from the representative date of wastewater sampling to the published date of the confirmed COVID-19 cases. For this analysis, three types of corrections were calculated: (a) LOD = 93 copies/L without correction by PMMoV; (b) LOD = 93 copies/L with correction by PMMoV (using the ratio "[SARS-CoV-2 RNA concentration]/[PMMoV RNA concentration]" instead of SARS-CoV-2 RNA concentration); (c) LOD = 9300 copies/L and without correction by PMMoV; and (d) LOD = 9300 copies/L with correction by PMMoV. The LOD of 9300 copies/L was used because the LOD of the commonly used PEG method was nearly 100 times higher than that of the EPISENS-S method used in our study, as described in Section 2.1. Additionally, further analysis was conducted for three-day and two-week moving averages of the number of newly confirmed COVID-19 cases, instead of the one-week moving average. Other data treatment conditions included replacement through distribution estimates, geometric mean calculation, and no correction using PMMoV.

Based on this analysis, the following conditions were used to calculate correlation coefficients in subsequent analyses: replacement by distribution estimates, calculation of the geometric mean, no time lag (+0 days), and no correction using PMMoV (detailed in Section 3.1).

2.3. Correction of sentinel surveillance data (Preliminary analysis 2)

In the analysis using the sentinel surveillance dataset, the relationship between virus concentration and confirmed COVID-19 cases changed before and after the legal reclassification of COVID-19 in Japan, which might have resulted in behavioral changes in individuals visiting health facilities. Therefore, a multiple regression analysis was conducted with confirmed COVID-19 cases ($\log_{10}$) as the objective variable and SARS-CoV-2 RNA concentration ($\log_{10}$) and a dummy variable for reclassification as explanatory variables. The dummy variable was set to 0 for the period before reclassification and 1 for the period after reclassification. IBM SPSS version 28 was used for all the analyses. Based on these results, in subsequent analyses, the confirmed COVID-19 cases per sentinel medical ($\log_{10}$) after the reclassification was treated as a corrected value by doing +0.32 (detailed in Section 3.2).

2.4. Assessment of survey methods based on correlations with the confirmed COVID-19 cases

Main analyses 1–4 were conducted to determine the sampling frequency, survey duration, analytical sensitivity, and analytical reproducibility required to adequately explain the number of infected individuals.

2.4.1. Main analysis 1: analytical sensitivity and sampling frequency

As shown in Section 2.1, the LOD in our wastewater analysis was 93 copies/L, approximately 100 times lower than that of the common PEG method. An exploratory simulation was conducted to estimated how the correlation coefficient would change if a higher LOD were applied in wastewater analysis. Previous studies utilized exploratory simulations using different LODs (Croghan and Egeghy, 2003), and similar simulations with varying detection rates have also been employed (Kato et al., 2013).

Conditions with LODs of 93, 186, 465, 930, 1860, 4650, 9300, 18,600, and 46,500 copies/L (corresponding to 1, 2, 5, 10, 20, 50, 100, 200, and 500 times the actual values in the dataset, respectively) were established. Hypothetically, by resetting these LODs, the data were treated as non-detect (for example, data detected as 100 copies/L were considered non-detect under the condition of an LOD of 186 copies/L). For each condition, as described in Section 2.2, replacement values based on distribution estimation for left-censored data were used as non-detect data (Table S3).

Furthermore, under these LOD conditions, 1–15 samples were assumed to be collected weekly. This corresponds to a hypothetical reduction in the total sampling frequency to 15 samples per week. The catchments and days of the week to be sampled were randomly determined according to the sampling frequencies and fixed for the duration of the WBE survey (same as in Main analysis 2 to 4).

2.4.2. Main analysis 2: multiple catchments on the same day of the week and the same catchment on multiple days of the week

As noted in Section 2.1, three samples per week were collected in each of the five catchments (15 samples per week). An exploratorily simulation was conducted to analyze correlation coefficients when one sample per week was taken at three catchments and three samples were taken at one catchment. Similarly, the correlation coefficients were compared between one sample per week at two catchments and two samples per week at one catchment. The catchments and days of the week were randomly determined according to the sampling frequencies from the five catchments or three days of the week.

2.4.3. Main analysis 3: survey duration and sampling frequency

Different survey durations of 101, 12, 25, 50, and 75 weeks were established from a notifiable disease surveillance dataset. The first week of the survey was randomly selected, and consecutive weeks were selected. The sampling frequency was analyzed from 1 to 15, as in the Main analysis. Only notifiable disease surveillance data were used in this analysis.

2.4.4. Main analysis 4: analytical reproducibility and sampling frequency

The analysis assumed conditions under which the analytical reproducibility of SARS-CoV-2 RNA concentrations varied. In accordance with a previous study (Ando et al., 2022), the analytical standard deviation of this dataset was conservatively set at 0.4 at $\log_{10}$ values, consistent with 0.03, 0.09, or 0.4 observed in wastewater samples from the same catchment (details in Section 2.1). Since the number of analyses in this dataset was one for each sample, the "true value" was estimated from the normal distribution when the "measured value" at $\log_{10}$ values was taken as the arithmetic mean and the standard deviation as 0.4. The hypothetically measured values were then estimated with standard deviations of 0.4, 0.6, 0.8, and 1, respectively. Simulations employing a unified analytical standard deviation were previously used in a study (Li et al., 2021). The sampling frequency was analyzed from 1 to 15, as in the Main analysis 1.

Under all conditions, the correlation coefficient between the SARS-CoV-2 RNA concentration in wastewater and the confirmed COVID-19 cases for one week was calculated using Monte Carlo simulations with 10,000 iterations for each. Oracle Crystal Ball version 11.1.2.4.900 was used for analysis. The arithmetic means and 95 % uncertainty intervals calculated from the 2.5th and 97.5th percentile values among the 10,000 iterations were used. The conditions were extracted such that the 2.5th percentile value was >0.7.

3. Results

3.1. Treatment of non-detect data and determination of a method for calculating representative values

Fig. 2 presents the results of the Preliminary analysis 1. With regard to the treatment of non-detect data, high correlation coefficients were obtained when the data were replaced with distribution estimates, or LOD/2. The correlation coefficients were high when the geometric mean, or median, was used to calculate representative values. Regardless of the type of moving averages for newly confirmed COVID-19 cases, the correlation coefficient remained ≥ 0.85 with a time gap ranging from −7 to +6 days (Fig. S4). PMMoV correction did not significantly affect

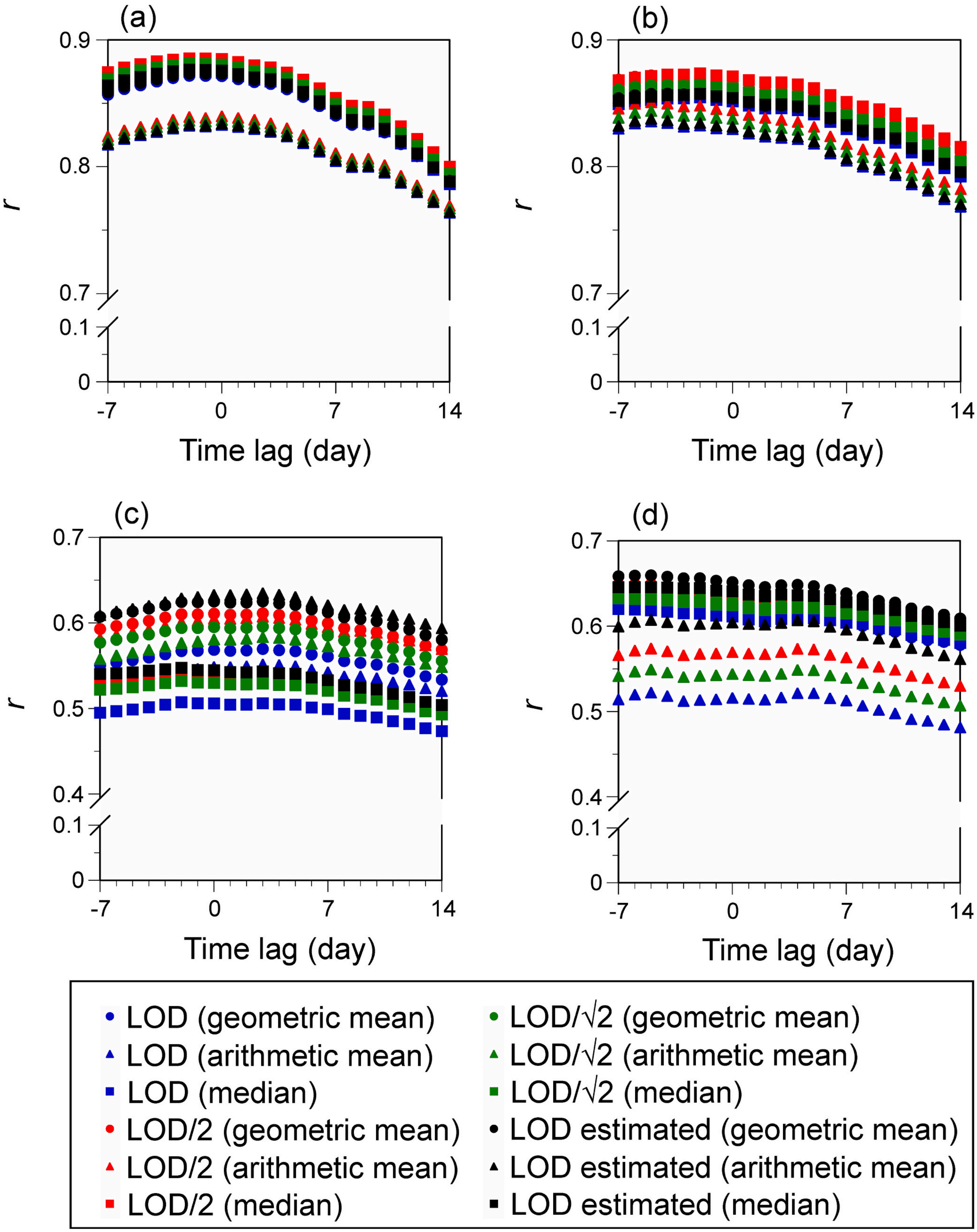


Fig. 2. Comparison of Pearson's correlation coefficients (*r*) based on treatment of non-detect data and different methods of calculating representative values. (a) Without PMMoV correction, limit of detection (LOD) = 93 copies/L; (b) with PMMoV correction, LOD = 93 copies/L; (c) without PMMoV correction, LOD = 9300 copies/L; and (d) with PMMoV correction, LOD = 9300 copies/L. Time lag represents the "published date of the number of infected individuals" minus the "representative date of wastewater sampling." LOD/2, LOD/$\sqrt{2}$, and LOD estimated represent the replacement of LOD values by LOD/2, LOD/$\sqrt{2}$, and distribution estimates, respectively.

the correlation coefficients.

Fig. 3(a) shows a scatter plot of SARS-CoV-2 RNA concentrations and confirmed COVID-19 cases in the notifiable disease surveillance under the conditions of replacement by distribution estimates, calculation of geometric mean, no gap (+0 days), and no correction of SARS-CoV-2 RNA concentrations by PMMoV. The correlation coefficient was as high as 0.87, and the slope of the regression equation using the logarithm of both SARS-CoV-2 RNA concentrations and confirmed COVID-19 cases was 0.96, which is close to 1, confirming a linear relationship between them.

3.2. *Determination of correction values for sentinel surveillance data*

Multiple regression analysis (Preliminary analysis 2) showed that the unstandardized partial regression coefficients (95 % confidence interval) for SARS-CoV-2 RNA concentration and reclassification were 0.80 (0.63–0.97) and −0.32 (−0.50 to −0.14), both significant. A scatter plot of SARS-CoV-2 RNA concentrations and corrected confirmed COVID-19 cases is shown in Fig. 3(b). The correlation coefficient and slope of the regression equation were 0.86 and 0.80, respectively, confirming a strong linear relationship between SARS-CoV-2 RNA concentrations in wastewater and confirmed COVID-19 cases, even in the sentinel surveillance data.

3.3. *Effects of sampling frequency, survey duration, analytical sensitivity, and analytical reproducibility on the correlation with the confirmed COVID-19 cases*

The results of Main analysis 1 showed that the correlation coefficients decreased as the sampling frequencies decreased for both notifiable disease surveillance and sentinel surveillance data (Fig. 4). Large differences in correlation coefficients were found between one and two samples per week, and three samples per week showed high correlation coefficients similar to those of more frequent sampling under the measurement condition of LOD = 93 copies/L (2.5th percentile values: 0.80 for notifiable disease surveillance and 0.76 for sentinel surveillance). The correlation coefficients decreased as the LOD value increased (i.e., as the non-detect rate increased): in the notifiable disease surveillance data, there was no large difference in the correlation coefficients for the measurement conditions with LOD between 93 and 930 copies/L (non-detect rate: 13–31 %), but the correlation coefficients decreased as the LOD increased from 1860 copies/L (non-detect rate: 42 %). Fifteen samples per week, with an LOD of 1860 copies/L, did not show a higher correlation than two samples per week, with an LOD of 93 copies/L. In the sentinel surveillance data, no large differences in the correlation coefficients were found between 93 and 9300 copies/L for the LOD (non-detect rate: 0.4–34 %), but the correlation coefficients decreased when the LOD was >18,600 copies/L (non-detect rate: 54 %). For both notifiable disease surveillance and sentinel surveillance data, the 2.5th percentile values of the correlation coefficients exceeded 0.7 for the three samples per week when the non-detect rate was <40 %.

Regarding Main analysis 2, no large differences in the correlation coefficients were found between multiple catchments on the same day of the week and the identical catchment on multiple days of the week (Fig. 5), although slightly less variation in the correlation coefficients existed for two or three samples per week in the same catchment than for sampling once per week in two or three catchments. The 2.5th percentile values of the correlation coefficients exceeded 0.7 for both sampling once at three catchments per week and sampling three times at one catchment per week.

In Main analysis 3, the 12-week surveys had lower correlation coefficients and wider uncertainty intervals (Fig. 6). Compared to the 50-week or longer surveys, the 25-week surveys had similar arithmetic mean correlation coefficients but lower 2.5th percentile values. The 2.5th percentile values in the 50-week or longer surveys with three or more samples per week exceeded 0.7.

In Main analysis 4, as the standard deviation increased, the correlation coefficients decreased (Fig. 7). In particular, the uncertainty interval widened with an increase in standard deviation when the number of samples per week was small. In the notifiable disease surveillance data, the 2.5th percentile values of the correlation coefficients exceeded 0.7 when the standard deviation was 0.4, with three or more samples per week. In the sentinel surveillance data, the arithmetic mean value of the correlation coefficient was below 0.7 for three samples per week. The 2.5th percentile value of the correlation coefficient was 0.66 and 0.71 for the five- and seven-weekly surveys, respectively.

4. Discussion

Using high analytical accuracy (LOD = 93 copies/L) and intensive surveys (15 samples per week for 122 weeks) as the gold standard for WBE, this study identified the treatment of non-detect data and the appropriate method for calculating representative values, and analyzed the impact of sampling frequency, survey duration, analytical sensitivity and analytical reproducibility on the correlation between SARS-CoV-2 RNA concentration in wastewater and confirmed COVID-19 cases.

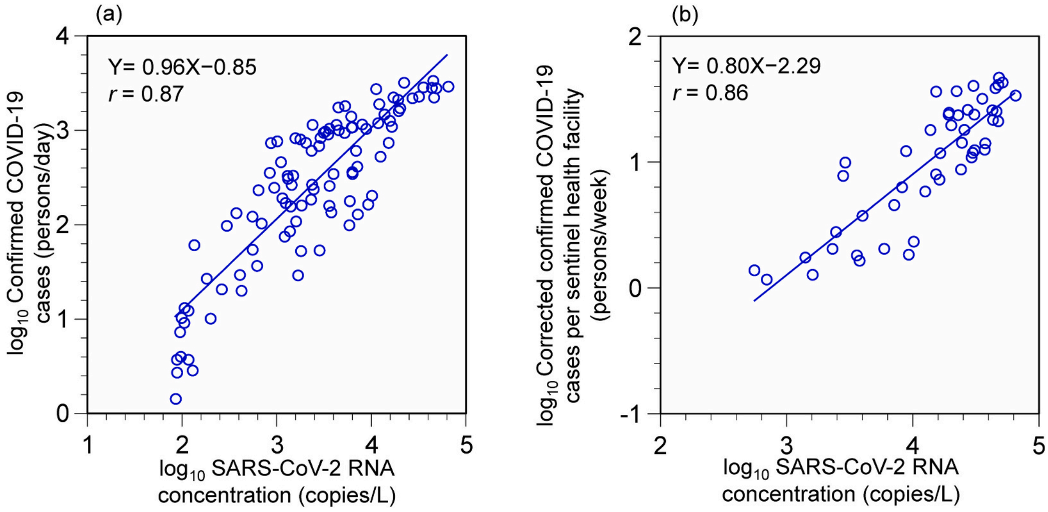


Fig. 3. Scatterplots of SARS-CoV-2 RNA concentration vs. confirmed COVID-19 cases. (a) Notifiable disease surveillance data; (b) sentinel surveillance data. *r*: Pearson correlation coefficients. Notifiable disease surveillance: clinical surveillance that captures all reported cases of infection. Sentinel surveillance: clinical surveillance that captures a representative portion of reported infection cases.

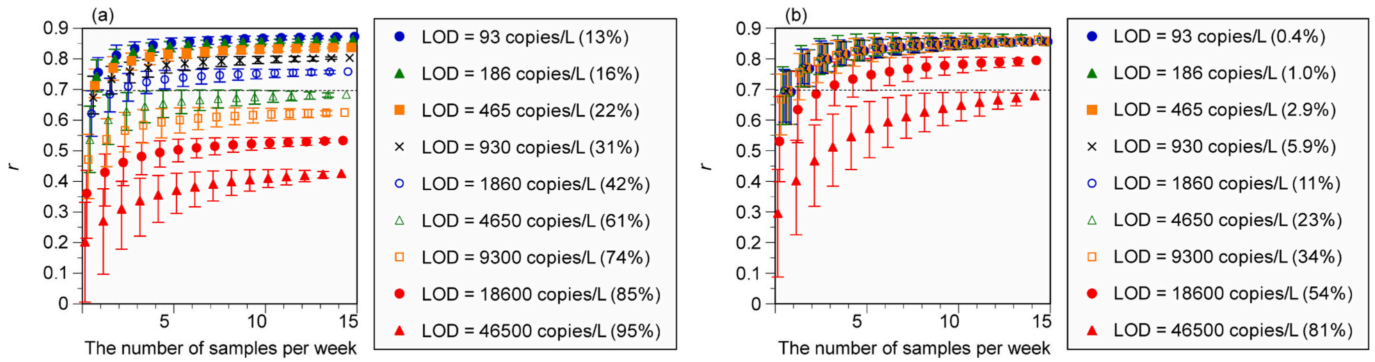


Fig. 4. Pearson correlation coefficients (*r*) between SARS-CoV-2 RNA concentrations and the confirmed COVID-19 cases according to the number of samples per week and analytical sensitivity. (a) Notifiable disease surveillance data; (b) sentinel surveillance data. LOD: limit of detection. The number in parenthesis represents the non-detect rate. An error bar represents a 95 % uncertainty interval. Notifiable disease surveillance: clinical surveillance that captures all reported cases of infection. Sentinel surveillance: clinical surveillance that captures a representative portion of reported infection cases.

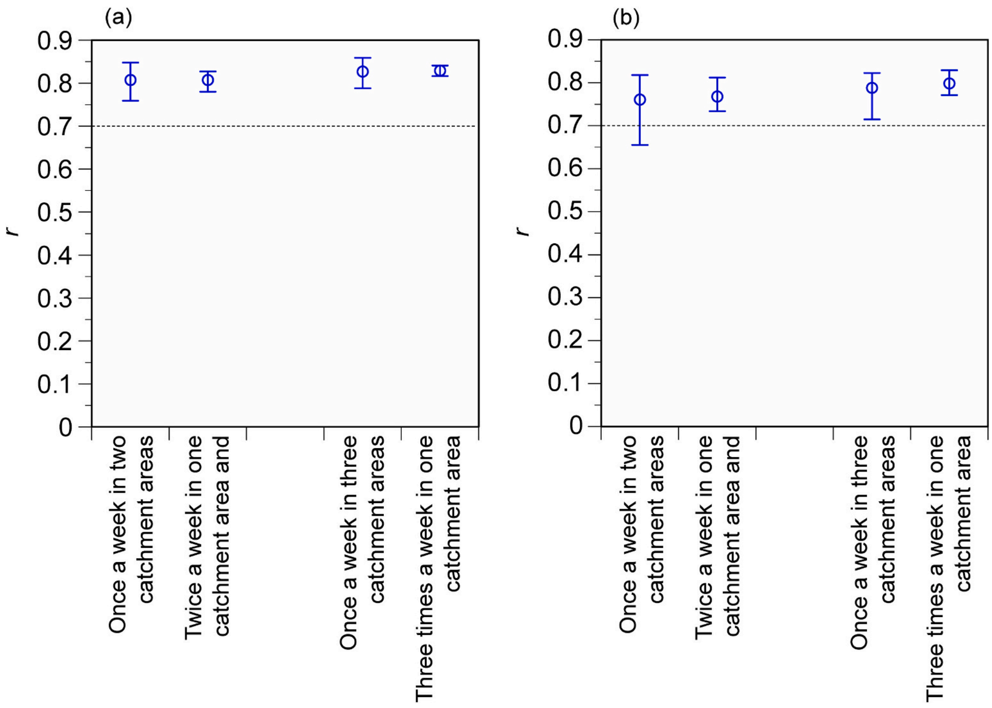


Fig. 5. Comparison of surveys at multiple catchments on the same day of the week and at the same catchment on multiple days of the week. (a) Notifiable disease surveillance data; (b) sentinel surveillance data. The *r* represents the Pearson correlation coefficient between SARS-CoV-2 RNA concentrations and the confirmed COVID-19 cases. An error bar represents a 95 % uncertainty interval. Notifiable disease surveillance: clinical surveillance that captures all reported cases of infection. Sentinel surveillance: clinical surveillance that captures a representative portion of reported infection cases.

First, a preliminary analysis identified the appropriate treatment for non-detect data and the appropriate method for calculating representative values. The validity of treating non-detect data using distribution estimates has been discussed in a previous report (Croghan and Egeghy, 2003). In this study, correlation coefficients were calculated by log-transforming virus concentrations in wastewater and confirmed COVID-19 cases, a method used in previous studies (Rabe et al., 2023; Xiao et al., 2022). Therefore, it is reasonable to use geometric means or medians rather than arithmetic means to calculate representative values. The PMMoV correction did not improve the correlation with confirmed COVID-19 cases, which is consistent with the findings of our previous study (Ando et al., 2023). This aligns with previous studies suggesting that PMMoV correction does not consistently improve the correlation (Islam et al., 2024; Maal-Bared et al., 2023), although conflicting results were reported in another study, which indicated that PMMoV correction enhanced the correlation (Zhan et al., 2023). The present study performed the main analyses with a time lag of 0 days, which did not negate the early detectability of WBE. In this study, the data were merged into one-week data to calculate representative values, which indicated that it was not necessary to consider the time lag. No substantial differences were identified in the correlation coefficients across a time gap of −7 to +6 days. The early detection of COVID-19 by the WBE in the same catchment in this study has been demonstrated in detail in our previous study, which revealed that SARS-CoV-2 RNA in wastewater samples can be detected 5 days before newly confirmed COVID-19 cases are reported (Ando et al., 2023). Therefore, our preliminary analysis supports the role of WBE in early detection.

When non-detect data were treated in this manner and representative values were calculated, strong correlations between SARS-CoV-2 RNA concentrations in wastewater and confirmed COVID-19 cases were observed (notifiable disease surveillance: $r = 0.87$, sentinel surveillance: $r = 0.86$). The slope with log-transformed data was almost 1,

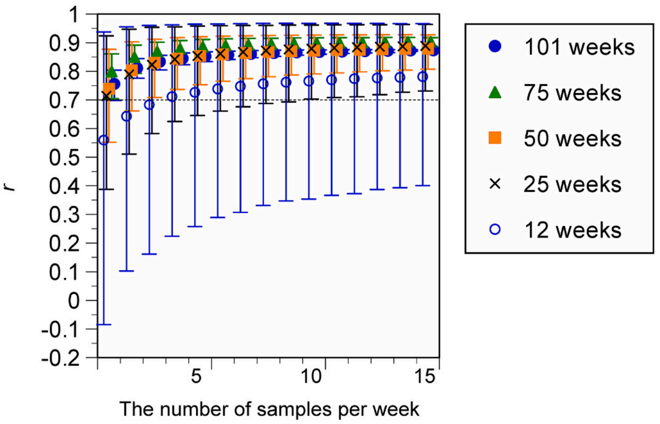


Fig. 6. Pearson correlation coefficients (*r*) between SARS-CoV-2 RNA concentrations and the confirmed COVID-19 cases according to the number of samples per week and survey duration. (a) Notifiable disease surveillance data; (b) sentinel surveillance data. An error bar represents a 95 % uncertainty interval. Notifiable disease surveillance: clinical surveillance that captures all reported cases of infection. Sentinel surveillance: clinical surveillance that captures a representative portion of reported infection cases.

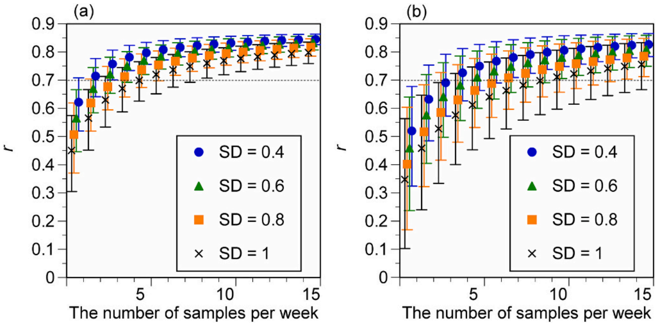


Fig. 7. Pearson correlation coefficients (*r*) between SARS-CoV-2 RNA concentrations and the confirmed COVID-19 cases according to the number of samples per week and analytical reproducibility. (a) Notifiable disease surveillance data; (b) sentinel surveillance data. SD: standard deviation. An error bar represents a 95 % uncertainty interval. Notifiable disease surveillance: clinical surveillance that captures all reported cases of infection. Sentinel surveillance: clinical surveillance that captures a representative portion of reported infection cases.

confirming that the WBE was sufficient to determine the COVID-19 infection incidence in the catchment. Interestingly, sentinel surveillance data showed a 0.32 decrease in confirmed COVID-19 cases at $\log_{10}$ values after reclassification. No major change in the prevalence of SARS-CoV-2 variants was observed during this period (Our World in Data, 2024). This suggests that the reclassification of COVID-19 led to a change in individuals' healthcare-seeking behaviors, such as hesitation in receiving medical examinations, and that the number of confirmed COVID-19 cases captured by clinics decreased by half (i.e., $10^{-0.32} = 48$ %). This result is consistent with US observations, where the ratio of reported cases to wastewater SARS-CoV-2 RNA load during a wave following the end of the Public Health Emergency of International Concern (PHEIC) was lower than during the Omicron wave under the PHEIC (Li et al., 2024). Our study highlights that the WBE can provide a good picture of the incidence of infection in catchments, even when changes in people's consultation behavior occur.

Regarding the analytical sensitivity, a decrease in the correlation coefficients was observed with increasing LOD values (Main analysis 1). In particular, both notifiable disease surveillance and sentinel surveillance data showed large decreases in correlation coefficients when the non-detect rate exceeded 40 %. Therefore, it is desirable to conduct the analysis with a sensitivity that the non-detect rate does not exceed 40 %. Fig. 4a and b shows different non-detect rates for the same LOD. Even with an LOD as high as 9300 copies/L, sentinel surveillance data showed a correlation coefficient of 0.86 for 15 samples per week. The non-detect rate depends not only on the LOD but also on the incidence of infection. This result indicates that it is desirable to use a dataset with a non-detect rate of <40 % (which is more achievable with high-sensitivity methods) to discuss correlations with COVID-19 cases. Ando et al. (2023) reported that a 50 % probability of detection corresponded to 0.69 out of 100,000 confirmed COVID-19 cases per day when the EPISENS for membrane (EPISENS-M) method with the LOD of 43.9 copies/L was used for SARS-CoV-2 RNA detection from wastewater. Although counting positive replicates in qPCR is an alternative in cases of high non-detect rates, numerous measurements are required for each sample, complicating the analysis (Zhao et al., 2023). Our study highlights the benefits of methods with low LOD for determining the low infection incidence in catchments. Application of pre-amplification prior to qPCR in a previous study, inspired by the EPISENS-S method, successfully increased the detection rate of mpox virus in wastewater (Bowes et al., 2023). As highly sensitive analytical methods are developed (El soufi et al., 2024), the findings of this study will serve as a reference for determining the infection incidence.

Regarding the sampling frequency, large differences in the correlation coefficients existed between one and two or more samples per week (main analyses 1, 3, and 4). This is consistent with the findings in previous studies demonstrating that at least 2–5 samples per week were needed to determine COVID-19 infection incidence in catchments (Chan et al., 2023; Corrin et al., 2024; Kuroita et al., 2024). The utility of twice-weekly sampling has been reported elsewhere (Okada and Nishiura, 2024). In particular, for both comprehensive notifiable and sentinel surveillance data, a sampling frequency of three or more times per week achieved a 2.5th percentile value of correlation coefficients >0.7 when the non-detect rate was <40 % (Main analysis 1). With three samples per week, no large difference in the correlation coefficients existed between three-day sampling in the same catchment and one-day sampling in three different catchments (Main analysis 2). When the standard deviation of the analysis was 0.4 (Main analysis 4), it was considered possible to survey at least five samples per week given that the correlation coefficient was low for the three-weekly surveys in the sentinel surveillance data.

Regarding survey duration, the correlation coefficient was notably low at 12 weeks (Main Analysis 3). At three samples per week, the 2.5th percentile values of the correlation coefficients exceeded 0.7 at 50 weeks or more. To determine the incidence of COVID-19 infection in the catchment by the WBE, a survey period that included 50 weeks or more would be necessary. This number of "50 weeks" as survey duration may depend on the number of infection waves as well as the number of data plots to discuss the correlation between virus concentration in wastewater and infection incidence. During the notifiable disease surveillance period, approximately two infection waves were observed over 50 weeks. High correlation coefficients were also observed in other analyses using sentinel surveillance data (49 weeks with two infection waves: Preliminary analysis 1 and Main analysis 1).

Regarding analytical reproducibility (Main analysis 4), when the standard deviations of the analysis were large, the uncertainty intervals were wide, particularly for a small number of samples. The finding that analytical standard deviation was the primary factor in estimating infectious disease incidence aligns with a previous study (Li et al., 2021). The 2.5th percentile values of the correlation coefficients were below 0.7 when the standard deviation was 0.6 or more for the three samples per week. A standard deviation of 0.4 or less was considered desirable in terms of analytical reproducibility.

Overall, it is considered desirable to use an analytical method that can quantify SARS-CoV-2 RNA in wastewater samples with high detectability and reproducibility (non-detect rate: <40 %; standard deviation: ≤0.4) and to survey at least three samples per week, preferably five or more samples, for 50 weeks or more.

This study has some limitations: First, although this study focused on wastewater survey methods, it did not examine clinical factors (e.g., COVID-19 prevalence or testing coverage) or environmental factors (e.g., air temperature or catchment population). Second, among the wastewater survey methods, this study did not examine factors that affect virus recovery rates during the analytical process, such as polymerase chain reaction inhibition. Third, the findings of this study were based on the City of Sapporo, and there is room for further research on the applicability of these findings to other catchments. The analyses examined in this study are expected to expand to various catchments as demonstrated in previous studies conducted in other countries (Duvallet et al., 2022; Krogsgaard et al., 2024; Tiwari et al., 2022), and the accumulation and integration of results will increase the generality of the findings.

5. Conclusions

The use of the WBE is sufficient to determine the incidence of COVID-19 in catchments. Furthermore, the WBE can present additional informational value with respect to understanding the infection incidence of a catchment, as estimated by the 48 % reduction in confirmed COVID-19 cases visiting health facilities after the reclassification of COVID-19 in Japan.

By examining the correlation between SARS-CoV-2 RNA concentrations in wastewater and confirmed COVID-19 cases under hypothetical conditions in which the quality of wastewater survey methods has declined, this study identified WBE survey methods that are necessary for understanding the infection situation in a catchment. The findings of the appropriate WBE survey methods obtained in this study are as follows:

- Non-detect data should be replaced by distribution estimates (or LOD/2).
- The geometric mean (or median) should be used to calculate representative values.
- A quantifiable and highly reproducible method (non-detect rate: $<$40 %; standard deviation: $\leq$0.4) is necessary for the analysis of SARS-CoV-2 RNA in samples.
- The sampling frequency required is at least three samples per week, preferably five samples per week.
- Surveys need to be conducted for a period of time that includes at least 50 weeks or longer.

Statements

During the preparation of this manuscript, the authors used DeepL solely for the purpose of the possible improvement of English language expression. The authors created the original texts before using this tool. The authors reviewed and edited the content as needed, after using this tool. Furthermore, the paper was carefully edited by native professional editors. The authors take full responsibility for the content of the publication.

CRediT authorship contribution statement

Michio Murakami: Writing – original draft, Visualization, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Hiroki Ando:** Writing – review & editing, Resources. **Ryo Yamaguchi:** Writing – review & editing, Resources. **Masaaki Kitajima:** Writing – review & editing, Resources, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Michio Murakami reports financial support was provided by The Nippon Foundation - Osaka University. Masaaki Kitajima reports financial support was provided by Japan Agency for Medical Research and Development. Masaaki Kitajima reports financial support was provided by Japan Science and Technology Agency. Michio Murakami reports a relationship with NJS CO LTD that includes: consulting or advisory. Masaaki Kitajima reports a relationship with AdvanSentinel that includes: funding grants and speaking and lecture fees. Masaaki Kitajima reports a relationship with Shimadzu Corporation that includes: funding grants. Masaaki Kitajima reports a relationship with Shionogi and Co Ltd that includes: funding grants and speaking and lecture fees. Masaaki Kitajima has patent pending to Shionogi & Co., Ltd. If there are other authors, they 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

We would like to thank Editage (www.editage.com) for English language editing. This work was supported by "The Nippon Foundation - Osaka University Project for Infectious Disease Prevention," the Japan Agency for Medical Research and Development (AMED) under grant number 24fk0108713h0001, and the Japan Science and Technology Agency (JST) through the JST-Mirai Program, under grant number JPMJMI22D1. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2024.176702.

References

Ahmed, W., Angel, N., Edson, J., Bibby, K., Bivins, A., O'Brien, J.W., et al., 2020. First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: a proof of concept for the wastewater surveillance of COVID-19 in the community. Sci. Total Environ. 728, 138764.

Ando, H., Iwamoto, R., Kobayashi, H., Okabe, S., Kitajima, M., 2022. The Efficient and Practical virus Identification System with ENhanced Sensitivity for Solids (EPISENS-S): a rapid and cost-effective SARS-CoV-2 RNA detection method for routine wastewater surveillance. Sci. Total Environ. 843, 157101.

Ando, H., Murakami, M., Ahmed, W., Iwamoto, R., Okabe, S., Kitajima, M., 2023. Wastewater-based prediction of COVID-19 cases using a highly sensitive SARS-CoV-2 RNA detection method combined with mathematical modeling. Environ. Int. 173, 107743.

Betancourt, W.Q., Schmitz, B.W., Innes, G.K., Prasek, S.M., Pogreba Brown, K.M., Stark, E.R., et al., 2021. COVID-19 containment on a college campus via wastewater-based epidemiology, targeted clinical testing and an intervention. Sci. Total Environ. 779, 146408.

Bowes, D.A., Henke, K.B., Driver, E.M., Newell, M.E., Block, I., Shaffer, G., et al., 2023. Enhanced detection of mpox virus in wastewater using a pre-amplification approach: a pilot study informing population-level monitoring of low-titer pathogens. Sci. Total Environ. 903, 166230.

Chan, E.M.G., Kennedy, L.C., Wolfe, M.K., Boehm, A.B., 2023. Identifying trends in SARS-CoV-2 RNA in wastewater to infer changing COVID-19 incidence: effect of sampling frequency. PLOS Water 2, e0000088.

City of Sapporo, 2024. https://www.city.sapporo.jp/hokenjo/f1kansen/2019n-covhassei.html (accsessed May 9, 2024). (in Japanese).

Corrin, T., Rabeenthira, P., Young, K.M., Mathiyalagan, G., Baumeister, A., Pussegoda, K., et al., 2024. A scoping review of human pathogens detected in untreated human wastewater and sludge. J. Water Health 22, 436–449.

Croghan, C.W., Egeghy, P.P., 2003. Methods of dealing with values below the limit of detection using SAS. https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NERL&dirEntryId=64046 (accsessed May 9, 2024).

Data-smart City Sapporo, 2023. https://ckan.pf-sapporo.jp/dataset/covid_19_patients (accsessed May 9, 2024). (in Japanese).

Duvallet, C., Wu, F., McElroy, K.A., Imakaev, M., Endo, N., Xiao, A., et al., 2022. Nationwide trends in COVID-19 cases and SARS-CoV-2 RNA wastewater concentrations in the United States. ACS ES&T Water 2, 1899–1909.

El soufi, G., Di Jorio, L., Gerber, Z., Cluzel, N., Van Assche, J., Delafoy, D., et al., 2024. Highly efficient and sensitive membrane-based concentration process allows quantification, surveillance, and sequencing of viruses in large volumes of wastewater. Water Res. 249, 120959.

Hart, O.E., Halden, R.U., 2020. Computational analysis of SARS-CoV-2/COVID-19 surveillance by wastewater-based epidemiology locally and globally: feasibility, economy, opportunities and challenges. Sci. Total Environ. 730, 138875.

Islam, G., Gedge, A., Ibrahim, R., de Melo, T., Lara-Jacobo, L., Dlugosz, T., et al., 2024. The role of catchment population size, data normalization, and chronology of public health interventions on wastewater-based COVID-19 viral trends. Sci. Total Environ. 937, 173272.

Karthikeyan, S., Levy, J.I., De Hoff, P., Humphrey, G., Birmingham, A., Jepsen, K., et al., 2022. Wastewater sequencing reveals early cryptic SARS-CoV-2 variant transmission. Nature 609, 101–108.

Kato, T., Miura, T., Okabe, S., Sano, D., 2013. Bayesian modeling of enteric virus density in wastewater using left-censored data. Food Environ. Virol. 5, 185–193.

Kitajima, M., Ahmed, W., Bibby, K., Carducci, A., Gerba, C.P., Hamilton, K.A., et al., 2020. SARS-CoV-2 in wastewater: state of the knowledge and research needs. Sci. Total Environ. 739, 139076.

Kitajima, M., Murakami, M., Kadoya, S.S., Ando, H., Kuroita, T., Katayama, H., et al., 2022. Association of SARS-CoV-2 load in wastewater with reported COVID-19 cases in the Tokyo 2020 Olympic and Paralympic Village from July to September 2021. JAMA Netw. Open 5, e2226822.

Klapsa, D., Wilton, T., Zealand, A., Bujaki, E., Saxentoff, E., Troman, C., et al., 2022. Sustained detection of type 2 poliovirus in London sewage between February and July, 2022, by enhanced environmental surveillance. Lancet 400, 1531–1538.

Krogsgaard, L.W., Benedetti, G., Gudde, A., Richter, S.R., Rasmussen, L.D., Midgley, S.E., et al., 2024. Results from the SARS-CoV-2 wastewater-based surveillance system in Denmark, July 2021 to June 2022. Water Res. 252, 121223.

Kuroita, T., Yoshimura, A., Iwamoto, R., Ando, H., Okabe, S., Kitajima, M., 2024. Quantitative analysis of SARS-CoV-2 RNA in wastewater and evaluation of sampling frequency during the downward period of a COVID-19 wave in Japan. Sci. Total Environ. 906, 166526.

Kuroiwa, M., Gahara, Y., Kato, H., Morikawa, Y., Matsui, Y., Adachi, T., et al., 2023. Targeted amplicon sequencing of wastewater samples for detecting SARS-CoV-2 variants with high sensitivity and resolution. Sci. Total Environ. 893, 164766.

La Rosa, G., Iaconelli, M., Mancini, P., Bonanno Ferraro, G., Veneri, C., Bonadonna, L., et al., 2020. First detection of SARS-CoV-2 in untreated wastewaters in Italy. Sci. Total Environ. 736, 139652.

Lee, L., 2022. Package 'NADA'. https://cran.r-project.org/web/packages/NADA/NADA.pdf (accsessed May 9, 2024).

Li, X., Zhang, S., Shi, J., Luby, S.P., Jiang, G., 2021. Uncertainties in estimating SARS-CoV-2 prevalence by wastewater-based epidemiology. Chem. Eng. J. 415, 129039.

Li, X., Liu, H., Gao, L., Sherchan, S.P., Zhou, T., Khan, S.J., et al., 2023a. Wastewater-based epidemiology predicts COVID-19-induced weekly new hospital admissions in over 150 USA counties. Nat. Commun. 14, 4548.

Li, X., Zhang, S., Sherchan, S., Orive, G., Lertxundi, U., Haramoto, E., et al., 2023b. Correlation between SARS-CoV-2 RNA concentration in wastewater and COVID-19 cases in community: a systematic review and meta-analysis. J. Hazard. Mater. 441, 129848.

Li, L., Haak, L., Carine, M., Pagilla, K.R., 2024. Temporal assessment of SARS-CoV-2 detection in wastewater and its epidemiological implications in COVID-19 case dynamics. Heliyon 10, e29462.

Maal-Bared, R., Qiu, Y., Li, Q., Gao, T., Hrudey, S.E., Bhavanam, S., et al., 2023. Does normalization of SARS-CoV-2 concentrations by Pepper Mild Mottle Virus improve correlations and lead time between wastewater surveillance and clinical data in Alberta (Canada): comparing twelve SARS-CoV-2 normalization approaches. Sci. Total Environ. 856, 158964.

Medema, G., Been, F., Heijnen, L., Petterson, S., 2020. Implementation of environmental surveillance for SARS-CoV-2 virus to support public health decisions: opportunities and challenges. Curr. Opin. Environ. Sci. Health 17, 49–71.

Ministry of Health Labour and Welfare, 2023. https://www.mhlw.go.jp/content/001065724.pdf (accsessed May 9, 2024). (in Japanese).

Mishra, P., Pandey, C.M., Singh, U., Gupta, A., Sahu, C., Keshri, A., 2019. Descriptive statistics and normality tests for statistical data. Ann. Card. Anaesth. 22, 67–72.

Miyazawa, S., Wong, T.S., Ito, G., Iwamoto, R., Watanabe, K., van Boven, M., et al., 2024. Wastewater-based reproduction numbers and projections of COVID-19 cases in three areas in Japan, November 2021 to December 2022. Euro Surveill. 29 pii=2300277.

Murakami, M., Hata, A., Honda, R., Watanabe, T., 2020. Letter to the editor: wastewater-based epidemiology can overcome representativeness and stigma issues related to COVID-19. Environ. Sci. Technol. 54, 5311.

Okada, Y., Nishiura, H., 2024. Estimating the effective reproduction number of COVID-19 from population-wide wastewater data: an application in Kagawa, Japan. Infect. Dis. Model. 9, 645–656.

Our World in Data, 2024. SARS-CoV-2 variants in analyzed sequences. https://ourworldindata.org/grapher/covid-variants-area?country=~JPN (accsessed May 13, 2024).

R Development Core Team, 2021. R 4.2.0. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rabe, A., Ravuri, S., Burnor, E., Steele, J.A., Kantor, R.S., Choi, S., et al., 2023. Correlation between wastewater and COVID-19 case incidence rates in major California sewersheds across three variant periods. J. Water Health 21, 1303–1317.

Randazzo, W., Truchado, P., Cuevas-Ferrando, E., Simón, P., Allende, A., Sánchez, G., 2020. SARS-CoV-2 RNA in wastewater anticipated COVID-19 occurrence in a low prevalence area. Water Res. 181, 115942.

Shah, S., Gwee, S.X.W., Ng, J.Q.X., Lau, N., Koh, J., Pang, J., 2022. Wastewater surveillance to infer COVID-19 transmission: a systematic review. Sci. Total Environ. 804, 150060.

Tiwari, A., Lipponen, A., Hokajärvi, A.-M., Luomala, O., Sarekoski, A., Rytkönen, A., et al., 2022. Detection and quantification of SARS-CoV-2 RNA in wastewater influent in relation to reported COVID-19 incidence in Finland. Water Res. 215, 118220.

Xiao, A., Wu, F., Bushman, M., Zhang, J., Imakaev, M., Chai, P.R., et al., 2022. Metrics to relate COVID-19 wastewater data to clinical testing dynamics. Water Res. 212, 118070.

Zhan, Q., Solo-Gabriele, H.M., Sharkey, M.E., Amirali, A., Beaver, C.C., Boone, M.M., et al., 2023. Correlative analysis of wastewater trends with clinical cases and hospitalizations through five dominant variant waves of COVID-19. ACS ES&T Water 3, 2849–2862.

Zhao, B., Fujita, T., Nihei, Y., Yu, Z., Chen, X., Tanaka, H., et al., 2023. Tracking community infection dynamics of COVID-19 by monitoring SARS-CoV-2 RNA in wastewater, counting positive reactions by qPCR. Sci. Total Environ. 904, 166420.