Delft University of Technology

Capturing the SARS-CoV-2 infection pyramid within the municipality of Rotterdam using longitudinal sewage surveillance

de Graaf, Miranda; Langeveld, Jeroen; Post, Johan; Carrizosa, Christian; Franz, Eelco; Izquierdo-Lara, Ray W.; Elsinga, Goffe; Heijnen, Leo; Medema, Gertjan; More Authors

DOI
10.1016/j.scitotenv.2023.163599

Publication date
2023

Document Version
Final published version

Published in
Science of the Total Environment

Citation (APA)
de Graaf, M., Langeveld, J., Post, J., Carrizosa, C., Franz, E., Izquierdo-Lara, R. W., Elsinga, G., Heijnen, L., Medema, G., & More Authors (2023). Capturing the SARS-CoV-2 infection pyramid within the municipality of Rotterdam using longitudinal sewage surveillance. *Science of the Total Environment*, *883*, Article 163599. https://doi.org/10.1016/j.scitotenv.2023.163599

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

https://www.openaccess.nl/en/you-share-we-take-care

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

Science of the Total Environment 883 (2023) 163599

Contents lists available at ScienceDirect

Science of the Total Environment

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

Capturing the SARS-CoV-2 infection pyramid within the municipality of Rotterdam using longitudinal sewage surveillance

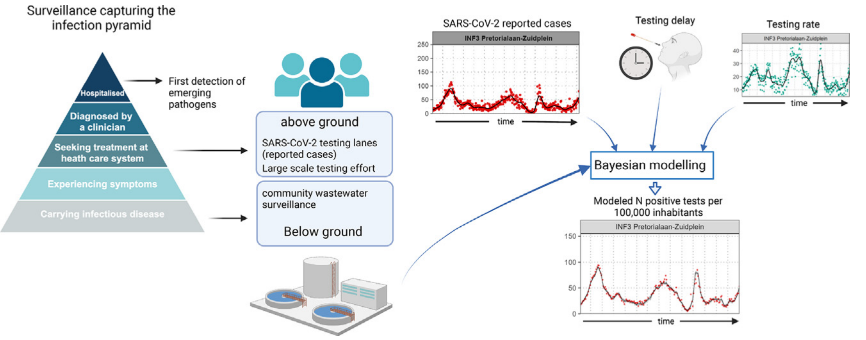
Miranda de Graaf[a,b,1,*], Jeroen Langeveld[c,h,1], Johan Post[c,1], Christian Carrizosa[d,i], Eelco Franz[d], Ray W. Izquierdo-Lara[a], Goffe Elsinga[e], Leo Heijnen[e], Frederic Been[e], Janko van Beek[a], Remy Schilperoort[c], Rianne Vriend[f], Ewout Fanoy[f], Evelien I.T. de Schepper[g], Marion P.G. Koopmans[a,b,2], Gertjan Medema[b,e,h,2]

[a] *Department of Viroscience, Erasmus University Medical Center, Rotterdam, the Netherlands*
[b] *Pandemic and Disaster Preparedness Centre Rotterdam and Delft, the Netherlands*
[c] *Partners4urbanwater, Nijmegen, the Netherlands*
[d] *Centre for Infectious Disease Control, National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands*
[e] *KWR Water Research Institute, Groningenhaven 7, 3433 PE Nieuwegein, the Netherlands*
[f] *Regional Public Health Service Rotterdam-Rijnmond, Rotterdam, the Netherlands*
[g] *Department of General Practice, Erasmus University Medical Center, Rotterdam, the Netherlands*
[h] *Delft University of Technology, Stevinweg 1, 2628 CN Delft, the Netherlands*
[i] *Department of Medical Genetics, Oslo University Hospital, Oslo, Norway.*

HIGHLIGHTS

- The incidence of positive coronavirus tests can be modeled based on sewage data.
- Normalization for testing delay and intensity of clinical data is needed for these models
- Shedding levels around symptom onset largely determine coronavirus levels in sewage.
- Sewage can accurately display coronavirus dynamics for small and large locations.
- Many SARS-CoV-2 cases go unreported, but can be measured through sewage surveillance

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Warish Ahmed

Keywords:
Public health
SARS-COV-2
Sewage
Surveillance
Early warning systems

ABSTRACT

Despite high vaccination rates in the Netherlands, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) continues to circulate. Longitudinal sewage surveillance was implemented along with the notification of cases as two parts of the surveillance pyramid to validate the use of sewage for surveillance, as an early warning tool, and to measure the effect of interventions.

Sewage samples were collected from nine neighborhoods between September 2020 and November 2021. Comparative analysis and modeling were performed to understand the correlation between wastewater and case trends. Using high resolution sampling, normalization of wastewater SARS-CoV-2 concentrations, and 'normalization' of reported positive tests for testing delay and intensity, the incidence of reported positive tests could be modeled based on sewage data, and trends in both surveillance systems coincided. The high collinearity implied that high levels of viral shedding around the onset of disease largely determined SARS-CoV-2 levels in wastewater, and that the observed relationship was independent of variants of concern and vaccination levels. Sewage surveillance alongside a large-scale testing

---

* Corresponding author at: Department of Viroscience, Erasmus University Medical Center, Rotterdam, the Netherlands.
E-mail address: m.degraaf@erasmusmc.nl (M. de Graaf).

[1] These authors contributed equally to this article.

[2] These senior authors contributed equally to this article.

http://dx.doi.org/10.1016/j.scitotenv.2023.163599
Received 12 December 2022; Received in revised form 7 April 2023; Accepted 15 April 2023
Available online 25 April 2023
0048-9697/© 2023 Published by Elsevier B.V.

effort where 58 % of a municipality was tested, indicated a five-fold difference in the number of SARS-CoV-2-positive individuals and reported cases through standard testing.

Where trends in reported positive cases were biased due to testing delay and testing behavior, wastewater surveillance can objectively display SARS-CoV-2 dynamics for both small and large locations and is sensitive enough to measure small variations in the number of infected individuals within or between neighborhoods. With the transition to a post-acute phase of the pandemic, sewage surveillance can help to keep track of re-emergence, but continued validation studies are needed to assess the predictive value of sewage surveillance with new variants. Our findings and model aid in interpreting SARS-CoV-2 surveillance data for public health decision-making and show its potential as one of the pillars of future surveillance of (re)emerging viruses.

1. Introduction

The coronavirus disease 2019 pandemic is a major challenge for society, and to control outbreaks the largest global testing program in history was started (Mercer and Salit, 2021). For most pathogens, there are differences in the severity of disease between infected individuals, with a small group experiencing symptoms that are so severe that hospital admission is warranted, while a much larger group experience mild symptoms or undergo asymptomatic infection (Sah et al., 2021; The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team, 2020). Since only people that seek medical care will be diagnosed, an outbreak can occur relatively unnoticed for some time (Munster et al., 2020; Zhu et al., 2020). SARS-CoV-2 is an example of a viral pathogen for which the disease pyramid is heavily skewed towards mild and asymptomatic cases (Munster et al., 2020).

Wastewater can be used for population surveillance of viruses that may spread by persons with mild or asymptomatic disease, and has been used for common endemic pathogens, antimicrobial resistance (Hellmer et al., 2014) and poliovirus (Hendriksen et al., 2019; Nieuwenhuijse et al., 2020). SARS-CoV-2 can be detected in domestic wastewater, and the levels of SARS-CoV-2 in wastewater correlate with reported cases and hospitalizations (Barrios et al., 2021; Fernandez-Cassi et al., 2021). In the Netherlands, during the spring of 2020, SARS-CoV-2 was already detectable in city wastewater six days before the first patients were diagnosed (Medema et al., 2020b). Our work and that of others showed that next-generation sequencing (NGS) and digital droplet PCR (ddPCR) can be used to test for the presence of SARS-CoV-2 variants of concern (VOC) and to unravel the genetic characteristics of SARS-CoV-2 viruses present in sewage (Heijnen et al., 2021; Izquierdo-Lara et al., 2021).

In other countries, wastewater surveillance also outperformed case reports as a tool to monitor SARS-CoV-2 circulation, where the detection of SARS-CoV-2 in wastewater preceded case reports by 3 to 14 days (Barrios et al., 2021; Claro et al., 2021; Fernandez-Cassi et al., 2021; Wu et al., 2021). The added value of sewage surveillance, depends on multiple factors, such as testing access and policy, healthcare access, demographics, and compliance to policies at the community level. Therefore, we conducted a high-resolution surveillance project in the Rotterdam Rijnmond area to assess how the dynamics of the number of clinical cases within a neighborhood are reflected in viral loads in sewage and which factors affect this correlation. We further investigated the potential of sewage surveillance to monitor SARS-CoV-2 trends, as an early warning tool, and to measure the effect of interventions. For this, we designed a study to collect the following data from different layers of the surveillance pyramid: 1) high-resolution sewage sampling in nine neighborhoods of different sizes and socioeconomic status within Rotterdam; 2) notification data based on clinical testing, with number of notifications, dates of probable day of onset of disease, and dates of test result from the national notification database for community cases in the same neighborhoods (OSIRIS); 3) data on total number of tests performed in the testing lanes of the regional Public Health Service (PHS). All data sources were matched to each neighborhood based on zip codes. During this study, sewage surveillance was carried out in real time and shared with municipal public health experts to assess the usefulness of this type of data for decision making on SARS-CoV-2 mitigation measures.

2. Methods

2.1. Selection of areas

For this observational study, nine catchment areas each representing between 6500 and 138,280 inhabitants were included. We first selected four neighborhoods for which the coverage area of general practices overlapped well with the sewer catchment areas: Katendrecht (6500 inhabitants), Ommoord (28,434 inhabitants), Pretorialaan (71,325 inhabitants), and Rozenburg (12,374 inhabitants). In addition, we selected four larger catchment areas: influent (INF)2 Everlo-Waalhaven (27,044 inhabitants), INF3 Pretorialaan-Zuidplein (121,118 inhabitants), INF4 Wolphaertsbocht (36,011 inhabitants) and INF5–6 Heemraadsplein (138,280 inhabitants) for which we could collect sewage samples at the wastewater treatment plant (WWTP) in Dokhaven (Figure1A). The sewer systems of Katendrecht, Pretorialaan and INF3 Pretorialaan-Zuidplein are nested; they form a cascade where upstream catchments discharge into larger downstream catchments. We later included Bergschenhoek (18,750 inhabitants) during a large-scale testing effort.

2.2. Sampling and analysis

Automated samplers were installed, and from September 1, 2020 to November 9, 2021 (except for Rozenburg where it was until October 13, 2021, and Ommoord where it was until September 6, 2021), wastewater specimens were collected three times per week as 24-h flow-dependent composite samples from each catchment. Later, during a large-scale testing effort to investigate an outbreak of the alpha VOC, we included the Bergschenhoek area in the sampling process from January 11, 2021 to October 13, 2021. The samples were processed as previously described (Langeveld et al., 2023; Medema et al., 2020b). RT-qPCR screening was performed on the samples using primer and probe sets targeting the SARS-CoV-2 N2 and *E*-gene as well as crAssphage, as previously described (Medema et al., 2020b; Stachler et al., 2017).

2.3. Normalization sewage

The sewage in Rotterdam consists of domestic wastewater from households, industrial wastewater, extraneous waters, such as infiltrating groundwater or inflowing surface water, and runoff. All non-domestic wastewater flows can dilute SARS-CoV-2 levels and vary strongly in time and per catchment, therefore flow normalization was applied to enable comparison over time and between catchments. The flow normalization process of these samples is described in Langeveld et al. (2023). The samples were normalized based on the quotient of the measured daily volumes of sewage and the expected amount of domestic wastewater (the average volume of domestic wastewater produced per person and day multiplied by the population per sewer district). This normalization was verified using the conductivity of the sewage and levels of crAssphage (an indicator of human fecal contamination) and erroneous data points were removed (Langeveld et al., 2023). The total inflow at WWTP Dokhaven was calculated as the weighted mean flow of the four larger catchment areas.

2.4. PHS SARS-CoV-2 test results

The regional PHS provided free SARS-CoV-2 testing during the pandemic in the Netherlands. Starting on June 1, 2020, testing capacity was expanded. All individuals experiencing symptoms, including mild symptoms, were eligible for testing. From December 1, 2020 onwards, the scope of testing was widened to include close contacts of positive cases who did not exhibit any symptoms. As the pandemic progressed, negative tests were required for travel and events, and self-testing became more common. Consequently, testing for SARS-CoV-2 varied across locations and over time (Fig. 1B). The testing process involved trained personnel at the PHS testing lanes obtaining a nasopharyngeal/throat swab, followed by either an RT-qPCR (Corman et al., 2020) or rapid antigen test.

We determined all zip codes within each catchment area, hereby including 382,011 inhabitants. Next, we extracted all negative and positive SARS-COV-2 test results from PHS testing lanes based on the residential address's zip code. We included all test results between September 1, 2020 and November 9, 2021, except for samples from the Bergschenhoek area, for which we included those from January 11, 2021 to November 9, 2021. In total, 360,086 test results (positive and negative) were included in this study. The process of establishing the large-scale testing initiative in Lansigerland, which encompasses Bergschenhoek, is outlined in detail by van Beek et al. (van Beek et al., 2022). We obtained the onset dates of the disease from the national notification database for community cases (OSIRIS). The Erasmus MC Medical Ethical Committee approved this study (MEC-2020-0617).

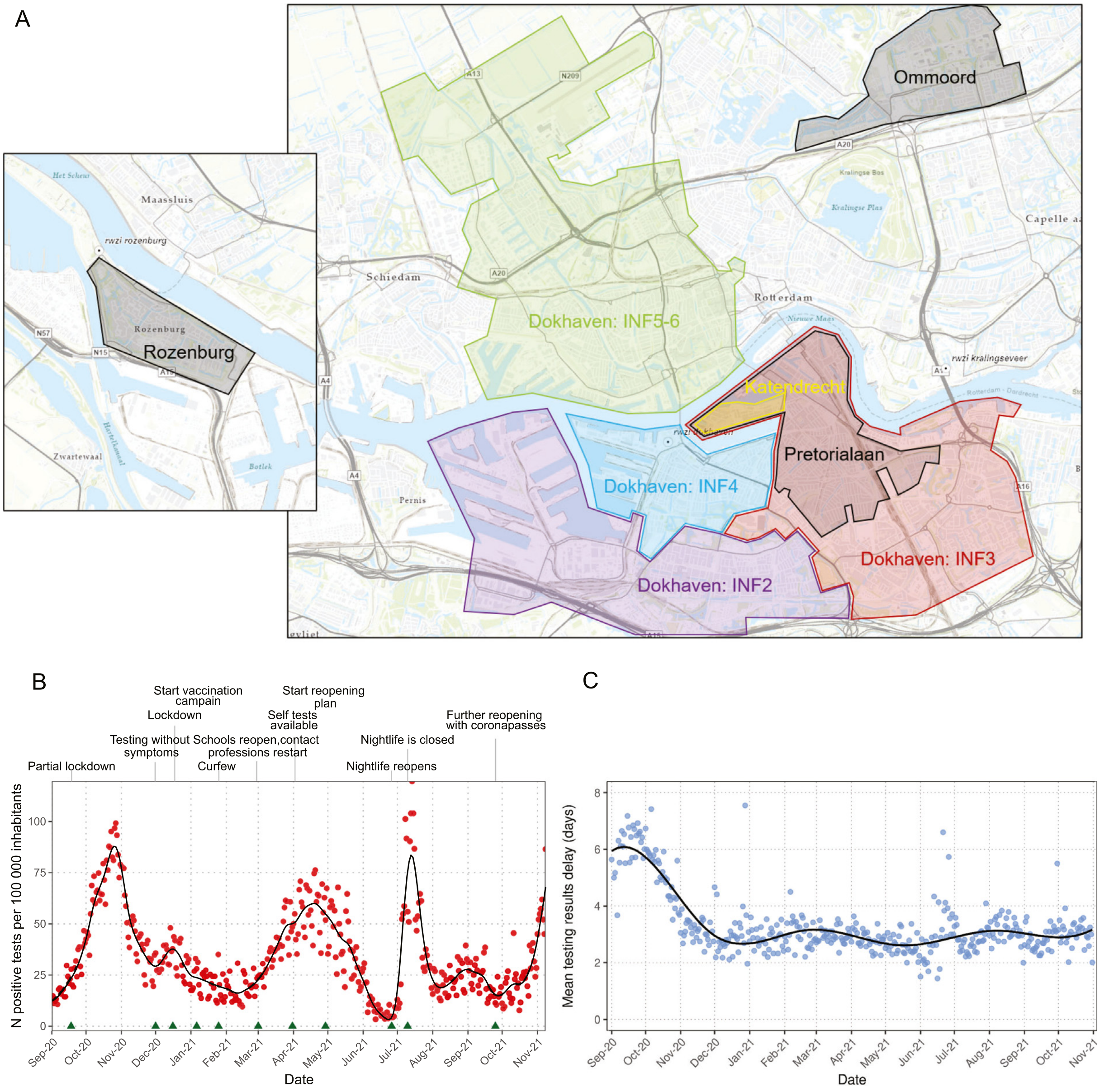


Fig. 1. A) Research areas within Rotterdam Rijnmond. Sample collection occurred at wastewater pumping stations Ommoord (28,434 inhabitants), Katendrecht (6500 inhabitants), Pretorialaan (71,325 inhabitants) and Bergschenhoek (18,750 inhabitants) and wastewater treatment plants (WWTP) Rozenburg (12,374 inhabitants) and Dokhaven. At WWTP Dokhaven samples were collected from 4 sewer pipes that service the catchment areas INF2 Everlo-Waalhaven (27,044 inhabitants), INF3 Pretorialaan-Zuidplein (121,118 inhabitants), INF4 Wolphaertsbocht (36,011 inhabitants) and INF5/6 Heemraadsplein (138,280 inhabitants). B) SARS-COV-2 positive tests per 100,000 inhabitants for all selected areas within Rotterdam-Rijnmond (excluding Bergschenhoek) and a selection of the implemented lock-down measures (green triangles) during the period from September 2020 till November 2021. C) The time that it took for inhabitants serviced by Dokhaven, from onset of symptoms to the result of the SARS-CoV-2 in the testing lanes. The smoothed curves were produced using cubic regression splines.

The lag time, or time difference between trends in the number of SARS-CoV-2 notifications per catchment area and SARS-CoV-2 levels in sewage was explored via the application of dynamic time warping (Giorgino, 2009), where the minimum distance between the two time series was computed using the Rabiner-Juang step function (Rabiner and Juang, 1993). The PHS data also showed that the time between the onset of disease and the notification date varied over the course of the pandemic. These delay data were used to transpose the notification date to the disease onset data for modeling. Since the time of onset of disease was not known for all cases, the date of first illness was assigned to each test result using the average number of days between first day of illness and test result on each date.

2.5. Data preparation and statistical modeling

To quantify how the dynamics of the reported number of clinical cases were corresponding with the normalized virus concentrations in sewage, time-shifted SARS-CoV-2 case data were modeled using a generalized linear mixed model (GLMM). The GLMM included a random effect that can account for in-group correlations since repeated observations nested within one catchment area may be more similar than observations from other catchment areas. The random effect captured the variation of a fixed-effect parameter among catchments with a Gaussian distribution with zero mean and a group variance that was estimated from the data. The selection procedure for the model components followed the steps in Supplementary Fig. S1 (Zuur et al., 2009).

2.5.1. Linear predictor

The linear predictor η is a linear function of the explanatory variables given by

$$\eta_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij}$$

where i represents day number, j the catchment, β refers to the weight assigned to the explanatory variables, and β_0 is the intercept. The first explanatory variable x_1 is the normalized SARS-CoV-2 concentrations in wastewater. The second explanatory variable, the reported total number of tests per 100,000 x_2, represents the testing intensity and served as a measure of the willingness to test. All continuous variables were standardized to avoid numerical estimation problems. The final structure of the linear predictor was obtained by applying backward selection based on the 95 % highest probability density interval. Initial visual data exploration revealed no outliers in the variables and a maximum variance inflation factor of 1.06, which is well below the threshold of 5–10 for correlation between independent variables (Montgomery and Peck, 1992).

2.5.2. Distribution of the response variable

The reported number of positive tests per 100,000 on day i in catchment j, time-shifted towards the first day of illness served as response variable y_{ij}. Several candidate distributions were considered, such as the Gaussian distribution, Poisson distribution and the negative-binomial distribution. The Gaussian distribution was specified by:

$$\mathbf{y} \sim N(\boldsymbol{\eta}, \mathbf{ZGZ'} + \sigma^2\mathbf{I})$$

$$E(\mathbf{y}) = \boldsymbol{\eta} \text{ and } var(\mathbf{y}) = \mathbf{ZGZ'} + \sigma^2\mathbf{I}$$

where design matrix $\mathbf{Z}$ provides the values of random effects for each observation and $\mathbf{G}$ provides a diagonal matrix with variance component σ_j^2.

In addition to the random noise term σ^2 to account for observation noise, other structures were considered to meet the underlying assumptions of linear regression models. One of the main underlying assumptions was a constant variance, or homogeneity. Another assumption was independence, while successive observations from the same catchment were expected to be similar compared to observations from other catchments. Therefore, two extensions of the random part were considered. First was a random effect that imposes a correlation structure on all observations from the same catchment. Second was an auto-regressive process of order 1 (AR1) to account for any temporal dependency present in the time series and is given by:

$$\varepsilon_{ij} = \phi \varepsilon_{i-1j} + \omega_{ij}$$

$$\omega_{ij} = N(0, \sigma^2)$$

where ϕ is the strength of the time-dependent correlation and ω_{ij} is the noise term. The different model structures were compared using the Watanabe–Akaike information criterion (WAIC) and conditional predictive ordinate (CPO) as metrics. WAIC is a generalized version of the AIC and limits the risk of over or underfitting by incorporating the goodness of fit while adding a penalty for model complexity (Watanabe, 2013). CPO is a cross-validatory metric given by

$$-\sum_{k=1}^{n} \log(f(y_k | y_{k-1}))$$

Residual plots were used to validate the model and investigate the assumptions for homogeneity of variance, independence, and normality.

2.5.3. Bayesian modeling

This study adopted a Bayesian approach for estimating the model variable weights β, hyperparameters for the variance parameters and ϕ from the autoregressive correlation term in the model structure. Within the Bayesian framework unknown parameters are considered as random variables for which a posterior distribution is estimated, based on observed data and a prior distribution. Since no prior information on the parameter values was available uninformative priors were specified. An advantage of the Bayesian approach over the maximum likelihood estimates is that it is more robust with respect to the bias of the derived intervals in the case of a random effect with limited number of levels (Stegmueller, 2013). We sampled from posterior marginal distributions using Integrated Nested Laplace Approximations (INLA) as implemented in the R-INLA package in R Version 4.2.1 (Dessau and Pipper, 2008; Martins et al., 2013).

3. Results

3.1. Trends in SARS-CoV-2 notifications per catchment area

Positive cases per 100,000 inhabitants over time were plotted for Rotterdam Rijnmond (Fig. 1B) and for each area (Fig. 2). Peaks are observed for October 2020, December 2020, April 2021, July 2021, and a rise in infections can be seen for October 2021. In September 2020, the time from onset of symptoms to reporting a positive test took an average of 4–6 days and this gradually reduced to an average of 3 days after ramping up test capacity (Fig. 1C).

Overall, all nine catchment areas show similar trends in incidence of reported SARS-CoV-2 cases over time (Fig. 2). For the (smaller) catchment areas of Katendrecht, Ommoord, and Rozenburg, the signal was noisier compared to the other catchment areas. This was most likely related to the size of the catchment population, as the noise was reduced between the cascading datasets of Katendrecht, Pretorialaan and INF3 Pretorialaan-Zuidplein. While overall trends were similar, the communities around the perimeter of the city, “Rozenburg” and “Ommoord”, had lower average levels of reported positive tests. Additionally, the rise in infections after the reopening of the Dutch nightlife in July 2021, leading to a massive spike in cases among young adults, was not observed in Rozenburg, which has an older population, while this increase was largest for the inner-city areas, such as “Katendrecht” and “center of Rotterdam” (INF5–6). The total number of tests varied over time (Supplementary Fig. S2), related to changes in testing policy, behavior, and availability.

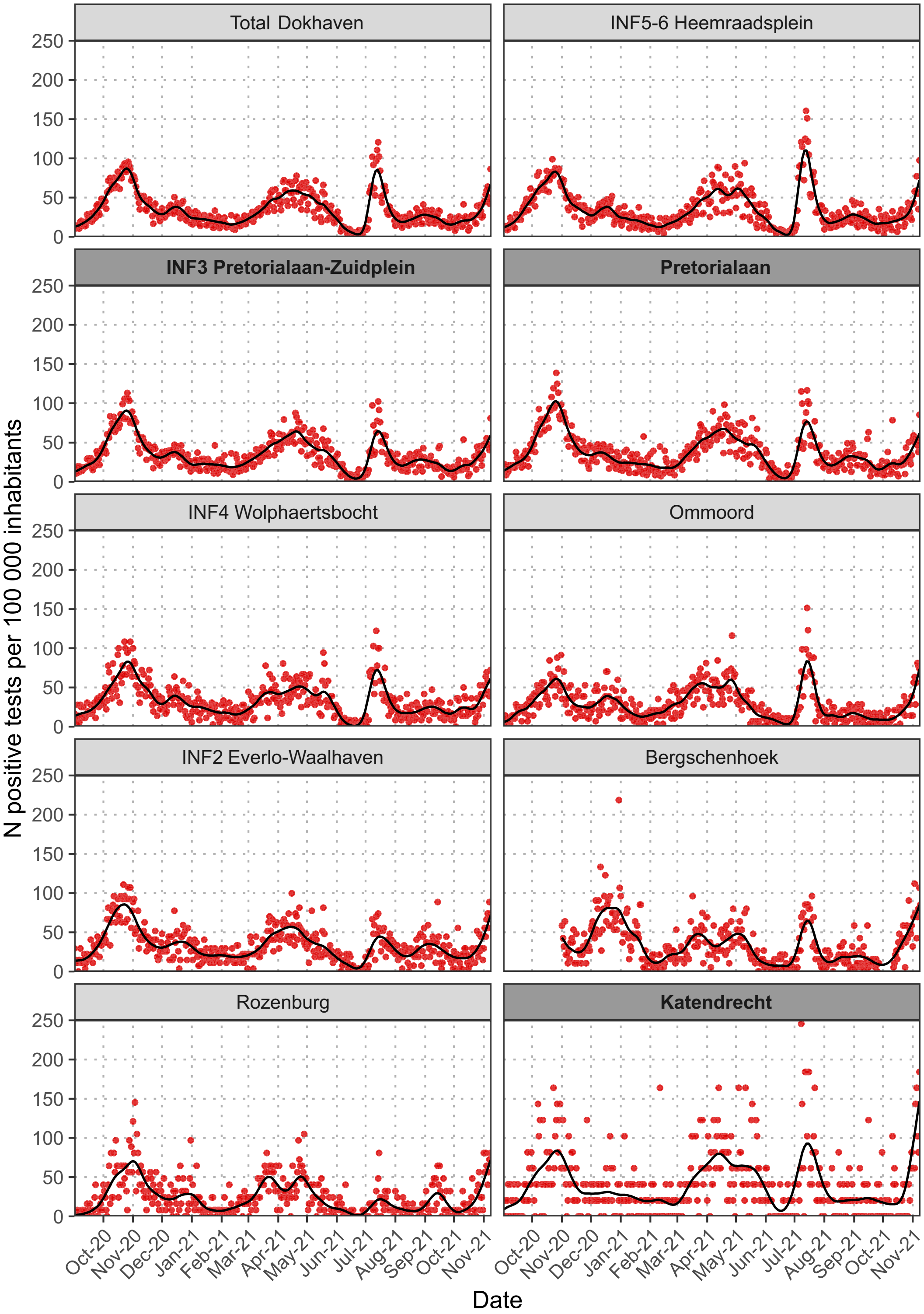


Fig. 2. SARS-COV-2 positive tests per 100,000 inhabitants for all ten areas for the period from September 2020 till November 2021.The areas are ordered based on population size. The areas from Katendrecht to Pretorialaan and INF3 (bold) overlap but are increasing in size. The smoothed curves were produced using cubic regression splines.

3.2. Validation of sewage surveillance by comparison with case notification data

The measured concentrations of SARS-CoV-2 in sewage were normalized for the flow of non-domestic wastewater (Langeveld et al., 2023). SARS-CoV-2 levels in wastewater were highest in October 2020 for all areas except for Rozenburg where it was highest in April 2021. SARS-CoV-2 levels were relatively low in Rozenburg and Bergschenhoek.

The initial plotting and calculation of the time-dependent correlation between the reported SARS-CoV-2 cases and viral loads in wastewater showed that, in the period from September to December 2020, the highest correlation was observed for a time lag of 6 days between an increase or decrease in the level of SARS-CoV-2 in sewage and reported SARS-CoV-2 cases (Supplementary Fig. S3). However, with increasing testing capacity, this time delay became shorter. After correcting the reported case data for the time of symptom onset, the incidence and concentrations of SARS-CoV-2 in sewage were highly collinear, especially during the rise of the October 2020 wave (Fig. 3). This high collinearity, observed when correcting for day of illness onset, implies that high viral shedding in the sewer around the onset day of new cases largely determines the SARS-CoV-2 concentration in wastewater, even though fecal shedding can last for more than two

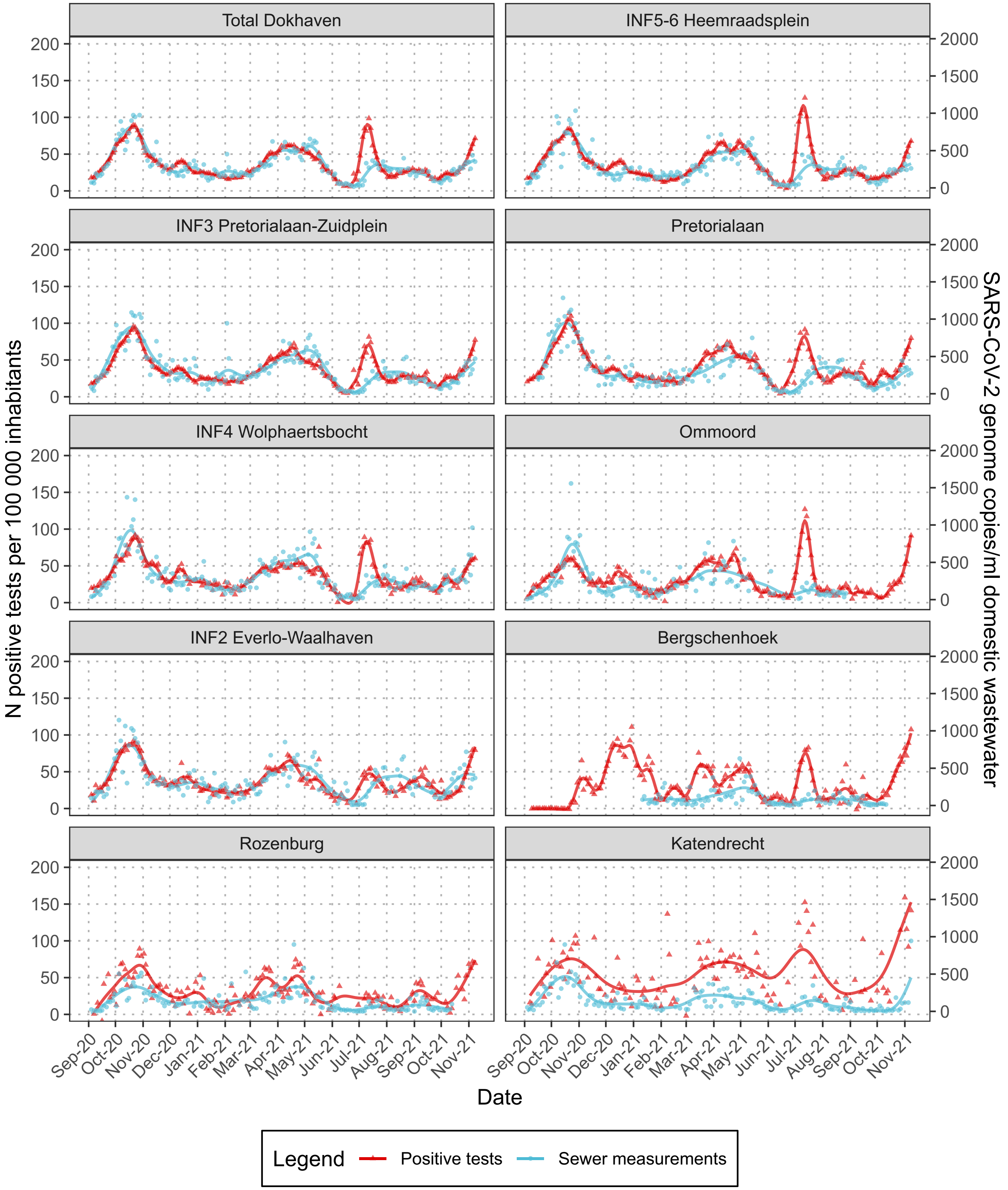


Fig. 3. Levels of normalized SARS-CoV-2 genome copies per ml of domestic wastewater in Rotterdam-Rijnmond (blue) and the number of SARS-CoV-2 positive tests per 100,000 corrected for first day of illness (red). The total inflow at WWTP Dokhaven is calculated as the flow averaged mean of the four larger catchment areas. The areas are ordered based on population size. Cubic regression splines were used to produce the smoothed curves.

weeks (Cevik et al., 2021; Chen et al., 2020; Holm-Jacobsen et al., 2021; Zhang et al., 2021). There were two discrepancies: a peak in incidence when the nightlife reopened (July 2021), and a peak in December 2020, possibly reflecting increased testing ahead of seasonal festivities (Fig. 3, Supplementary Fig. S4).

To investigate these differences, we applied a Bayesian approach to model the expected incidence of reported positive cases by first day of illness based on the normalized SARS-CoV-2 concentrations in wastewater (Fig. 4).

3.3. Model selection

Several nested model structures were defined and compared. Each model included a full set of explanatory variables and a random effect by design. Table 1 shows the WAIC and CPO values for the three models. The Gaussian GLMM with AR1 component and different variances per catchment outperformed the other models. A model validation revealed that the models without an AR1 correlation structure suffered from a significant time-dependent correlation, thus violating the assumption of independence. This is confirmed by the sample variogram shown in Supplementary Fig. S5 that revealed a time-dependent correlation of up to 20 days. The derived standard deviations per catchment depicted in Fig. 5 show a reduction in residual noise associated with an increasing catchment population. In agreement with statistical theory, noise reduction was roughly proportional to the square root of the population sampled. Differences between consecutive observations of the sewage surveillance data did not increase with a decreasing populations size. This suggests the model fit was predominantly

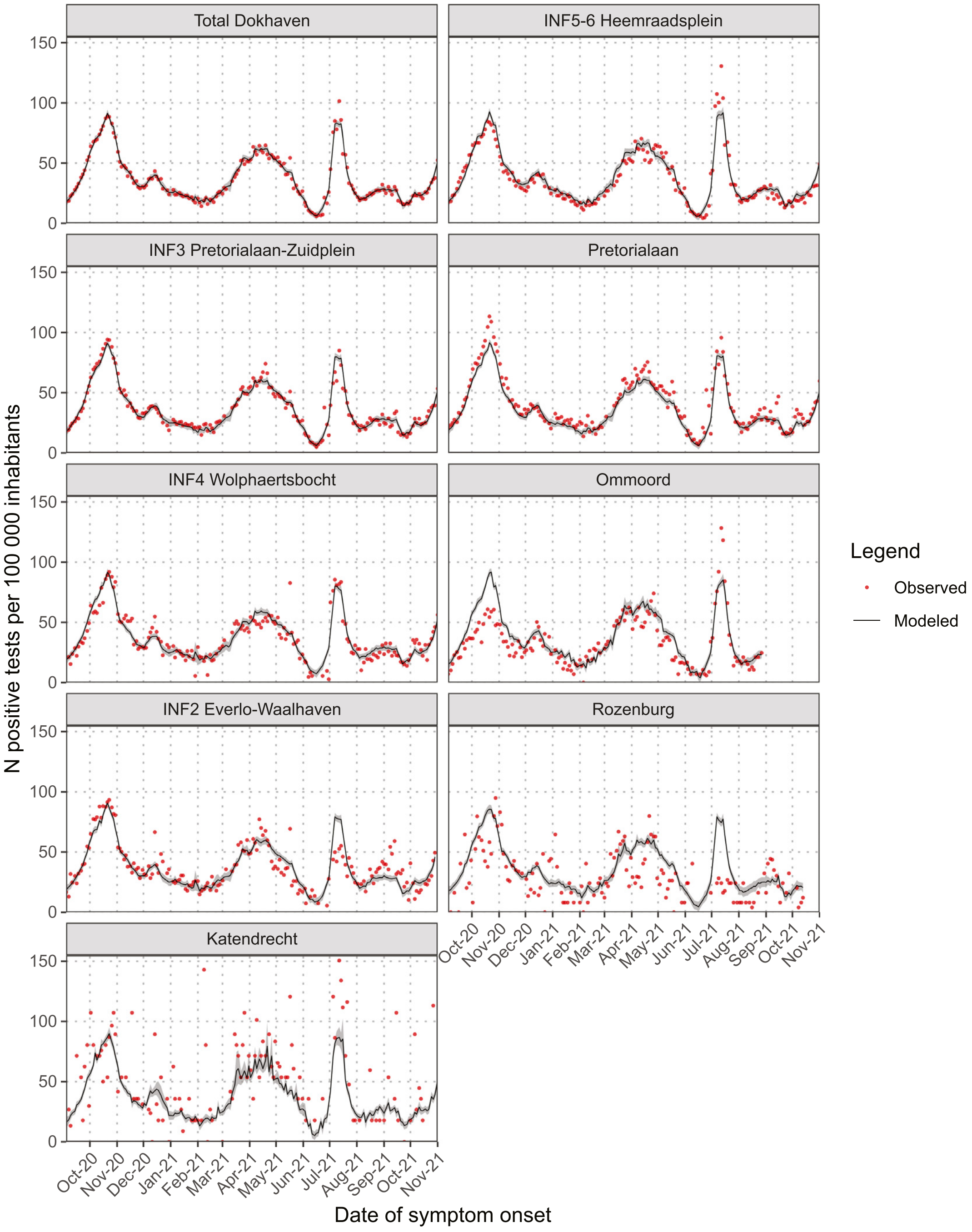


Fig. 4. Modeling of the expected incidence of reported cases based on wastewater. Reported SARS-CoV-2 positive tests over time are shown in red, along with the predicted incidence of reported SARS-CoV-2 positive tests based on the normalized SARS-CoV-2 concentrations in sewage for each of the areas.

Table 1
Model comparison with the Watanabe-Akaike information criterion (WAIC) and the Conditional predictive ordinate (CPO) as metrics.

Model	WAIC	CPO
Gaussian GLMM	3416.6	1.19
Gaussian GLMM with variance structure	2962.4	1.00
Gaussian GLMM with AR1	11,216.9	0.96
Gaussian GLMM with AR1 and variance structure	1386.5	0.47

Table 2
Summary statistics for the selected model.

Parameter	Mean	Median	SD	2.5 %	97.5 %
β_1	0.056	0.056	0.016	0.024	0.088
β_2	0.117	0.116	0.029	0.060	0.176
ϕ	0.784	0.789	0.065	0.645	0.896

limited by noise in the incidence data and not in the sewage surveillance data.

The structure of the final model is:

$$y_{ij} = \beta_1 x_{1ij} + (\beta_2 + u_j) x_{2ij} + \varepsilon_{ij}$$

$$u_j \sim N\left(0, \sigma_j^2\right)$$

$$\varepsilon_{ij} = \phi \varepsilon_{i-1\,j} + \omega_{ij}$$

$$\omega_{ij} = N\left(0, \sigma_j^2\right)$$

where the y_{ij} is the reported number of positive tests per 100,000 on day i in catchment j, time-shifted towards the first day of illness. The explanatory variable x_1 is the normalized SARS-CoV-2 concentration in wastewater, and x_2 is the reported total number of tests per 100,000, representing testing intensity and serving as a measure of the willingness to test. Since the effect of the latter variable may differ per catchment j, the random slope u_j was added to capture the variability. A sensitivity analysis revealed that the addition of a random slope improved model performance as the WAIC decreased by 26.2. Weights β for the explanatory variables x_1 and x_2 are shown in Table 2 and the marginal distributions for each weight are depicted in Supplementary Fig. S6A. The positive median values in Table 2 mean that both high SARS-CoV-2 levels in wastewater and a large number of total tests are associated with a large number of positive tests. The intercept was not significantly different from 0 at a 5 % significance level and was subsequently omitted from the model. The parameter ϕ for the correlation was strong, indicating a high sampling frequency throughout the time course of the project.

A model validation revealed no problems with homogeneity, independence, or normality. A plot of the model residuals is provided in Supplementary Fig. S7 and showed that the choice of the Gaussian distribution was appropriate given the absence of negative fitted values and signs of none-normality in the scaled residuals.

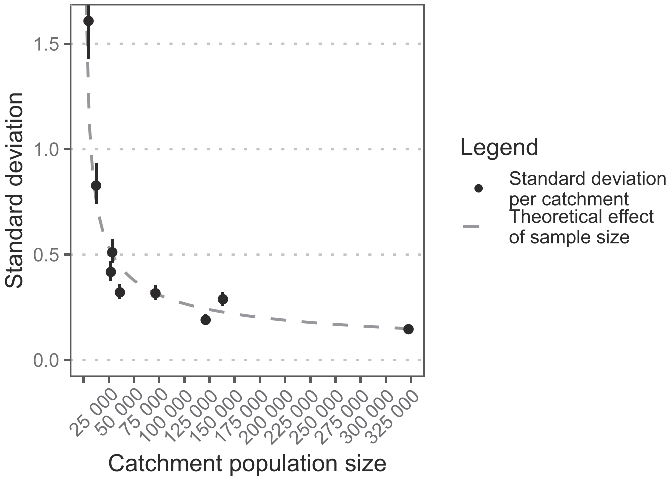


Fig. 5. Model estimated standard deviations per catchment area. The dashed line represents the theoretical relation of the residual spread in a catchment being proportional to the square root of the population size.

Using this model, the December 2020 peak of reported cases that was not evident from sewage surveillance (Fig. 3), could now be modeled based on a small increase in sewage and a large increase in clinical tests. However, at the height of the outbreak in July 2021, the values predicted via sewage surveillance were still significantly lower than the reported incidence of SARS-CoV-2 positive individuals, after this outbreak the fit between wastewater and positive cases was restored.

During the sampling period, the initial circulating SARS-CoV-2 strains were replaced by alpha and later by the delta VOC. As differences in viral characteristics can lead to changes in tropism or shedding rates, these replacements could affect the ratio of SARS-CoV-2 levels in wastewater versus the incidence of cases. This was investigated by dividing the data in three different time periods based on VOC circulation; September 1, 2020 till December 31, 2020 (non-VOC), February 26, 2021 till June 13, 2021 (alpha VOC) and July 23, 2021 till November 9, 2021 (delta VOC) (coronadashboardrijksoverheid, 2022). Next, the final model was extended with interaction terms between SARS-CoV-2 levels in wastewater and the VOC time periods. This showed that the ratio of SARS-CoV-2 levels in wastewater versus the incidence of cases was not affected by VOC since the parameters did not significantly differ between the different time periods (Table 3 and Supplementary Fig. S6B).

3.4. Sewage surveillance to address public health questions, and to monitor the effect of interventions

In December 2020 there was an outbreak of the alpha VOC in the village Bergschenhoek within municipality of Lansingerland (van Beek et al., 2022). During this outbreak a large-scale testing effort was carried out for Lansingerland and a sampling cabinet was placed in Bergschenhoek for parallel wastewater surveillance. In Lansingerland between the 11th and 22nd of January 36,534 of 63,338 (58 %) residents were tested. For Bergschenhoek, 118 positive SARS-CoV-2 tests were reported (629 per 100,000 inhabitants). The average normalized concentration of SARS-CoV-2 in Bergschenhoek wastewater in this period was 66 genome copies (GC)/ml of domestic wastewater, indicating that 1/100.000 positive tests would corresponded to 66/629 = 0.10 GC/ml of domestic wastewater. The larger catchment areas in Rotterdam (INF3 and INF5/6) where only symptomatic individuals were tested (1.9–2.7 % of the residents) in that same time period, reported lower incidences of 269 and 293 positives per 100,000 inhabitants. The corresponding average concentrations in sewage in these catchment areas were higher: 151–164 GC/ml of domestic wastewater. This would compute to 0.53–0.56 GC/ml per 1/100.000 positive tests, yielding a fivefold difference between large-scale population testing (Bergschenhoek) and symptomatic case testing (INF3 and 5/6). With comparable shedding rates between symptomatic and asymptomatic cases (Lee et al., 2020). This suggested a five-times-higher incidence than that reported in INF3 and 5/6.

During the large-scale testing effort levels SARS-CoV-2 in wastewater declined for Bergschenhoek but increased for INF3 and INF5/6 indicating that this intervention potentially had an effect on SARS-CoV-2 circulation.

4. Discussion

Sewage surveillance is now widely used in parallel to testing for suspected SARS-CoV-2 cases, and several groups showed a correlation

Table 3
Summary statistics for the variant of concern (VOC).

VOC	Mean	Median	SD	2.5 %	97.5 %
None	0.079	0.079	0.045	−0.009	0.167
Alpha	−0.070	−0.070	0.049	−0.166	0.026
Delta	0.005	0.005	0.058	−0.109	0.119

between the number of reported cases and SARS-CoV-2 levels in wastewater (Barrios et al., 2021; Fernandez-Cassi et al., 2021; Medema et al., 2020b; Peccia et al., 2020; Wu et al., 2021). Our high-resolution data in well-matched populations along the surveillance pyramid clearly show that wastewater surveillance was ahead of clinical testing at the start of the second wave in Rotterdam in October 2020, and that this was due to the delay between disease onset with a high virus shedding, and the time a case was reported. Over time, the ramping up of clinical testing has reduced this delay, and thus the lead-time of sewage testing for early warning in this setting. Similar results were found for a 14-month-long wastewater surveillance time series in Massachusetts where wastewater data served as an early warning system for the first wave but not during the second, also most likely due to the ramping up of testing availability (Aberi et al., 2021; Xiao et al., 2021).

Our model allowed us to investigate which variables in clinical testing affect the relation between clinical and sewage surveillance, such as testing intensity and testing delay. When including these variables, a peak of reported cases that was not evident from sewage surveillance, could be modeled based on a small increase in sewage combined with a large increase in clinical testing. While trends based on sewage testing and the modeled incidence of reported cases were highly colinear, there was a clear delineation in July 2021, after the re-opening of the Dutch nightlife and events. Starting from June 18, 2021 events were accessible after a negative test and many more (young) people were tested during this period via an antigen test. We hypothesized that a high percentage of individuals were tested in the testing-street after a positive antigen test, while the negative antigen tests do not show up in the testing street results. This was supported by the strong increase in the percentage of positive tests at that time. It also showed that with standard testing a sudden increase in a niche group can quickly be discovered, while this is not possible with sewage surveillance.

We selected multiple neighborhoods and areas within Rotterdam to investigate if SARS-CoV-2 dynamics can vary within a city and if these differences would be evident through wastewater surveillance. The comparison of smaller areas enclosed in larger areas showed similar dynamics. However, in smaller areas relatively few infected individuals can have a large impact on SARS-CoV-2 trends and a large spread was observed between reported and predicted SARS-CoV-2 cases for the smallest catchment areas. Our model showed that this was mostly caused by noise in the above ground data, rather than noise in the sewage data, meaning that sewage testing even for small catchment areas (serving 6500 individuals) can be used to identify trends in SARS-CoV-2 circulation. Trends of reported SARS-CoV-2 positives and sewage were similar for each area except Rozenburg and Bergschenhoek. This was supported by our model which showed that wastewater surveillance can be used to compare neighborhoods.

From our results we deduced that peak shedding of SARS-CoV-2 in feces occurs around the time of symptom onset. This correlates with the viral load in respiratory samples, as the highest levels of virus shedding occurred around onset of symptoms (Cevik et al., 2021; He et al., 2020; Lewis et al., 2021). Available human fecal shedding data on SARS-CoV-2 does not show a high initial peak shedding, but most studies do not include data around symptom onset (Medema et al., 2020a). However, animal studies showed that shedding kinetics in the feces were similar or slightly delayed compared to the respiratory tract (Gaudreault et al., 2020; Richard et al., 2020; Sia et al., 2020).

This study showed the sensitivity of sewage surveillance in relation to the dynamics in case-based surveillance: when the wastewater concentrations were just above the limit of detection in the larger catchments in June 2021, the reported incidence was around 2 per 100.000. During the large-scale surveillance of Bergschenhoek, the ratio between reported incidence and normalized wastewater concentration was around fivefold lower than in the large catchment areas in Rotterdam where only a small percentage of the population was tested, which is indicative of significant undertesting. Not surprisingly, differences in testing behavior affect the relationship between sewage and clinical surveillance, highlighting that the value of sewage surveillance is independent from testing behavior. Another potential variable is the emergence of SARS-CoV-2 variants since variant-specific phenotypical changes could result in differences in tropism, fecal shedding, age range, symptoms, or vaccine breakthrough infections which could affect the correlation between the incidence of reported cases and wastewater surveillance. For example, 1000-fold higher viral loads were reported for patients infected with the delta VOC compared to the 19A/19B strain (Li et al., 2021), and 10-fold higher viral loads for the alpha VOC compared to previous non-VOC strains. Whether these VOC strains also result in higher fecal shedding is unknown. Nevertheless, the model parameter for wastewater concentrations during the period when non-VOC SARS-CoV-2 was predominant were not significantly different from the periods when VOC alpha and beta were predominant, indicating that for wastewater concentrations, there was no measurable effect of the emergence of these VOC. During this study SARS-CoV-2 vaccination was rolled out in the Netherlands, but the effect on SARS-CoV-2 levels in wastewater appears to be limited. SARS-CoV-2 breakthrough infections after vaccination could result in lower levels of shedding, but only a 2.8–4.5-fold decrease in respiratory viral load in vaccinated compared to non-vaccinated individuals was reported (Levine-Tiefenbrun et al., 2021).

5. Conclusions

This study shows that wastewater surveillance can, independently of testing behavior, accurately display SARS-CoV-2 dynamics within a city and is sensitive enough to measure small variations in the number of infected individuals within or between neighborhoods. Sewage surveillance alongside the large-scale testing effort at Lansingerland indicated that many SARS-CoV-2 cases go unreported and that sewage surveillance can be used to monitor the effect of interventions. This is especially relevant at times and in locations where the willingness to test is lower, testing facilities have scaled down or when home antigen testing is applied as an alternative to clinical tests. As clinical test data provides insights into patient characteristics, sewage data supplement this with insights into the overall infection pressure in a specific catchment area. Sewage surveillance combined with NGS or ddRT-PCR can also provide rapid insight in the spread of new SARS-CoV-2 variants without the need to analyze thousands of patient samples. We anticipate that wastewater testing should be considered as one of the pillars of future surveillance for (re) emerging viruses.

Funding

This study was financed by STOWA, TKI Watertechnology, Erasmus Foundation, Adessium Foundation, European Union's Horizon H2020 grant VEO (grant no. 874735), 2018 NWO Stevin Award Koopmans, Ministry of Health, Welfare and Sport, H2020 and Waterboards Waterschap Hollandse Delta, Hoogheemraadschap van Delfland and Hoogheemraadschap Schieland en Krimpenerwaard.

CRediT authorship contribution statement

Miranda de Graaf: Conceptualization, Formal analysis, Funding acquisition, Writing – original draft. **Jeroen Langeveld:** Conceptualization, Formal analysis, Writing – original draft. **Johan Post:** Conceptualization, Formal analysis, Methodology, Visualization, Writing – original draft. **Christian Carrizosa:** Data curation, Formal analysis, Methodology, Visualization. **Eelco Franz:** Conceptualization, Writing – review & editing. **Ray W. Izquierdo-Lara:** Investigation. **Goffe Elsinga:** Investigation. **Leo Heijnen:** Investigation. **Frederic Been:** Investigation. **Janko van Beek:**

Data curation. **Remy Schilperoort:** Investigation. **Rianne Vriend:** Data curation. **Ewout Fanoy:** Resources, Conceptualization, Funding acquisition, Writing – review & editing. **Evelien I.T. de Schepper:** Conceptualization, Writing – review & editing. **Marion P.G. Koopmans:** Conceptualization, Funding acquisition, Writing – review & editing. **Gertjan Medema:** Conceptualization, Funding acquisition, Writing – review & editing.

Data availability

Data will be made available on request.

Declaration of competing interest

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

Acknowledgements

The authors are very grateful for the assistance of Roan Pijnacker, the Water Authorities and WWTP and pumping station operators of Waterschap Hollandse Delta and municipality of Rotterdam, staff of IMD for installing and maintaining the autosamplers and sampling specialists of AQUON.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2023.163599.

References

Aberi, P., Arabzadeh, R., Insam, H., Markt, R., Mayr, M., Kreuzinger, N., et al., 2021. Quest for optimal regression models in SARS-CoV-2 wastewater based epidemiology. Int. J. Environ. Res. Public Health 18.

Barrios, M.E., Diaz, S.M., Torres, C., Costamagna, D.M., Blanco Fernandez, 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.

The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team, 2020. The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) - China, 2020. China CDC Wkly 2, 113–122.

van Beek, J., Teesing, G., Oude Munnink, B.B., Meima, A., Vriend, H.J., Elzakkers, J., et al., 2022. Population-based screening in a municipality after a primary school outbreak of the SARS-CoV-2 alpha variant, the Netherlands, December 2020-February 2021. PLoS One 17, e0276696.

Cevik, M., Tate, M., Lloyd, O., Maraolo, A.E., Schafers, J., Ho, A., 2021. SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. Lancet Microbe 2, e13–e22.

Chen, Y., Chen, L., Deng, Q., Zhang, G., Wu, K., Ni, L., et al., 2020. The presence of SARS-CoV-2 RNA in the feces of COVID-19 patients. J. Med. Virol. 92, 833–840.

Claro, I.C.M., Cabral, A.D., Augusto, M.R., Duran, A.F.A., Graciosa, M.C.P., Fonseca, F.L.A., et al., 2021. Long-term monitoring of SARS-COV-2 RNA in wastewater in Brazil: a more responsive and economical approach. Water Res. 203, 117534.

Corman, V.M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D.K., et al., 2020. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill. 25.

coronadashboardrijksoverheid, 2022. www.coronadashboardrijksoverheid.nl.

Dessau, R.B., Pipper, C.B., 2008. "R"--project for statistical computing. Ugeskr Laeger 170, 328–330.

Fernandez-Cassi, X., Scheidegger, A., Banziger, C., Cariti, F., Tunas Corzon, A., Ganesanandamoorthy, P., et al., 2021. Wastewater monitoring outperforms case numbers as a tool to track COVID-19 incidence dynamics when test positivity rates are high. Water Res. 200, 117252.

Gaudreault, N.N., Trujillo, J.D., Carossino, M., Meekins, D.A., Morozov, I., Madden, D.W., et al., 2020. SARS-CoV-2 infection, disease and transmission in domestic cats. Emerg. Microbes Infect 9, 2322–2332.

Giorgino, T., 2009. Computing and visualizing dynamic time warping alignments in R: the dtw package. J. Stat. Softw. 31, 1–24.

He, X., Lau, E.H.Y., Wu, P., Deng, X., Wang, J., Hao, X., et al., 2020. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat. Med. 26, 672–675.

Heijnen, L., Elsinga, G., de Graaf, M., Molenkamp, R., Koopmans, M.P.G., Medema, G., 2021. Droplet digital RT-PCR to detect SARS-CoV-2 signature mutations of variants of concern in wastewater. Sci. Total Environ. 799, 149456.

Hellmer, M., Paxeus, N., Magnius, L., Enache, L., Arnholm, B., Johansson, A., et al., 2014. Detection of pathogenic viruses in sewage provided early warnings of hepatitis Avirus and norovirus outbreaks. Appl. Environ. Microbiol. 80, 6771–6781.

Hendriksen, R.S., Munk, P., Njage, P., van Bunnik, B., McNally, L., Lukjancenko, O., et al., 2019. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. Nat. Commun. 10, 1124.

Holm-Jacobsen, J.N., Bundgaard-Nielsen, C., Rold, L.S., Jensen, A.M., Shakar, S., Ludwig, M., et al., 2021. The prevalence and clinical implications of rectal SARS-CoV-2 shedding in Danish COVID-19 patients and the general population. Front. Med. 8, 804804 Lausanne.

Izquierdo-Lara, R., Elsinga, G., Heijnen, L., Munnink, B.B.O., Schapendonk, C.M.E., Nieuwenhuijse, D., et al., 2021. Monitoring SARS-CoV-2 circulation and diversity through community wastewater sequencing, the Netherlands and Belgium. Emerg. Infect. Dis. 27, 1405–1415.

Langeveld, J., Schilperoort, R., Heijnen, L., Elsinga, G., Schapendonk, C.E.M., Fanoy, E., et al., 2023. Normalisation of SARS-CoV-2 concentrations in wastewater: the use of flow, electrical conductivity and crAssphage. Sci. Total Environ. 865, 161196.

Lee, S., Kim, T., Lee, E., Lee, C., Kim, H., Rhee, H., et al., 2020. Clinical course and molecular viral shedding among asymptomatic and symptomatic patients with SARS-CoV-2 infection in a community treatment center in the Republic of Korea. JAMA Intern. Med. 180, 1447–1452.

Levine-Tiefenbrun, M., Yelin, I., Katz, R., Herzel, E., Golan, Z., Schreiber, L., et al., 2021. Initial report of decreased SARS-CoV-2 viral load after inoculation with the BNT162b2 vaccine. Nat. Med. 27, 790–792.

Lewis, N.M., Duca, L.M., Marcenac, P., Dietrich, E.A., Gregory, C.J., Fields, V.L., et al., 2021. Characteristics and timing of initial virus shedding in severe acute respiratory syndrome coronavirus 2, Utah, USA. Emerg. Infect. Dis. 27, 352–359.

Li, B., Deng, A., Li, K., Hu, Y., Li, Z., Xiong, Q., et al., 2021. Viral infection and transmission in a large, well-traced outbreak caused by the SARS-CoV-2 Delta variant. medRxiv 2021.07.07.21260122.

Martins, T.G., Simpson, D., Lindgren, F., Rue, H., 2013. Bayesian computing with INLA: new features. Comput.Stat.Data Anal. 67, 68–83.

Medema, G., Been, F., Heijnen, L., Petterson, S., 2020a. 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.

Medema, G., Heijnen, L., Elsinga, G., Italiaander, R., Brouwer, A., 2020b. 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, 511–516.

Mercer, T.R., Salit, M., 2021. Testing at scale during the COVID-19 pandemic. Nat. Rev. Genet. 22, 415–426.

Montgomery, D.C., Peck, E.A., 1992. Introduction to Linear Regression Analysis. Wiley.

Munster, V.J., Koopmans, M., van Doremalen, N., van Riel, D., de Wit, E., 2020. A novel coronavirus emerging in China - key questions for impact assessment. N. Engl. J. Med. 382, 692–694.

Nieuwenhuijse, D.F., Oude Munnink, B.B., Phan, M.V.T., Munk, P., Venkatakrishnan, S., et al., Global Sewage Surveillance project c, 2020. Setting a baseline for global urban virome surveillance in sewage. Sci. Rep. 10, 13748.

Peccia, J., Zulli, A., Brackney, D.E., Grubaugh, N.D., Kaplan, E.H., Casanovas-Massana, A., et al., 2020. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. Nat. Biotechnol. 38, 1164–1167.

Rabiner, L.J., Juang, B.H., 1993. Fundamentals of Speech Recognition. Prentice Hall, Englewood Cliffs, NJ.

Richard, M., Kok, A., de Meulder, D., Bestebroer, T.M., Lamers, M.M., Okba, N.M.A., et al., 2020. SARS-CoV-2 is transmitted via contact and via the air between ferrets. Nat. Commun. 11, 3496.

Sah, P., Fitzpatrick, M.C., Zimmer, C.F., Abdollahi, E., Juden-Kelly, L., Moghadas, S.M., et al., 2021. Asymptomatic SARS-CoV-2 infection: a systematic review and meta-analysis. Proc. Natl. Acad. Sci. U. S. A. 118.

Sia, S.F., Yan, L.M., Chin, A.W.H., Fung, K., Choy, K.T., Wong, A.Y.L., et al., 2020. Pathogenesis and transmission of SARS-CoV-2 in golden hamsters. Nature 583, 834–838.

Stachler, E., Kelty, C., Sivaganesan, M., Li, X., Bibby, K., Shanks, O.C., 2017. Quantitative CrAssphage PCR assays for human fecal pollution measurement. Environ. Sci. Technol. 51, 9146–9154.

Stegmueller, D., 2013. How many countries for multilevel modeling? A comparison of frequentist and Bayesian approaches. Am. J. Polit. Sci. 57, 748–761.

Watanabe, S., 2013. A widely applicable Bayesian information criterion. J. Mach. Learn. Res. 14, 867–897.

Wu, F., Xiao, A., Zhang, J., Moniz, K., Endo, N., Armas, F., et al., 2021. SARS-CoV-2 RNA concentrations in wastewater foreshadow dynamics and clinical presentation of new COVID-19 cases. Sci. Total Environ. 805, 150121.

Xiao, A., Wu, F., Bushman, M., Zhang, J., Imakaev, M., Chai, P.R., et al., 2021. Metrics to relate COVID-19 wastewater data to clinical testing dynamics. medRxiv.

Zhang, Y., Cen, M., Hu, M., Du, L., Hu, W., Kim, J.J., et al., 2021. Prevalence and persistent shedding of fecal SARS-CoV-2 RNA in patients with COVID-19 infection: a systematic review and meta-analysis. Clin. Transl. Gastroenterol. 12, e00343.

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., et al., 2020. A novel coronavirus from patients with pneumonia in China, 2019. N. Engl. J. Med. 382, 727–733.

Zuur, A., Ieno, E.N., Walker, N., Saveliev, A.A., Smith, G.M., SpringerLink, 2009. Mixed Effects Models And Extensions in Ecology With R. Springer New York, New York, NY Imprint: Springer.