communications medicine

A Nature Portfolio journal

Article

https://doi.org/10.1038/s43856-025-00908-5

Post-recovery viral shedding shapes wastewater-based epidemiological inferences

Tin Phan[1], Samantha Brozak[2], Bruce Pell[3], Stanca M. Ciupe[4,5], Ruian Ke[1], Ruy M. Ribeiro[1], Anna Gitter[6], Kristina D. Mena[6], Alan S. Perelson[1,7], Yang Kuang[2] & Fuqing Wu[6]

Abstract

Background The prolonged viral shedding from the gastrointestinal tract is well documented for numerous pathogens, including SARS-CoV-2. However, the impact of prolonged viral shedding on epidemiological inferences using wastewater data is not yet fully understood.

Methods To gain a better understanding of this phenomenon at the population level, we extended a wastewater-based modeling framework that integrates viral shedding dynamics, viral load data in wastewater, case report data, and an epidemic model.

Results Our results indicate that as an outbreak progresses, the viral load from recovered individuals gradually becomes predominant, surpassing that from the infectious population. This phenomenon leads to a dynamic relationship between model-inferred and reported daily incidence over the course of an outbreak. Sensitivity analyses on the duration and rate of viral shedding for recovered individuals reveal that accounting for this phenomenon can considerably advance prediction of transmission peak timing. Furthermore, extensive viral shedding from the recovered population toward the conclusion of an epidemic wave may overshadow viral signals from newly infected cases carrying emerging variants, which can delay the rapid recognition of emerging variants based on viral load.

Conclusions These findings highlight the necessity of integrating post-recovery viral shedding to enhance the accuracy and utility of wastewater-based epidemiological analysis.

Plain language summary

Infected people shed virus to the sewer. Thus, measurements of virus in wastewater can be used to track and predict the progression of an outbreak, as is the case for the COVID-19 pandemic. However, some people continue shedding the virus weeks to months after they recover and can no longer transmit the disease. This complicates the analysis of wastewater data. Using a mathematical model based on viral shedding and wastewater data, we demonstrate that prolonged viral shedding can substantially alter interpretation of results from wastewater data, such as predicting when viral transmission is peaking, and delaying the detection of new viral variants. Our work emphasizes the importance of accounting for post-recovery viral shedding in the analysis of wastewater data to maximize its accuracy and utility.

Wastewater-based epidemiology (WBE) has emerged as an essential public health tool for monitoring the spread of infectious diseases[1–9]. Wastewater collects human excretions including pathogenic elements like viruses from communities. By testing viral signals in wastewater samples, WBE offers a non-invasive, cost-effective approach to track the outbreak progression and estimate the prevalence of infections in the sewershed. Understanding viral shedding, the process by which infected individuals release virus particles into the environment, is crucial for maximizing the accuracy and utility of WBE[10].

Viral shedding is a dynamic process that generally progresses through several stages. Initially, during the incubation phase, the virus replicates in the host without causing symptoms. Low viral shedding is expected during the incubation phase. This is usually followed by an infectious stage characterized by symptom manifestations and high viral shedding, though some individuals may not show any symptoms. Subsequently, the individual enters a recovery phase, where symptoms wane and viral shedding declines[11–14]. Recovered individuals are defined as those no longer capable of transmitting the infection - as determined by upper respiratory viral load[15–19], regardless of their symptoms or viral shedding. This disease progression exhibits great heterogeneity among individuals, with variations in symptom manifestations[20–23] and in viral

[1]Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, NM, USA. [2]School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ, USA. [3]Department of Mathematics and Computer Science, Lawrence Technological University, Southfield, MI, USA. [4]Department of Mathematics, Virginia Tech, Blacksburg, VA, USA. [5]Virginia Tech Center for the Mathematics of Biosystems, Blacksburg, VA, USA. [6]Department of Environmental and Occupational Health Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, USA. [7]Santa Fe Institute, Santa Fe, NM, USA. e-mail: fuqing.wu@uth.tmc.edu

shedding dynamics across the respiratory[24–27] and gastrointestinal (GI) tracts[13,28–30].

It is well documented that some individuals infected with SARS-CoV-2 experience prolonged GI tract viral shedding, which can be detected in stool samples for weeks to months post-recovery[13,28–33]. Yet, the impact of prolonged viral shedding on epidemiological insights derived from wastewater data remains largely unexplored. One reason could be that during the early phase of an outbreak, the shedding from recovered individuals seems minor due to a lower post-recovery shedding and a relatively small number of recovered individuals. However, its significance likely increases as the outbreak unfolds due to prolonged shedding, especially when the number of recovered individuals surpasses the number of new infections. This dynamic relationship between recovery shedding and outbreak progression suggests the need to account for post-recovery viral shedding in the wastewater-based epidemiological framework.

In this study, our objective is to investigate how viral shedding from the recovered population affects epidemiological inferences drawn from wastewater data. We use a dynamic model integrating wastewater data, viral shedding, and transmission dynamics to demultiplex the viral shedding contribution from infectious and recovered populations[10,34,35]. Model fittings, simulations, and analyses reveal an increasing contribution of virus shed by the recovered population as an outbreak unfolds. This phenomenon introduces nonlinearities into the relationship between viral load and case data, thus altering epidemiological inferences of transmission dynamics and emerging variants based on wastewater data. By providing a more nuanced understanding of how viral shedding dynamics influence the relationship between wastewater viral load and case report data, this study enhances the capability of WBE to capture the epidemic landscape.

Methods

Data

Wastewater data, including SARS-CoV-2 concentrations and flow rate, were collected during our previous study[1]. Composite (24-h) wastewater samples were collected from the Greater Boston area from 10/01/2020 to 02/27/2021. The time frame was chosen to limit the data to a single wave of the COVID-19 pandemic (Alpha variant B.1.1.7). We selected the Alpha wave to minimize confounding effects from vaccination introduced in December of 2020, as vaccination may impact viral shedding and transmission dynamics. Additionally, because the viral shedding kinetics is parametrized based on the ancestral strain[13] and extensive wastewater data during the initial outbreak is not available, Alpha—which shares similar viral kinetics with the ancestral strain[24,26]—is an appropriate candidate. The reported case data in the sewershed were obtained from the Massachusetts government website and are publicly accessible at www.mass.gov. The University of Texas Health Science Center at Houston Institutional Review Board determined that this study does not meet the regulatory definition of human subjects research (Reference number: 262788) and so specific IRB approval was not required.

SEIR-V model with temperature variation

SEIR-V type models have been developed and evaluated for their ability to infer and predict transmission dynamics from wastewater data[2,9,34–45]. Our previous study[34] showed that viral shedding during the incubation period can be negligible, due to its short duration and likely low gastrointestinal (GI) tract shedding rate; however, the net viral contribution from the recovered phase can be substantial. Hence, we modify the basic SEIR-V model[34] to include viral shedding from both the infectious and recovered populations, as described in Eq. 5. This model assumes a closed population with uniform interaction among individuals[46]. For the analysis of a single COVID-19 wave, we simplify our SEIR model by omitting natural births and deaths[47–49], in contrast to models that generate long-term projections[50,51].

$$S' = -\lambda IS \tag{1}$$

$$E' = \lambda IS - kE \tag{2}$$

$$I' = kE - \delta I \tag{3}$$

$$R' = \delta I - \sigma R \tag{4}$$

$$V' = \alpha\left(1 - \gamma(T)\right)\left(\beta_I I + \beta_R R\right) \tag{5}$$

S represents the susceptible population, E is the exposed population in the incubation phase, I is the infectious population, and R is the recovered population who continue to shed the virus. Susceptible individuals are infected at a rate λI. The average duration of the incubation phase is $1/k$ days. Infectious individuals recover after an average duration of $1/\delta$ days. Recovered individuals shed virus for an average duration of $1/\sigma$ days. This classification of infection stage is based on previous studies of viral kinetics in the upper respiratory tract and its relationship to infectiousness[15–19]. Both infectious and recovered individuals contribute to the cumulative viral RNA in wastewater (V) at rates β_I and β_R, respectively. α is the average fecal load per day. $\gamma(T)$ is the fraction of virus lost in wastewater between shedding and collection (for measurement). Previous studies have shown that the viral degradation is temperature-dependent (T)[52]. Therefore, we modeled temperature variation throughout the study period by fitting the model to the average daily wastewater temperature data in Celsius[34], which is given in Eq. 6.

$$T(t) = 3.6249\sin(0.0202t - 4.4665) + 16.2298 \tag{6}$$

where t is measured in days and $t = 0$ is 10/01/2020. The temperature-adjusted half-life is modeled using an exponential decay of the form[34,39,53]:

$$\eta(T) = \eta_0 Q_{10}^{-\left(T(t)-T_0\right)/10} \tag{7}$$

with η_0 being the half-life (in hours) at ambient temperature T_0, and Q_{10} ($= 2.5$) being the temperature-dependent rate of change[34,53]. The temperature-dependent decay rate $\xi(T)$ is approximated by a first-order decay function[54,55], which gives: $\xi(T) = \frac{\ln 2}{\eta(T)}$. Assuming that the viral RNA decays in an exponential manner $V' = -\xi(T)V$ or $V(t) = V_0 e^{-\xi(T)t}$ [34,39,53,56], the fraction of virus decay between shedding and collection is given by:

$$\gamma(T) = \frac{V_0 - V\left(t_{\text{arrive}}\right)}{V_0} = 1 - e^{-\xi(T)t_{\text{arrive}}} \tag{8}$$

In Eq. 8, t_{arrive} is the time between shedding and sample collection, taken to be 18 h[34]. $\gamma(T)$ is estimated using daily wastewater temperature data and arrival time. $V(t)$ is a proxy variable for the cumulative virus shed in wastewater. The viral load measured on day t, i.e., $\Delta V_t = V(t) - V(t-1)$, is compared to 24-h composite wastewater viral load data. This approach is inspired by how daily incidence is tracked through cumulative incidence in classic epidemiological models. Accordingly, we also estimate daily incidence using the cumulative number of infectious individuals via another proxy variable $C(t)$ with $C'(t) = kE$. Similarly, the daily incidence on day t is given by $\Delta C_t = C(t) - C(t-1)$. While this is a standard approach, ΔC_t does not perfectly correspond to the daily reported cases due to factors such as symptom onset, testing capacity, and underreporting rate[26,57].

The population-level viral shedding function

SARS-CoV-2 dynamics in the upper respiratory tract can be described by standard models of viral dynamics[15–17,19,58], which could be extended to the GI tract[10]. However, for simplicity, we describe the viral shedding from the

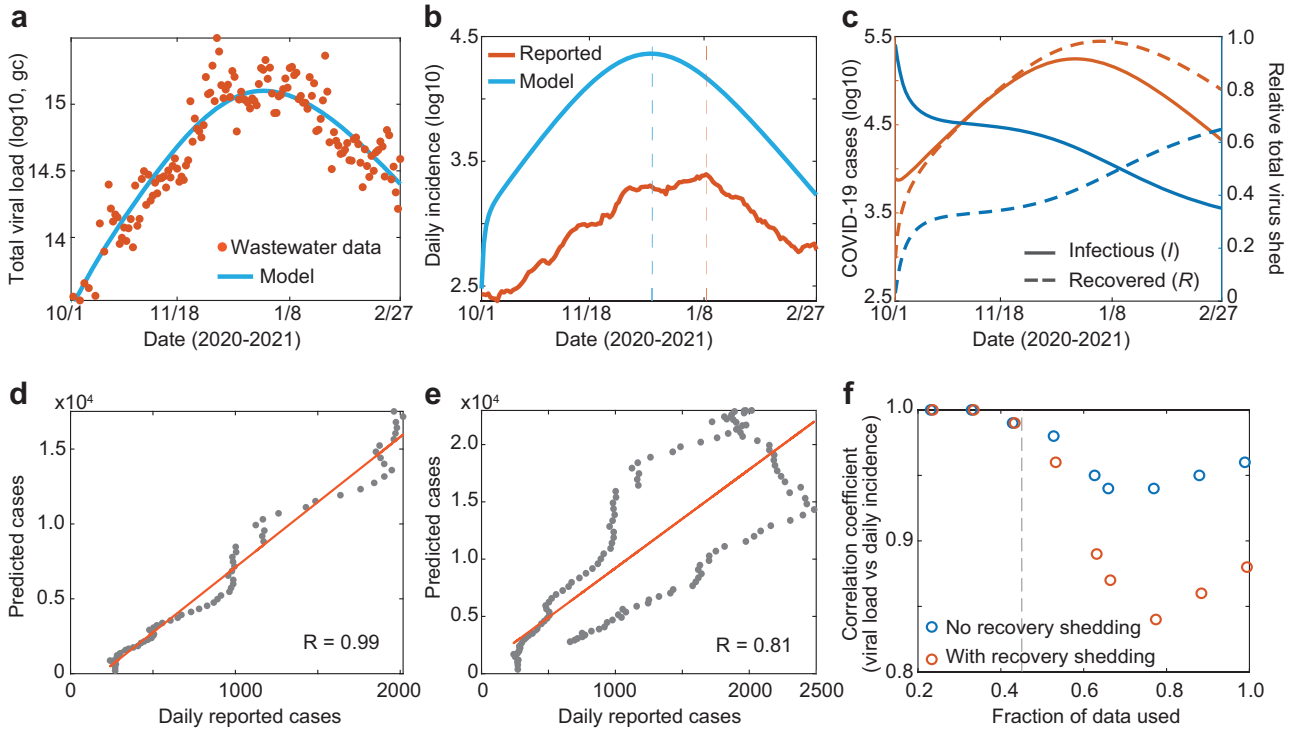

Fig. 1 | The impact of post-recovery viral shedding on the relationship between wastewater data and reported case data. a Best fit (blue curve) to daily measurements of viral RNA (brown circles) in wastewater. **b** Model-inferred daily incidence (blue curve) vs. reported daily incidence (brown curve). The vertical dashed lines mark the transmission peak time. **c** Model-predicted number of infectious/recovered (brown solid/dashed curve) and their relative contribution to the viral load in wastewater (blue solid/dashed curve). **d**, **e** Correlation between model-inferred and reported daily incidence, using only pre-transmission-peak data (**d**) versus using the entire wave data (**e**). Case report data were not used for fitting or informing any model parameters and initial conditions. **f** Model-predicted correlation between daily viral measurements in wastewater and daily incidence with and without the post-recovery shedding. The vertical dashed line marks the transmission peak.

GI tract using a simple phenomenological function (Eq. 9)[34].

$$f(t) = \frac{\omega_1 t}{\omega_2^2 + t^2} \tag{9}$$

The function $f(t)$ peaks at $\frac{\omega_1}{2\omega_2}$ when $t = \omega_2$. Here, ω_1 (log10 viral RNA copy per g) is a magnitude modifier and ω_2 (day) influences the timing and magnitude of the viral shedding peak. The viral shedding rate represents the viral concentration (per g of fecal matter) per day. This function has been shown to be a good approximation for the viral shedding dynamics of the ancestral SARS-CoV-2 strain up to about one month post infection[13,34]. Using this function, the total shedding between times t_1 and t_2 is given by

$$\int_{t_1}^{t_2} f(t)dt = \frac{\omega_1}{2}\log\left(\frac{\omega_2^2 + t_2^2}{\omega_2^2 + t_1^2}\right) \tag{10}$$

Here $0 \le t_1 < t_2$ and log denotes the natural logarithm. The unit of $f(t)$ is log10 viral RNA copy per g per day, meaning the integral $\int f(t)dt$ has a unit of log10 viral RNA copy per g. The average exposed E, infectious I, and recovered R phases of SARS-CoV-2 infection are classified based on the time since infection, roughly corresponding to viral progression in the upper respiratory tract[34]. Thus, we define t_E, t_I, and t_R (or $1/k$,$1/\delta$, and $1/\sigma$) to be the average time an infected person spends in these infection phases, respectively. The average viral shedding rate during the infectious phase β_I is:

$$\beta_I = \frac{1}{t_I}\int_{t_E}^{t_E+t_I} f(t)dt = \frac{1}{t_I}\frac{\omega_1}{2}\log\left(\frac{\omega_2^2 + (t_E + t_I)^2}{\omega_2^2 + t_E^2}\right) \tag{11}$$

The unit of β_I is log10 viral RNA copy per g per day. Based on the best-fit to GI tract viral shedding data[13,34], we have $\omega_1 = 71.97$ log10 viral RNA copy per g and $\omega_2 = 4$ days. The estimated value of β_I with $t_E = 3$ days and $t_I = 8$ days is 7.65 log10 viral RNA copy per g per day, consistent with previous estimates[59,60]. However, there is considerable uncertainty in the estimate of the viral shedding profile due to the lack of samples throughout the course of infection and a rudimentary understanding of how GI viral shedding relates to the standard course of infection (exposed, infectiousness, recovered). Thus, instead of using precise values for the viral shedding rate β_R and the duration t_R of the recovered phase, we will examine a range of values of β_R relative to β_I and t_R relative to t_I. For instance, the example in Fig. 1f assumes $\beta_R = 0.5 \times \beta_I$ and $t_R = 14$ days.

Data fitting

When fitting the SEIR-V model to the viral RNA data in wastewater, we minimize the sum of square error (SSE) between the measurement collected every 24 h and the total virus produced at each data point (Eq. 12). $\hat{V}(t_d)$ is the total viral load, e.g., viral RNA concentration × total flow, measured on day t_d. $V(t)$, as given by Eq. 1, is a proxy of the cumulative viral RNA in wastewater. Thus, the viral RNA produced daily, given by $\int_{t_d-1}^{t_d} V'(s)ds$, or ΔV_t, is the comparable quantity to $\hat{V}$.

$$\mathrm{SSE} = \sum_{t_d}\left(\log\int_{t_d-1}^{t_d} V'(s)ds - \log\hat{V}(t_d)\right)^2 \tag{12}$$

The minimization is done using the function *fmincon* and *multistart* (with 500 random initial guesses) in MATLAB. The initial susceptible population $S(0)$ is fixed to 2.3×10^6 persons[1]. The best fit in Fig. 1a gives $\lambda = 9.36 \times 10^{-8}$ per day per person, $\alpha = 126$ g, and $E(0) = 145$ people.

Table 1 | Parameters in the model

	Definition	Unit	Value/range	References
S	Susceptible population	Persons	$S(0) = 2.3 \times 10^6$ or 10^6 (*)	1
E	Exposed population	Persons	$E(0) = 0$ or 1 (*)	
I	Infectious population	Persons	$I(0) = \frac{V(0)}{\alpha\beta(1-\gamma)}$ or 1 (*)	
R	Recovered population	Persons	$R(0) = 0$ or 1 (*)	
λ	Transmission rate	Per day per person	Fitting	
$1/k$	Exposed duration	Day	3	11,13,99
$1/\delta$	Infectious duration	Day	8	11,13,99
$1/\sigma$	Recovered duration	Day	14 or $[7, 84]$ (**)	
α	Daily fecal load	Gram	Fitting: $[51, 796]$	
β_I	Infectious viral shedding rate	Viral RNA copies per gram	7.65	34
β_R	Recovered viral shedding rate	Viral RNA copies per gram	0.001 – 1 (***)	
$\gamma(T)$	Fraction of viral loss in sewer	Unitless	Eqs. (6) and (8)	
ω_1	Magnitude modifier	Log_{10} viral RNA per gram	71.97	13,34
ω_2	Peak timing for viral shedding	Day	4	13,34

Note: single asterisk (*) indicates $E(0), I(0), R(0)$ are fixed to the first value in the model fitting and to the second value in the simulation experiment. Double asterisk (**) indicates $1/\sigma$ is fixed to 14 for the fitting but otherwise vary over the provided range in the simulation. Triple asterisk (***) indicates β_R is a fraction of β_I.

The initial infected population is approximated by $I(0) = \frac{V(0)}{\alpha\beta(1-\gamma)}$, where $V(0)$ is taken as the first viral RNA measurement and $R(0)$ is assumed to be 0. The parameters k and δ are fixed to 1/3 and 1/8 per day, respectively[11], and σ is fixed to 1/14 per day for the scenario in Fig. 1. A summary of model parameters is provided in Table 1.

Theoretical correlation between daily viral measurement and incidence with and without recovered shedding

We use the model to demonstrate how post-recovery shedding β_R affects the correlation between viral measurements and incidence. The methods used to relate daily viral measurements to incidence data vary in literature[4,61]. Here, we will simply look at the correlation of daily viral measurement (ΔV_t) and incidence (ΔC_t). To produce Fig. 1f, we use the best-fit parameters for Fig. 1a. The model without post-recovery shedding has $\beta_R = 0$.

Simulation of the impact of post-recovery shedding on the inference capability of WBE

We systematically simulate the model with variations in β_R and σ (or $1/t_R$) to assess the impact of post-recovery shedding on model-inferred transmission dynamics. All other parameters in the model are fixed as follows: $k = 1/3$ per day, $\delta = 1/8$ per day, $\lambda = 4 \times 10^{-7}$ per person per day, $\alpha = 125$ g, $\gamma(T) = 0.1$, $\beta_I = 7.65$ log10 viral RNA copy per g per day, $S(0) = 10^6$ persons, $E(0) = 1$ person, $I(0) = 1$ person, $R(0) = 0$ person, and $V(0) = 0$ viral RNA copies. These simplified initial conditions and parameter values align with the beginning of an outbreak.

For the detection of an emerging variant, we limit the analysis to the end of an outbreak, where the virus shed by the recovered population contributes close to 100% of viral measurements. In this scenario, to detect a notable shift in the wastewater viral load, we need enough new infections (or infectious cases) that can generate a notable fraction κ of the virus shed by the recovered population on day t_d denoted by:

$$C_R(t_d) = \int_{t_d-1}^{t_d} \alpha(1-\gamma)\beta_R R(t)dt \tag{13}$$

Let $N(t_d)$ be the minimum number of new infections needed to match the viral signal from the recovered population, then $N(t_d)$ must satisfy:

$$N(t_d)\alpha(1-\gamma)\beta_I > \kappa C_R(t_d) \tag{14}$$

Solving for N, we have

$$N(t_d) > \frac{\kappa C_R(t_d)}{\alpha(1-\gamma)\beta_I} \tag{15}$$

For reporting purposes (Fig. 2c), we use $\kappa = 0.05$ (or 5%) and round up $N(t_d)$ to the nearest integer. The heat map (Fig. 2c) is generated by calculating $N(t_d)$ for the average post-recovery shedding rate β_R varying between $0.001\beta_I$ and β_I and the average post-recovery shedding duration σ varying between 1/7 and 1/84 per day. To represent the end of an outbreak, we use the time t_f when the cumulative recovered population given by $\int_0^{t_f} \delta I(s)ds$ reaches 99.9% of its final size $\lim_{t\to\infty}\int_0^t \delta I(s)ds$.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Results

Post-recovery viral shedding alters the correlation of wastewater viral load and case report data

To reconstruct the temporal dynamics of the infectious and recovered populations in the sewershed, we fit the SEIR-V model to wastewater viral load data. As shown in Fig. 1a, the model accurately captures the wastewater viral load dynamics, and the inferred daily incidence mirrors the reported incidence data. Comparison of peak timing, as marked by the vertical dashed lines in Fig. 1b, suggests model-predicted transmission dynamics precede reported incidence by about 3 weeks. The model-inferred incidence also exceeds the reported daily incidence by 3–5 times in early October and late February and by over 10 times in December. These findings are consistent with the expected lead time of wastewater data relative to case report data and the underreported incidence rates[62,63]. However, without further analysis, the inferred lead-time and underreported rates should not be taken as definitive due to noisy case report data around the peak and possible issue of model identifiability, as discussed in Supplementary Note A1.

As the outbreak unfolds, the number of recovered individuals accumulates steadily, ultimately surpassing the number of infectious individuals just before the transmission peak, depicted in Fig. 1c. Consequentially, while the infectious population is the primary source of viral RNA in wastewater at the beginning of the epidemic wave, shortly after the transmission peak, the viral contribution from the recovered population becomes predominant. This turns out to be a logical consequence of the natural progression that

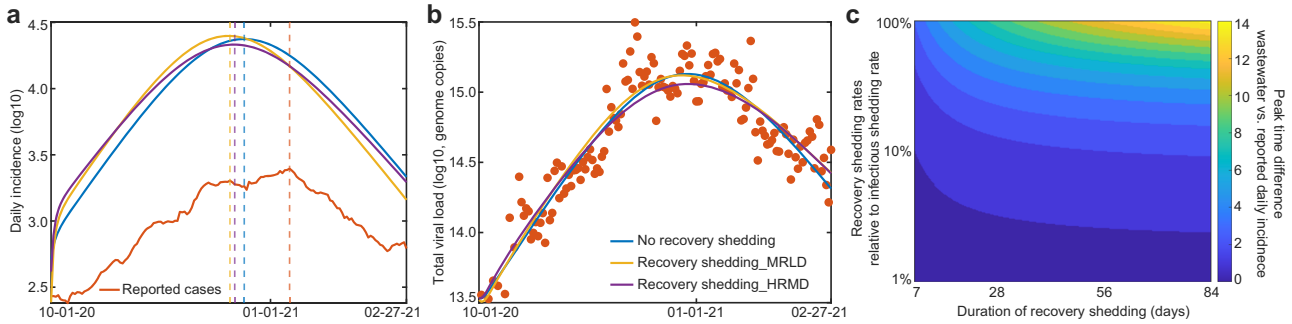

Fig. 2 | The impact of post-recovery shedding on model-inferred transmission dynamics. a Model-inferred daily incidence under three scenarios with varied post-recovery shedding. The vertical dashed curves mark the transmission peak times. **b** Model best fits to viral RNA in wastewater under different recovered shedding scenarios. MRLD: moderate post-recovery shedding rate (25% of the infectious shedding rate) with a long duration of 21 days; HRMD: high post-recovery shedding rate (50% of the infectious shedding rate) for a moderate duration of 14 days. **c** The difference in viral RNA peak in wastewater and transmission peak varies with post-recovery shedding rate and duration. The color represents the lead time difference with warmer colors indicating a larger lead time.

infected individuals go through (exposed, infectious, and recovered phases), which is shown mathematically in Supplementary Note A2. Moreover, post-recovery shedding rate and duration influence when the recovered population starts to shed more viruses than the infectious population, with higher post-recovery shedding rates or durations resulting in an earlier time.

Next, we examined the relationship between SARS-CoV-2 wastewater viral load and case report data, which has been widely demonstrated to be strongly correlated[1–3]. We hypothesized that as the outbreak progresses, the shift toward the recovered population as the predominant viral source alters this correlation. We tested this hypothesis by fitting the same model to the pre-transmission-peak wastewater data versus the entire data set. This reveals a clear decrease in the correlation between the model-inferred daily incidence and the reported daily incidence as more post-transmission-peak data is included, see Fig. 1d, e. Additionally, the theoretical correlation between daily viral measurements and incidence, depicted in Fig. 1f, shows that the inclusion of viral shedding by the recovered population results in a faster decline of the correlation coefficient as the outbreak unfolds. This supports the notion that the relationship between wastewater and case report data within a single outbreak is dynamic and nonlinear[64,65].

It is worth noting that the increasing trend near the end is a result of the convergence of daily wastewater viral load and incidence to zero at the end of the outbreak. Equally interesting is the clear decreasing trend in correlation – albeit to a lesser extent, around the time of the transmission peak observed in Fig. 1f, even when the recovered population is assumed to not shed the virus! This stems from an intrinsic asynchronization in the natural infection and viral shedding processes, as thoroughly explained in Supplementary Note A3.

Inclusion of post-recovery shedding results in earlier estimates for the timing of transmission peak

Due to the heterogeneity in viral shedding, we performed sensitivity analyses on the duration and magnitude of post-recovery shedding. Figure 2a shows the model-inferred daily incidence based on the best fit to viral load in wastewater under three scenarios: (1) no shedding from recovered individuals, (2) MRLD: moderate post-recovery shedding rate (25% of the infectious shedding rate) with a long duration of 21 days, and (3) HRMD: high post-recovery shedding rate (50% of the infectious shedding rate) for a moderate duration of 14 days. The model best fits under the three scenarios are similar, but the higher shedding rates and longer shedding durations (MRLD and HRMD) result in slightly slower decreasing trends in the total viral load in Fig. 2b. A remarkable difference is observed in the daily incidence. The inclusion of viral shedding by the recovered population shifts the inferred transmission peak 7 days earlier under the MRLD scenario and 5 days earlier under the HRMD scenario, as indicated by the vertical dashed lines in Fig. 2a. This suggests that post-recovery viral shedding can considerably affect the estimation of the transmission peak, potentially modulating the lead time of wastewater to case data.

To better understand this observation, we tested varying shedding duration ranging from 7 to 84 days and post-recovery shedding rate from 1% to 100% of the infectious shedding rate. Within this parameter landscape, we compared the lead time between the viral RNA peak in wastewater and the daily incidence peak. The analysis clearly shows that higher shedding magnitudes and/or longer durations of the recovery phase result in a larger difference in the peak time, as depicted by the warmer colors in Fig. 2c. This supports the notion that the shedding rate and duration can have a substantial impact on the estimated lead time difference between the wastewater measurement and incidence data, and in a predictable manner. We mathematically demonstrated this property in Supplementary Note A4.

The virus shed by the recovered population may lead to a delay in the detection of an emerging variant by wastewater data

Several studies have investigated the minimum number of new cases required to detect viral signal in wastewater[61,66–68], which applies to the detection of emerging outbreaks. However, none has looked into this issue taking into consideration the virus shed by the recovered population. This is especially important when a new variant emerges near the end of an outbreak, as illustrated in Fig. 3a. In this scenario, the virus shed by a sufficiently large number of recovered individuals can overshadow the virus shed by the new cases infected with the emerging variant, therefore delaying the recognition of the emergence of the new variant based on the wastewater viral load data. Figure 3b shows that the number of cases infected with a new variant X required to signal an emerging outbreak is an increasing function of post-recovery viral shedding rate and duration associated with the existing strain. The more infections needed to reach a detectable level, the longer the delay in identifying an emerging outbreak. This analysis extends to scenarios where one wishes to estimate in near real-time the emergence speed, or doubling rate, of variant X based on wastewater data.

Discussion

In this study, we showed that viral shedding dynamics are crucial for the precise and accurate implementation and interpretation of wastewater-based viral load data. In particular, we demonstrated that post-recovery shedding introduces nonlinearities that continuously alter the relationship between wastewater viral load and case report data throughout an outbreak. Neglecting this factor could lead to inaccurate representations of transmission dynamics and delay the detection of emerging variants when analyzing wastewater viral load data. Our study provides a framework to quantify the impact of viral shedding on wastewater-based epidemiological inferences, such as estimates of new infections and predictions of outbreak progression.

Our results suggest that the relationship between viral RNA in wastewater and case report data is dynamically nonlinear over the course of an outbreak. While this observation has been noted in previous

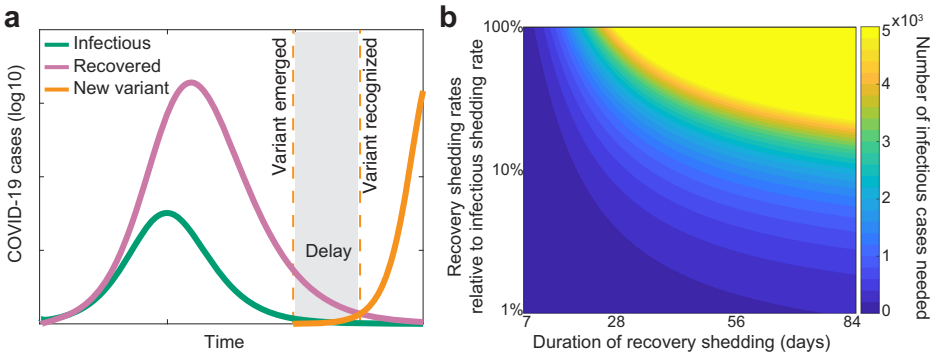

Fig. 3 | The potential delay in the detection of an emerging variant X due to the virus shed by the recovered population. a An illustration of the challenge of tracking an emerging variant X near the end of an outbreak when viral signal is mainly contributed by the recovered population. The extent of detection delay depends on the viral shedding rate and duration of the recovered individuals from the existing strain. **b** The number of new cases needed to produce a sufficiently large viral signal as a function of post-recovery viral shedding (Eqs. 13–15). Warmer color indicates higher numbers of cases needed.

studies[4,36,64,65,69–71], we provided a mechanistic basis driving this dynamic relationship. During the exponential growth phase of an outbreak, the viral load in wastewater is primarily driven by two factors: the rapid increase in new cases, which reflects the exponential growth of the outbreak, and the high viral shedding rate during the infectious phase. Thus, the trends of viral RNA in wastewater and reported daily incidence mirror each other in the early phases, leading to a high correlation. However, as the outbreak slows down around the time of transmission peak, the rate of new infections declines while recoveries accumulate. This leads to an increase in the relative contribution of virus shed into wastewater by the recovered population, meaning wastewater viral load no longer reflects just the number of new cases, but also an increasing number of recovered cases. As a result, the correlation between viral RNA in wastewater and case report data changes. For readers interested in the mathematical details, we refer to Supplementary Note A5.

The near-linear relationship between wastewater and case report data in the exponential growth phase suggests that not only can the trend of wastewater viral load be simply interpreted as the growth trend of an outbreak in the initial phase, but it can also capture important early characteristics, such as R_0[72,73]. Furthermore, we hypothesize that, if viral shedding is most concentrated during the infectious phase, one could simply use the exponential growth phase of wastewater data to quickly assess the severity of the outbreak, especially when case report data is unavailable or limited. However, this linear relationship is short-lived and followed by a more complex, nuanced phase where post-recovery shedding becomes an important factor. Analyses based on wastewater data can be biased without accounting for this factor. For instance, suppose we wish to use only wastewater data in the later phase, without knowing the post-recovery shedding, to estimate the effective reproduction number R_e (the expected number of new infections caused by an infectious individual at any given time). Then, because the trends observed in wastewater data vary with different viral shedding profiles even for the same transmission dynamics (see an illustration in Fig. 4), the estimation of R_e also varies based on the assumed viral shedding profile. This observed dependency is analogous to the dependence of R_e estimates on asymptomatic SARS-CoV-2 cases[74,75]. Refining models to incorporate these shedding dynamics will significantly improve the accuracy of wastewater-based epidemic monitoring.

Post-recovery shedding may also delay the identification of emerging outbreaks or variants. To proactively minimize this delay, we need a better understanding of the variant-specific viral shedding dynamics to enable strategies that adapt testing capacity according to different phases of an outbreak. For instance, integrating modeling analysis of routine wastewater surveillance data and genomic sequencing of wastewater samples[76–79] may better facilitate the identification of emerging variants circulating in the community.

One limitation of this study is the uncertainty surrounding GI tract viral shedding. The viral shedding rate (of infectious individuals) is often estimated based on the tail of the shedding distribution due to the scarcity of GI tract viral shedding data prior to symptom onset. However, without the constraints of shedding data during the initial infection phase, accurately determining the incline phase of the viral shedding curve is challenging[80,81]. There could be an infinite number of viral shedding profiles with different early kinetics that all have the same tail, as depicted in Supplementary Fig. 1. This could be problematic because viral dynamics, at least in the case of nasal viral load, are highly heterogeneous[24,25], which is further aggravated by the differences in viral shedding dynamics among SARS-CoV-2 variants[26,82]. Additionally, GI tract viral shedding data alone is not sufficient to fully connect wastewater to case report data. This is because there is no well-established GI tract viral shedding according to the disease stages for SARS-CoV-2. In fact, even the first and only comprehensive human challenge study for SARS-CoV-2 did not report GI tract viral shedding data[11]. This problem could be alleviated by examining the impact of the uncertainty around GI viral shedding dynamics on model inference based on wastewater data. For this reason, perhaps noroviruses would be a more appropriate disease choice for the initial development of WBE, given that data on viral shedding, symptomology, and disease severity are available[83–85].

Our analyses did not include the viral shedding during the exposed phase, or the incubation period, which could impact the relationship between wastewater and case report data. However, SARS-CoV-2 infected individuals in the incubation period likely shed virus in fecal matter at a much lower rate compared to infectious, and for a much shorter duration compared to recovered individuals[34]. Thus, the virus shed by an exposed individual during an ongoing outbreak is likely negligible, which leads to a small intrinsic detection delay of new cases using WBE. Nevertheless, case report data is also affected by reporting delays due to the typical asymptomatic nature of the incubation phase. Thus, the effect of viral shedding during the exposed phase might have a minimal impact on the inferences that use wastewater and case report data.

Another potentially important factor not addressed in our analyses is instances of non-monotonic viral shedding (as illustrated in the upper right plot in Fig. 4a). This phenomenon has been observed in numerous clinical cohorts of SARS-CoV-2 infected individuals with and without treatments[25,86–88]. The increase(s) in viral shedding in the later phase further complicates the analysis, as it paints the illusion of multiple infections. Nevertheless, because non-monotonic viral shedding for SARS-CoV-2 is often associated with treatments[58,86], it needs to be integrated into realistic WBE modeling frameworks to capture pharmaceutical interventions. The extent of its impact requires further studies.

Additional factors such as the frequency of sampling, reporting rates and delays, vaccination, the presence of different variants, mobility patterns within and between sewersheds, and geographical variations can also affect the relationship between viral load and case report data[1,2,34,64,70,89,90]. While this study examines the effects of temperature and travel time on viral decay rate, the impact of other in-sewer parameters, including organic matter content, particle concentration, pH levels, solvent presence, detergent concentrations, and microbial activity remains largely unexplored[91–94]. These knowledge gaps highlight the need for standardized sampling

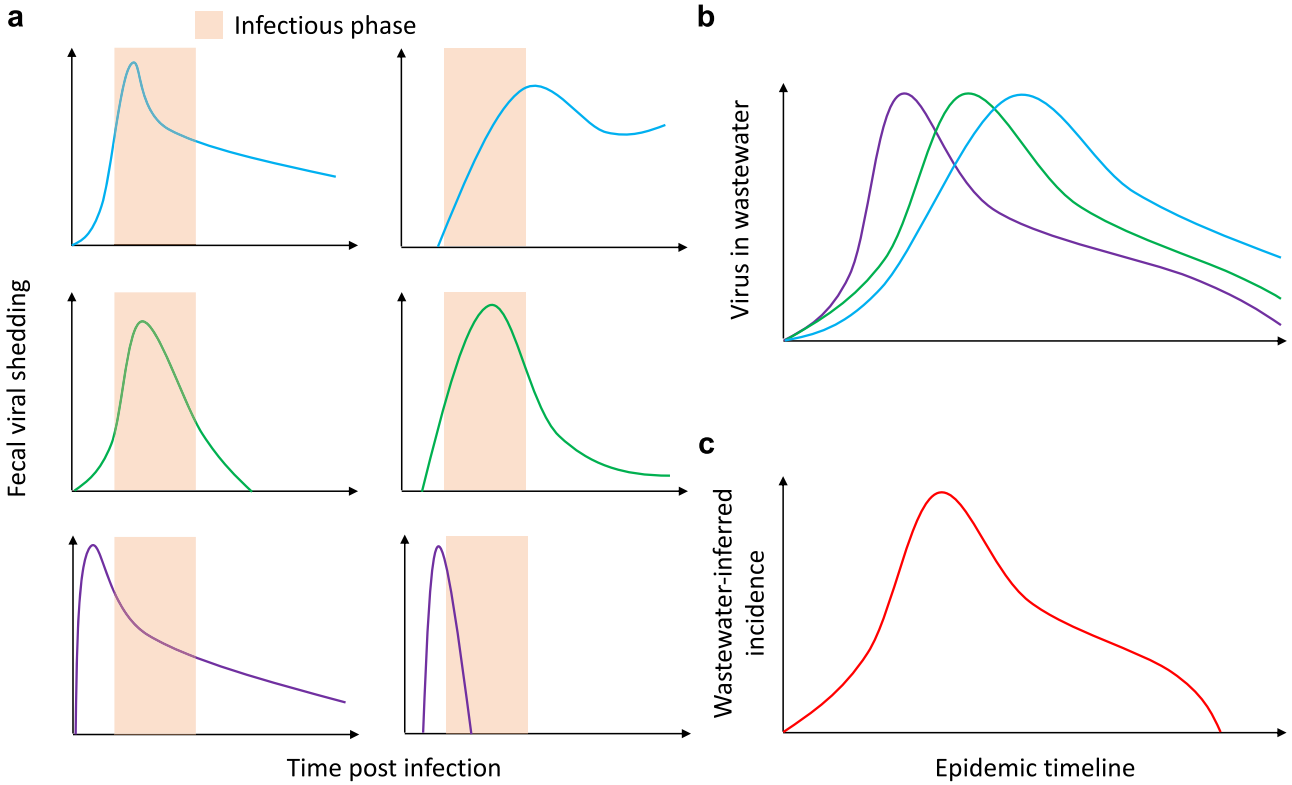

Fig. 4 | Illustration of how the viral shedding pattern may influence the relationship between viral load in wastewater and reported case data. a Various population-level viral shedding profiles for a generic disease. Blue illustrates prolonged viral shedding. Green represents concentrated viral shedding during the infectious phase. Purple shows viral shedding concentrated in the exposed phase. **b** Each profile yields a distinct pattern of longitudinal viral load in wastewater over the course of an epidemic. **c** However, these diverse patterns of wastewater viral load may reflect identical underlying incidence trends. Conversely, a single transmission pattern may also result in different viral load trends in wastewater, depending on the underlying viral shedding dynamics.

protocols and comprehensive integration of demographic and environmental data to improve the accuracy and reliability of WBE as a tool for public health surveillance and early detection of emerging threats.

Although our study examines SARS-CoV-2, the framework for understanding how post-recovery viral shedding shapes inferences based on wastewater data is broadly applicable to other infectious diseases, such as norovirus[95], influenza[96], RSV[97], and mpox[98]. The model structure can be adapted for different pathogens by incorporating specific viral shedding characteristics, symptomology, and transmission dynamics. To summarize, we illustrate our results and hypothesis of how various viral shedding profiles for a generic disease affect the relationship between wastewater and case report data in Fig. 4. This hypothesis highlights the importance of a thorough understanding of viral shedding dynamics on the development of wastewater-based epidemiology.

Data availability

All data used in this study are included in the Supplementary Data. Source data for Fig. 1 are available in Supplementary Data 1, for Fig. 2 in Supplementary Data 2, and for Fig. 3 in Supplementary Data 3. The reported COVID-19 case data in the sewershed are publicly available and can be freely accessed without restrictions through the official Massachusetts government website (www.mass.gov).

Code availability

All code developed and utilized in this study is provided in the Supplementary Data. Analysis code for Fig. 1 is available in Supplementary Data 1, for Fig. 2 in Supplementary Data 2, and for Fig. 3 in Supplementary Data 3.

Received: 28 August 2024; Accepted: 12 May 2025;
Published online: 22 May 2025

References

1. Wu, F. et al. SARS-CoV-2 RNA concentrations in wastewater foreshadow dynamics and clinical presentation of new COVID−19 cases. *Sci. Total Environ.* **805**, 150121 (2022).
2. Mattei, M., Pintó, R. M., Guix, S., Bosch, A. & Arenas, A. Analysis of SARS-CoV-2 in wastewater for prevalence estimation and investigating clinical diagnostic test biases. *Water Res.* **242**, 120223 (2023).
3. Ai, Y. et al. Wastewater SARS-CoV-2 monitoring as a community-level COVID−19 trend tracker and variants in Ohio, United States. *Sci. Total Environ.* **801**, 149757 (2021).
4. Xiao, A. et al. Metrics to relate COVID−19 wastewater data to clinical testing dynamics. *Water Res.* **212**, 118070 (2022).
5. Kilaru, P. et al. Wastewater surveillance for infectious disease: a systematic review. *Am. J. Epidemiol.* **192**, 305–322 (2023).
6. Varkila, M. R. J. et al. Use of wastewater metrics to track COVID-19 in the US. *JAMA Netw. Open* **6**, e2325591 (2023).
7. Gitter, A. et al. Not a waste: wastewater surveillance to enhance public health. *Front. Chem. Eng.* **4**, 1112876 (2023).
8. Peccia, J. et al. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nat. Biotechnol.* **38**, 1164–1167 (2020).
9. Chen, C. et al. Wastewater-based epidemiology for COVID-19 surveillance and beyond: a survey. *Epidemics* **49**, 100793 (2024).
10. Phan, T. et al. Making waves: Integrating wastewater surveillance with dynamic modeling to track and predict viral outbreaks. *Water Res.* **243**, 120372 (2023).
11. Killingley, B. et al. Safety, tolerability and viral kinetics during SARS-CoV-2 human challenge in young adults. *Nat. Med.* **28**, 1031–1041 (2022).

12. Gunawardana, M. et al. Early SARS-CoV-2 dynamics and immune responses in unvaccinated participants of an intensely sampled longitudinal surveillance study. *Commun. Med.* **2**, 129 (2022).
13. Wölfel, R. et al. Virological assessment of hospitalized patients with COVID-2019. *Nature* **581**, 465–469 (2020).
14. Perelson, A. S. & Ke, R. Mechanistic modeling of SARS-CoV-2 and other infectious diseases and the effects of therapeutics. *Clin. Pharmacol. Ther.* **109**, 829–840 (2021).
15. Ke, R., Zitzmann, C., Ho, D. D., Ribeiro, R. M. & Perelson, A. S. In vivo kinetics of SARS-CoV-2 infection and its relationship with a person's infectiousness. *Proc. Natl. Acad. Sci. USA* **118**, e2111477118 (2021).
16. Marc, A. et al. Quantifying the relationship between SARS-CoV-2 viral load and infectiousness. *eLife* **10**, e69302 (2021).
17. Heitzman-Breen, N. & Ciupe, S. M. Modeling within-host and aerosol dynamics of SARS-CoV-2: the relationship with infectiousness. *PLoS Comput. Biol.* **18**, e1009997 (2022).
18. Goyal, A., Reeves, D. B., Cardozo-Ojeda, E. F., Schiffer, J. T. & Mayer, B. T. Viral load and contact heterogeneity predict SARS-CoV-2 transmission and super-spreading events. *eLife* **10**, e63537 (2021).
19. Iyaniwura, S. A. et al. The kinetics of SARS-CoV-2 infection based on a human challenge study. *Proc. Natl. Acad. Sci. USA* **121**, e2406303121 (2024).
20. Sanche, S. et al. A simple model of COVID-19 explains disease severity and the effect of treatments. *Sci. Rep.* **12**, 14210 (2022).
21. Gandhi, R. T., Lynch, J. B. & Del Rio, C. Mild or moderate Covid-19. *N. Engl. J. Med.* **383**, 1757–1766 (2020).
22. Avila, J., Long, B., Holladay, D. & Gottlieb, M. Thrombotic complications of COVID-19. *Am. J. Emerg. Med.* **39**, 213–218 (2021).
23. Zhang, J., Dong, X., Liu, G. & Gao, Y. Risk and protective factors for COVID-19 morbidity, severity, and mortality. *Clin. Rev. Allergy Immunol.* **64**, 90–107 (2022).
24. Ke, R. et al. Daily longitudinal sampling of SARS-CoV-2 infection reveals substantial heterogeneity in infectiousness. *Nat. Microbiol.* **7**, 640–652 (2022).
25. Hay, J. A. et al. Quantifying the impact of immune history and variant on SARS-CoV-2 viral kinetics and infection rebound: a retrospective cohort study. *eLife* **11**, e81849 (2022).
26. Puhach, O., Meyer, B. & Eckerle, I. SARS-CoV-2 viral load and shedding kinetics. *Nat. Rev. Microbiol*. https://doi.org/10.1038/s41579-022-00822-w (2022).
27. Owens, K., Esmaeili, S. & Schiffer, J. T. Heterogeneous SARS-CoV-2 kinetics due to variable timing and intensity of immune responses. *JCI Insight* **9**, e176286 (2024).
28. Natarajan, A. et al. Gastrointestinal symptoms and fecal shedding of SARS-CoV-2 RNA suggest prolonged gastrointestinal infection. *Med* **3**, 371–387.e9 (2022).
29. Gupta, S., Parker, J., Smits, S., Underwood, J. & Dolwani, S. Persistent viral shedding of SARS-CoV-2 in faeces – a rapid review. *Colorectal Dis.* **22**, 611–620 (2020).
30. Cevik, M. et al. 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 (2021).
31. Zhang, Y. et al. 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 (2021).
32. Du, W. et al. Persistence of SARS-CoV-2 virus RNA in feces: a case series of children. *J. Infect. Public Health* **13**, 926–931 (2020).
33. Arts, P. J. et al. Longitudinal and quantitative fecal shedding dynamics of SARS-CoV-2, pepper mild mottle virus, and crAssphage. *mSphere* **8**, e00132–23 (2023).
34. Phan, T. et al. A simple SEIR-V model to estimate COVID-19 prevalence and predict SARS-CoV-2 transmission using wastewater-based surveillance data. *Sci. Total Environ.* **857**, 159326 (2023).
35. Pell, B., Brozak, S., Phan, T., Wu, F. & Kuang, Y. The emergence of a virus variant: dynamics of a competition model with cross-immunity time-delay validated by wastewater surveillance data for COVID-19. *J. Math. Biol.* **86**, 63 (2023).
36. Brouwer, A. F. et al. The role of time-varying viral shedding in modelling environmental surveillance for public health: revisiting the 2013 poliovirus outbreak in Israel. *J. R. Soc. Interface* **19**, 20220006 (2022).
37. Ahmadini, A., Msmali, A., Mutum, Z. & Raghav, Y. S. The mathematical modeling approach for the wastewater treatment process in Saudi Arabia during COVID-19 pandemic. *Discrete Dyn. Nat. Soc.* **2022**, 1–15 (2022).
38. Nourbakhsh, S. et al. A wastewater-based epidemic model for SARS-CoV-2 with application to three Canadian cities. *Epidemics* **39**, 100560 (2022).
39. McMahan, C. S. et al. COVID-19 wastewater epidemiology: a model to estimate infected populations. *Lancet Planet. Health* **5**, e874–e881 (2021).
40. Polcz, P. et al. Wastewater-based modeling, reconstruction, and prediction for COVID-19 outbreaks in Hungary caused by highly immune evasive variants. *Water Res.* **241**, 120098 (2023).
41. Proverbio, D. et al. Model-based assessment of COVID-19 epidemic dynamics by wastewater analysis. *Sci. Total Environ.* **827**, 154235 (2022).
42. Pájaro, M., Fajar, N. M., Alonso, A. A. & Otero-Muras, I. Stochastic SIR model predicts the evolution of COVID-19 epidemics from public health and wastewater data in small and medium-sized municipalities: a one year study. *Chaos Solitons Fractals* **164**, 112671 (2022).
43. Meadows, T. et al. Epidemiological model can forecast COVID-19 outbreaks from wastewater-based surveillance in rural communities. *Water Res.* **268**, 122671 (2025).
44. Pant, B., Safdar, S., Ngonghala, C. N. & Gumel, A. B. Mathematical assessment of wastewater-based epidemiology to predict SARS-CoV-2 cases and hospitalizations in Miami-Dade County. *Acta Biotheor.* **73**, 2 (2025).
45. Joung, M. J. et al. Coupling wastewater-based epidemiological surveillance and modelling of SARS-COV-2/COVID-19: practical applications at the Public Health Agency of Canada. *Can. Commun. Dis. Rep.* **49**, 166–174 (2023).
46. Hunter, E. & Kelleher, J. D. Understanding the assumptions of an SEIR compartmental model using agentization and a complexity hierarchy. *J. Comput. Math. Data Sci.* **4**, 100056 (2022).
47. Brauer, F. Some simple epidemic models. *Math. Biosci. Eng.* **3**, 1–15 (2006).
48. Eikenberry, S. E. et al. To mask or not to mask: modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infect. Dis. Model.* **5**, 293–308 (2020).
49. Brauer, F. & Castillo-Chavez, C. *Mathematical Models in Population Biology and Epidemiology*, Vol. 40 (Springer New York, 2012).
50. Saikia, D., Bora, K. & Bora, M. P. COVID-19 outbreak in India: an SEIR model-based analysis. *Nonlinear Dyn.* **104**, 4727–4751 (2021).
51. Saad-Roy, C. M. et al. Immune life history, vaccination, and the dynamics of SARS-CoV-2 over the next 5 years. *Science* **370**, 811–818 (2020).
52. Bivins, A. et al. Persistence of SARS-CoV-2 in water and wastewater. *Environ. Sci. Technol. Lett.* **7**, 937–942 (2020).
53. Hart, O. E. & Halden, R. U. 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 (2020).
54. Hiatt, C. W. Kinetics of the inactivation of viruses. *Bacteriol. Rev.* **28**, 150–163 (1964).
55. Ahmed, W. et al. Decay of SARS-CoV-2 and surrogate murine hepatitis virus RNA in untreated wastewater to inform application in wastewater-based epidemiology. *Environ. Res.* **191**, 110092 (2020).

56. De Oliveira, L. C. et al. Viability of SARS-CoV-2 in river water and wastewater at different temperatures and solids content. *Water Res.* **195**, 117002 (2021).
57. Lau, H. et al. Evaluating the massive underreporting and undertesting of COVID-19 cases in multiple global epicenters. *Pulmonology* **27**, 110–115 (2021).
58. Phan, T. et al. Modeling the emergence of viral resistance for SARS-CoV-2 during treatment with an anti-spike monoclonal antibody. *PLoS Pathog.* **20**, e1011680 (2024).
59. Han, M. S. et al. Viral RNA load in mildly symptomatic and asymptomatic children with COVID-19, Seoul, South Korea. *Emerg. Infect. Dis.* **26**, 2497–2499 (2020).
60. Schmitz, B. W. et al. Enumerating asymptomatic COVID-19 cases and estimating SARS-CoV-2 fecal shedding rates via wastewater-based epidemiology. *Sci. Total Environ.* **801**, 149794 (2021).
61. Wu, F. et al. Making waves: wastewater surveillance of SARS-CoV-2 in an endemic future. *Water Res.* **219**, 118535 (2022).
62. Wu, F. et al. SARS-CoV-2 titers in wastewater are higher than expected from clinically confirmed cases. *mSystems* **5**, e00614-20 (2020).
63. Angulo, F. J., Finelli, L. & Swerdlow, D. L. Estimation of US SARS-CoV-2 infections, symptomatic infections, hospitalizations, and deaths using seroprevalence surveys. *JAMA Netw. Open* **4**, e2033706 (2021).
64. Radu, E. et al. Emergence of SARS-CoV-2 Alpha lineage and its correlation with quantitative wastewater-based epidemiology data. *Water Res.* **215**, 118257 (2022).
65. Schill, R., Nelson, K. L., Harris-Lovett, S. & Kantor, R. S. The dynamic relationship between COVID-19 cases and SARS-CoV-2 wastewater concentrations across time and space: considerations for model training data sets. *Sci. Total Environ.* **871**, 162069 (2023).
66. Hewitt, J. et al. Sensitivity of wastewater-based epidemiology for detection of SARS-CoV-2 RNA in a low prevalence setting. *Water Res.* **211**, 118032 (2022).
67. Hong, P.-Y. et al. Estimating the minimum number of SARS-CoV-2 infected cases needed to detect viral RNA in wastewater: to what extent of the outbreak can surveillance of wastewater tell us? *Environ. Res.* **195**, 110748 (2021).
68. Nauta, M. et al. Early detection of local SARS-CoV-2 outbreaks by wastewater surveillance: a feasibility study. *Epidemiol. Infect.* **151**, e28 (2023).
69. Hegazy, N. et al. Understanding the dynamic relation between wastewater SARS-CoV-2 signal and clinical metrics throughout the pandemic. *Sci. Total Environ.* **853**, 158458 (2022).
70. Armas, F. et al. Contextualizing wastewater-based surveillance in the COVID-19 vaccination era. *Environ. Int.* **171**, 107718 (2023).
71. Peng, K. K. et al. An exploration of the relationship between wastewater viral signals and COVID-19 hospitalizations in Ottawa, Canada. *Infect. Dis. Model.* **8**, 617–631 (2023).
72. Chowell, G., Sattenspiel, L., Bansal, S. & Viboud, C. Mathematical models to characterize early epidemic growth: a review. *Phys. Life Rev.* **18**, 66–97 (2016).
73. Pell, B., Phan, T., Rutter, E. M., Chowell, G. & Kuang, Y. Simple multi-scale modeling of the transmission dynamics of the 1905 plague epidemic in Bombay. *Math. Biosci.* **301**, 83–92 (2018).
74. Saikia, D., Bora, K. & Bora, M. P. Counting the uncounted: estimating the unaccounted COVID-19 infections in India. *Nonlinear Dyn.* **112**, 9703–9717 (2024).
75. Jewell, N. P. & Lewnard, J. A. On the use of the reproduction number for SARS-COV-2: estimation, misinterpretations and relationships with other ecological measures. *J. R. Stat. Soc. Ser. A Stat. Soc.* **185**, S16–S27 (2022).
76. Karthikeyan, S. et al. Wastewater sequencing reveals early cryptic SARS-CoV-2 variant transmission. *Nature* **609**, 101–108 (2022).
77. Amman, F. et al. Viral variant-resolved wastewater surveillance of SARS-CoV-2 at national scale. *Nat. Biotechnol.* **40**, 1814–1822 (2022).
78. Vigil, K. et al. Long-term monitoring of SARS-CoV-2 variants in wastewater using a coordinated workflow of droplet digital PCR and nanopore sequencing. *Water Res.* **254**, 121338 (2024).
79. Xu, X. et al. High-resolution and real-time wastewater viral surveillance by nanopore sequencing. *Water Res.* **256**, 121623 (2024).
80. Ciupe, S. M. & Tuncer, N. Identifiability of parameters in mathematical models of SARS-CoV-2 infections in humans. *Sci. Rep.* **12**, 14637 (2022).
81. Zitzmann, C., Ke, R., Ribeiro, R. M. & Perelson, A. S. How robust are estimates of key parameters in standard viral dynamic models? *PLoS Comput. Biol.* **20**, e1011437 (2024).
82. Eyre, D. W. et al. Effect of Covid-19 vaccination on transmission of Alpha and Delta variants. *N. Engl. J. Med.* **386**, 744–756 (2022).
83. Atmar, R. L. et al. Norwalk virus shedding after experimental human infection. *Emerg. Infect. Dis.* **14**, 1553–1557 (2008).
84. Leon, J. S. et al. Randomized, double-blinded clinical trial for human norovirus inactivation in oysters by high hydrostatic pressure processing. *Appl. Environ. Microbiol.* **77**, 5476–5482 (2011).
85. Rouphael, N. et al. Dose-response of a norovirus GII.2 controlled human challenge model inoculum. *J. Infect. Dis.* **226**, 1771–1780 (2022).
86. Phan, T. et al. Modeling suggests SARS-CoV-2 rebound after nirmatrelvir-ritonavir treatment is driven by target cell preservation coupled with incomplete viral clearance. *J. Virol*. **99**, e0162324 (2025).
87. Edelstein, G. E. et al. SARS-CoV-2 virologic rebound with nirmatrelvir–ritonavir therapy: an observational study. *Ann. Intern. Med.* **176**, 1577–1585 (2023).
88. Anderson, A. S., Caubel, P. & Rusnak, J. M. Nirmatrelvir–ritonavir and viral load rebound in Covid-19. *N. Engl. J. Med.* **387**, 1047–1049 (2022).
89. Krivoňáková, N. et al. Mathematical modeling based on RT-qPCR analysis of SARS-CoV-2 in wastewater as a tool for epidemiology. *Sci. Rep.* **11**, 19456 (2021).
90. Sanjuán, R. & Domingo-Calap, P. Reliability of wastewater analysis for monitoring COVID-19 incidence revealed by a long-term follow-up study. *Front. Virol.* **1**, 776998 (2021).
91. Bertels, X. et al. Factors influencing SARS-CoV-2 RNA concentrations in wastewater up to the sampling stage: a systematic review. *Sci. Total Environ.* **820**, 153290 (2022).
92. Gundy, P. M., Gerba, C. P. & Pepper, I. L. Survival of coronaviruses in water and wastewater. *Food Environ. Virol.* **1**, 10 (2009).
93. Wiesner-Friedman, C. et al. Characterizing spatial information loss for wastewater surveillance using crAssphage: effect of decay, temperature, and population mobility. *Environ. Sci. Technol.* **57**, 20802–20812 (2023).
94. Chahal, C. et al. Pathogen and particle associations in wastewater. *Adv. Appl. Microbiol*. **97**, 63–119 (2016).
95. Walker, D. I. et al. Piloting wastewater-based surveillance of norovirus in England. *Water Res.* **263**, 122152 (2024).
96. Wolfe, M. K. et al. Wastewater-based detection of two influenza outbreaks. *Environ. Sci. Technol. Lett.* **9**, 687–692 (2022).
97. Hughes, B. et al. Respiratory syncytial virus (RSV) RNA in wastewater settled solids reflects RSV clinical positivity rates. *Environ. Sci. Technol. Lett.* **9**, 173–178 (2022).
98. Oghuan, J. et al. Wastewater analysis of Mpox virus in a city with low prevalence of Mpox disease: an environmental surveillance study. *Lancet Reg. Health Am.* **28**, 100639 (2023).
99. Van Kampen, J. J. A. et al. Duration and key determinants of infectious virus shedding in hospitalized patients with coronavirus disease-2019 (COVID-19). *Nat. Commun.* **12**, 267 (2021).

Acknowledgements
This work is supported by the U.S. National Science Foundation (DMS-2421257) and the National Institute of Drug Abuse (R01DA059394-01A1) to F.W. F.W. acknowledges partial support from the Texas Epidemic Public Health Institute (TEPHI). T.P. was supported by Director's postdoctoral fellowship at Los Alamos National Laboratory (20220791PRD2). Y.K. and S.B. are partially supported by the US National Science Foundation (DMS-2421258 and DEB-1930728) and the NIH grant (5R01GM131405-02). Y.K. is also supported by the eMB program (DMS-2325146). B.P. is supported by the U.S. National Science Foundation grant (DMS-2421260) and the LEAPS program (DMS-2316809). R.M.R. was supported by cooperative agreement CDC-RFA-FT-23-0069 from the CDC's Center for Forecasting and Outbreak Analytics. S.M.C. acknowledges partial support from the NIH NIGMS grant (1R01GM152743-01) and National Science Foundation grant (2051820). The contents of the manuscript are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention, the National Institutes of Health or any of the funders.

Author contributions
Conceptualization: T.P., S.M.C., R.K., A.S.P., F.W. Methodology: T.P., R.M.R., F.W. Software: T.P., S.B., B.P. Investigation: T.P., S.B., B.P. Validation: S.B., B.P., A.G. Funding acquisition: B.P., R.K., K.D.M., Y.K., F.W. Project administration: T.P., F.W. Supervision: R.K., R.M.R., A.S.P., Y.K., F.W. Visualization: T.P., F.W. Manuscript drafting: T.P., S.B., B.P., F.W. All authors reviewed the manuscript, contributed to critical revisions for important intellectual content, and approved the final version for submission.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43856-025-00908-5.

Correspondence and requests for materials should be addressed to Fuqing Wu.

Peer review information *Communications Medicine* thanks Madhurjya P. Bora and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.