

Environmental Science Water Research & Technology

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: K. Ensor, J. Palacio, S. A. Keller, R. Schneider, K. Domakonda, L. Hopkins and L. Stadler, *Environ. Sci.: Water Res. Technol.*, 2026, DOI: 10.1039/D5EW01170A.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

Water Impact

Wastewater surveillance provides a powerful approach for tracking viral infection dynamics in a community. Monitoring pathogens through sewer systems yields population-level information that informs public health guidance, resource allocation, and risk assessment. RSV wastewater monitoring provides timely estimates of the effective reproduction number and the number of infected individuals.



ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Understanding Transmission and Infections of Respiratory Syncytial Virus through Wastewater-Based Epidemiology[†]Katherine B. Ensor,^{*a} Jose Palacio,^b Sallie A. Keller,^c Rebecca Schneider,^d Kaavya Domakonda,^d Loren Hopkins,^e Lauren Stadler,^fReceived Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

Wastewater-based epidemiology (WBE) is an effective tool for tracking community circulation of respiratory viruses. We address a scientific gap that takes measured wastewater viral load of Respiratory Syncytial Virus (RSV) and estimate the effective reproduction number and the number of infections in the population. We advocate a modular approach to the analysis. We first estimate the trend and current level of the RSV viral load and quantify the uncertainty. These estimates become input for our Bayesian renewal model for both the infection rate and the number of infected individuals. The modular approach simplifies the analysis pipeline while maintaining scientific integrity. Further, the modular approach supports translation to other viruses by using disease-specific models for estimated transmission and cases in the second phase of the analysis.

1 INTRODUCTION.

Wastewater-based epidemiology (WBE) has become an essential tool for monitoring infectious diseases at the population level. By analyzing viral genetic material in wastewater, WBE offers a non-invasive and cost-effective way to detect early signs of community spread, regardless of whether individuals seek medical care or undergo clinical testing.¹

For respiratory syncytial virus (RSV), a leading cause of severe respiratory illness in infants, older adults, and immunocompromised people, clinical surveillance faces challenges from underreporting and delays in reporting. Many RSV cases are managed at home without hospitalization, so clinical records do not fully reflect community transmission. RSV measured in wastewater provides insight into community disease dynamics.² In a comprehensive study of 176 sites during the 2022-2023 RSV season, Zulli et al.³ observed that RSV RNA concentrations at state and national levels were linked to infection positivity and hospitalization rates. Allen et al.⁴ detail the implementation of a WBE approach in Northern Ireland to track RSV community transmission over the 2021 and 2022 seasons, correlating wastewater RSV levels with clinical cases. Through sequencing and phylogenetic analy-

sis, they compared RSV A and B G-gene sequences from wastewater and clinical samples to elucidate transmission patterns. A study of an active international land border found that wastewater signals peaked in Detroit (MI, USA) for the 2022-2023 RSV season approximately 5 weeks prior to the peak in Windsor (ON, Canada).⁵ The authors further found a strong positive relationship between wastewater disease concentrations and hospitalization rates in the Canada location.

Building on the success of WBE for RSV, we address a scientific gap that takes measured wastewater viral load of RSV and estimate the effective reproduction number and the number of infections in the population, advocating a modular approach to the analysis. We first estimate the trend and current level of the RSV viral load and quantify the uncertainty in the trend. These estimates become input for a Bayesian renewal model for both the infection rate and the number of infected individuals. The modular approach simplifies the analysis pipeline while maintaining scientific integrity. This approach further supports translation to other viruses by refinement of the estimate of the trend and using disease-specific models for estimated transmission and cases in the second phase of the analysis.

2 METHODS.

We estimate the weekly reproduction rate, R_t , and number of infections, I_t , from RSV RNA concentrations collected each Monday between January 16, 2023, and December 30, 2024. The sample is taken from the largest wastewater treatment plant in Houston, TX, which serves an estimated 551,150 residents. The facility's size and regular sampling ensure it provides one of the most de-

^a Department of Statistics, MS 138, Rice University, Houston, TX, USA. Tel: 01 713 348 4687; E-mail: ensor@rice.edu

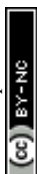
^b Department of Statistics, Rice University, Houston, TX, USA.

^c Biocomplexity Institute, University of Virginia, Charlottesville, NC, USA.

^d Houston Health Department, Houston, TX, USA.

^e Houston Health Department & Department of Statistics, Rice University, Houston, TX, USA.

^f Department of Civil & Environmental Engineering, Rice University, Houston, TX, USA.



pendable wastewater signals in the city.

The temporal relationship between the number of infections, I_t , is compared to the estimated number of RSV cases for the city during the same time period, proportionally allocated to the population served by the single treatment plant.

2.1 Wastewater Measurements.

Twenty-four-hour composite samples were collected weekly each Monday between January 16, 2023, and December 30, 2024. Laboratory analysis of the RSV RNA signal is available each Friday.⁶ Wastewater samples are collected on RSV RNA viral loads, expressed in 10^9 genome copies per day (B gc/day).^{7,8} Viral loads are constructed by multiplying observed concentrations (gc/L) by each plant's median daily influent flow (L/day) and rescaling by 10^9 . In this analysis, values below the laboratory limit of detection (LOD) were set to a uniform random number between one-half of the LOD and the LOD.

2.2 RSV Syndromic Data.

RSV healthcare encounter data are sourced from the Houston Health Department's syndromic surveillance system, HHD-ESSENCE.⁹ Health encounter data reflect city-wide counts and include routine, urgent care, emergency department, and hospitalization encounters from healthcare facilities that are on-boarded. The city-wide counts are proportionally adjusted based on the percentage of the population served by the wastewater treatment plant.

2.3 Model and Estimation.

We use a two-step modular approach to the data analysis and estimating the temporal trajectory of the reproduction number and number of infections in the community.

2.3.1 State-Space Model for WW Trend Estimate

We assume that the RSV WW RNA copies per day follow a lognormal distribution with time varying mean and variance. Based on this assumption we estimate the trend and its uncertainty through a non-linear Gaussian state-space model (SSM), by modeling the log of the observed series. The SSM framework defines two equations, namely the state equation and the observation equation. The state equation represents the trend or true dynamics of the system being studied, whereas, the observation equation models the observed series as a function of the true state and other features but most importantly the observation noise. Through these dynamics we are able to separate the uncertainty in the trend or state, and uncertainty due to sampling and measurement error. In short, our filtered estimates and their uncertainty of the trend represent our best estimate of the true RSV WW viral load at any time t , given measurements up to and including time t .

Following the approach in¹⁰, let x_t denote the trend in wastewater RNA viral load, measured on a log10 basis, at time t . The dynamics of the trend are specified by the state equation as

$$(x_t - x_{t-1}) = (x_{t-1} - x_{t-2}) + w_t, \quad w_t \sim N(0, \sigma_w^2),$$

where w_t is the noise term which follows a Normal distribution

with mean zero and variance σ_w^2 . The rationale for this specific state equation is it represents the first differences twice, and most closely mimics a nonparametric cubic spline.^{10,11} The trend in RSV viral load is nonlinear and nonstationary, this model allows for cubic local dynamics. It is important to note, that the specific model for the state or trend can be altered to reflect the temporal dynamics of any given series.

The observation equation of the state-space model is given by

$$y_t = x_t + v_t, \quad v_t \sim N(0, \sigma_v^2),$$

where y_t denotes the measured viral load for week t , and v_t represents measurement and sampling error. This error is assumed normally distributed with mean zero and variance σ_v^2 . More complexity can be encapsulated in the observation equation if necessary; the primary purpose of our parameterization is to isolate the measurement and sampling error from the state or trend.

We estimate the parameters of the SSM by maximum likelihood. Parameter estimation is an iterative process of likelihood evaluation using the Kalman filter at a given set of parameters and nonlinear optimization moving us through the parameter space. For a fixed parameterization, the Kalman filter provides efficient calculation of the filtered states, or conditional means, and their uncertainties, or conditional variances. We implement this process using the MARSS package in R. A nice feature, is that this algorithm and software can easily manage missing values in our time series. The SSM framework coupled with the Kalman Filter technology, can also manage irregularly spaced observations as seen in.^{11,12}

The output from stage one of our analysis is the estimated state, or filtered value which we denote by $\hat{x}_{t|t} = E[x_t | y_1, \dots, y_t]$, at time t throughout our time series. In other words, the estimated state at time t is the expected value of the trend (on a log10 scale) at time t , given all observations through time t . We also obtain the estimated variances of the state at time t , $\hat{P}_{t|t} = \text{Var}[x_t | y_1, \dots, y_t]$, or the conditional variance of x_t at time t .

Importantly, estimation is performed on the log10 scale and subsequently converted to the natural measurement scale (gc/L), ensuring consistency with the log-normal likelihood adopted in the measurement layer of the renewal model. You can simplify this step, by taking the natural log rather than log10, each work equally as well as we ultimately transform back to the original measurement scale.

2.3.2 Bayesian Renewal Model

Stage two of the analysis takes the estimated trend and variance of the RSV viral load, and estimates the infection dynamics and counts. We use a Bayesian renewal framework, common in wastewater epidemiology¹³⁻¹⁶. The framework incorporates both new infections and assumes a common shedding profile across all individuals in the population. The renewal equation connects past infections to the viral RNA signal.

2.4 Disease transmission dynamics

The effective reproduction number R_t is modeled through a scaled softplus transformation of an unconstrained unobserved variable



151 z_t for disease transmission, or

$$R_t = \frac{\log(1 + \exp(kz_t))}{k}, \quad k > 0.$$

152 This transformation guarantees positivity of R_t while behaving ap-199
153 proximately linear around zero and is numerically stable.

154 The unobserved variable for disease transmission, namely $\{z_t\}$,201
155 is modeled as a random walk with normally distributed noise.202
156 This parameterization provides a parsimonious yet flexible rep-203
157 resentation of week-to-week changes in disease transmission.204
158 Specifically, $z_t = z_{t-1} + \varepsilon$ where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ and $z_1 \sim N(1, \sigma_\varepsilon^2)$ 205
159 with $\sigma_\varepsilon > 0$. The parameter k controls the curvature of the link206
160 between z_t and R_t and regulates how fluctuations in the process207
161 translate into changes in transmission.

162 Further, let I_t denote the number of new infections during week209
163 t . Following the renewal formulation¹³, the expected number of210
164 cases, denoted by λ_t , is expressed as a convolution of past infec-211
165 tions with the transmission-interval distribution:

$$\lambda_t = R_t \sum_{g=1}^G w_g^{\text{inf}} I_{t-g}, \quad t \geq 2,$$

166 where w_g^{inf} are the infection transmission interval weights and G is216
167 the maximum number of weeks of history considered. The num-217
168 ber of infections at time t is assumed to follow a Poisson distribu-218
169 tion with time-specific parameter λ_t .

170 The weekly temporal weights, w_g^{inf} , for newly infected individ-220
171 uals are set by segments of a Gamma probability density func-221
172 tion (pdf) for weeks $1, \dots, G$, and rescaled to sum to one. This222
173 construction is standard in renewal-based epidemic models^{13–15}.223
174 The weights are explicitly defined in,¹⁶ and demonstrated in the224
175 code.

176 Following¹⁷, we assume a mean of 7.5 days and a standard226
177 deviation of 2.1 days for RSV for the Gamma pdf to define the227
178 infection transmission weights.

179 There is not a one-to-one relationship between wastewater229
180 RNA copies and the number of infections, rather this relation-230
181 ship is inferred through our renewal process. The magnitude of I_t 230
182 depends on our assumed initial condition. For our purposes, we231
183 set the initial number to the citywide average weekly RSV health-232
184 care encounters, scaled by the share of the city's population in the233
185 service area.¹⁸ This initial setting is considered conservative, as234
186 it does not account for individuals who did not seek medical care.

187 2.5 Linking wastewater to infections

188 We link the number of infections to RNA virus c/L through the238
189 shedding weights $\{w_d^{\text{shed}}\}_{d=0}^D$, which incorporate the timing of vi-239
190 ral shedding after infection. The expected wastewater value is

$$\pi_t = \beta \sum_{d=0}^D I_{t-d} w_d^{\text{shed}},$$

191 where $\beta > 0$ converts infections into 10^9 genome copies per day245
192 per infection.

193 Similar to the infection transmission weights, the weekly247
194 weights $\{w_d^{\text{shed}}\}_{d=0}^D$ for the shedding distribution for a given in-248

195 dividual are also modeled as a discretized Gamma distribution,
196 with maximum history of D weeks, and rescaled to sum to one.
197 Clinical evidence provided in¹⁹, leads us to assume a mean of 4.6
198 days and a standard deviation of 2.0 days of the Gamma pdf to
obtain the temporal weights for the duration of viral shedding.
Because our data are aggregated weekly, the interpretation of
lag zero differs from daily formulation. Specifically, $d = 0$ corre-
sponds to shedding occurring within the same calendar week as
infection, which is biologically plausible given that RSV shedding
can begin within the first few days of illness^{20,21}.

The wastewater RNA trend estimate, say $\hat{u}_{t|t}$, is provided by
step 1, where $\hat{u}_{t|t}$ follows a lognormal distribution, and the
 $\log(\hat{u}_{t|t})$ is normally distributed with mean $\hat{x}_{t|t}$ and variance $\hat{P}_{t|t}$.
The latter are the output from the state equation of our SSM, in
step 1. This parametrization implies that the expected value of u_t
is in fact π_t , defining our link between wastewater trend and the
infection dynamics.

For the state-space model, we performed the estimations using
maximum likelihood with BFGS optimization through the MARSS
package,²² which utilizes Kalman filtering for efficient likelihood
evaluation. The uncertainty in the parameter estimates for σ_v and
 σ_w , of the SSM is not incorporated in the Bayesian renewal model.
We estimate the parameters of the Bayesian renewal model using
`nimble`²³ via MCMC. Prior distributions for the parameters
 β , k , and σ_ε , include a log-scale prior for β , a half-normal prior
for the random-walk volatility σ_ε , and a log-normal prior on k
to regularize the curvature of the softplus link. Hyperparame-
ters are chosen to ensure weakly informative priors. The MCMC
sampler jointly estimated the unobserved values of R_t and I_t along
with key distribution parameters, producing posterior samples for
each. Multiple chains were implemented with warm-up iterations
discarded. Convergence of the MCMC algorithm was assessed using
the rank-normalized split \hat{R} . Mixing was verified through the
relative Monte Carlo standard error (relMCSE).

3 Results and Discussion

3.1 Fitted Model and Diagnostics

The SSM defining the trend requires estimation of three param-
eters, specifically the initial state x_0 , the variance of the noise
for the observation equation σ_v^2 , and the variance of the noise
for the state equation σ_w^2 . These unbiased estimated paramete-
rs for the SSM defining the trend in the RSV wastewater viral
load are $\hat{x}_0 = 0.0074$ (SE 0.6783), $\hat{\sigma}_v^2 = 0.02145$ (SE 0.0034), and
 $\hat{\sigma}_w^2 = 0.0005$ (SE 0.0002). The important variance parameters are
both small, with small standard errors, indicating the model sepa-
rates the trend from the sampling and measurement error. The
uncertainty in these parameters is not included in the Bayesian
renewal model. The estimated WW trend and its variance are
conditional on the point estimates for the variance parameters.
The accuracy of the estimates of these parameters supports this
analysis decision and does not influence the stage 2 outcomes.

Additional assessment of the SSM assumptions is performed
by examination of the standardized residuals. The standardized
residuals of the fitted SSM model do not exhibit any additional
autocorrelation over a 4 week history, based on the Ljung-Box



test (p-value = 0.33). Further, the Kolmogorov-Smirnov test for normality of the standardized residuals indicate normality is a reasonable assumption (p-value = 0.09). The SSM model does an excellent job of capturing the WW viral trend.

Table 1 summarizes posterior estimates and convergence diagnostics for the scalar parameters of the Bayesian renewal model. We observe strong convergence with rank-normalized split \hat{R} values between 1.005 and 1.018. Further the chains are well mixed as indicated by the small relative Monte Carlo standard errors (relMCSE) (≈ 0.01 – 0.02), across all parameters. The link-curvature parameter k centers around 8, consistent with a moderately sharp softplus mapping from z_t to R_t . The innovation standard deviation σ_ε is small (0.089), suggesting gradual week-to-week changes in transmission. The wastewater, the scaling factor β is 14.865 (SD 3.388 with 95% Credible Interval [9.465, 22.352]). The wastewater observation dynamic variance is provided through the SSM, in other words, the time-varying \hat{P}_t .

Table 1 Posterior means, standard deviations, 95% credible intervals, and convergence diagnostics (\hat{R} /relMCSE) for scalar parameters. The SSM-filtered viral-load inputs with week-specific variances \hat{P}_t are treated as known.

Parameter	Mean	SD	95% CI	\hat{R} /relMCSE
β	14.737	3.969	[8.591, 23.945]	1.018 / 0.021
k	8.026	4.727	[1.765, 18.600]	1.005 / 0.009
σ_ε	0.078	0.027	[0.032, 0.137]	1.017 / 0.019

3.2 Number of infections and transmission dynamics from wastewater

Figure 1 displays posterior means and 95% credible bands for the estimated number of infections (I_t). As a reference, the estimated wastewater viral load in 10^9 RNA copies/day is also included. I_t rises sharply in late 2023, drops in January 2024, remains low through mid-2024, and shows a moderate uptick in autumn 2024. The peak in the estimated number of cases I_t occurs approximately one week before the peak in the wastewater trend. Because the absolute scale of I_t is not identified from wastewater alone, we interpret I_t in relative terms and anchor its level with an initial starting point of 64 case, estimated from the healthcare encounter data.

Figure 2 depicts the time series of the effective reproduction number (R_t). For the majority of the two-year period, the reproduction number is 1, indicating no change in disease levels. The reproduction number increases in October 2023 and again in October/November 2024, corresponding to the increase of disease measured in WW. Disease transmission quickly returned to stable levels after the rise, as indicated by reproduction numbers less than one in January 2024, and then stabilized at one.

The results shown in figures 1 and 2 and were driven solely by wastewater trends. The only case information used in modeling was the initial starting value, which was set to 64 cases based on the observed RSV healthcare encounter cases for the week of January 2, 2023. Additionally, our modeling decisions for the transmission and shedding weights were based on scientific references, not estimated from our own data. For comparison, figure

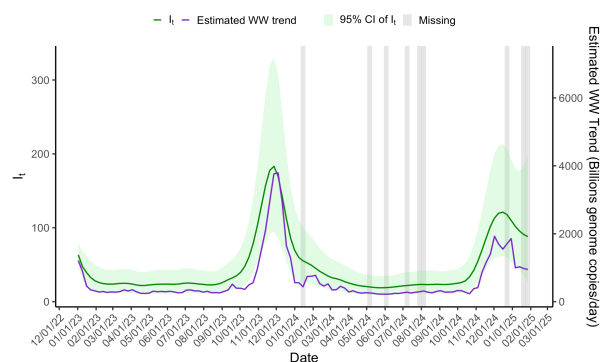


Fig. 1 Time series plot of estimated infections from wastewater, I_t , with 95% credible intervals (green). The estimated WW trend (violet) is included with scale on the right axis. Grey vertical bars indicate when WW measurements are missing.

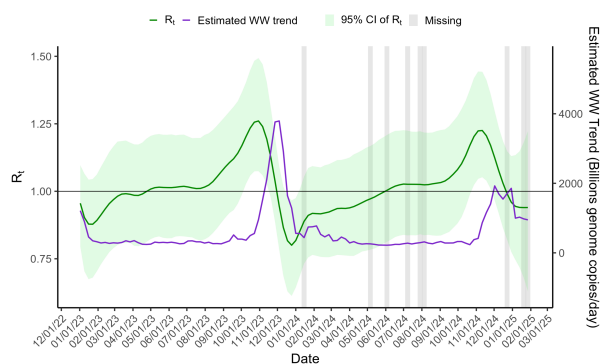


Fig. 2 Time series plot of effective reproduction number (R_t), with 95% credible intervals (green). The estimated WW trend (violet) is included, with scale on the right axis. Grey vertical bars indicate when WW measurements are missing.

shows posterior means and 95% credible bands for the estimated number of infections (I_t) and the observed time series of weekly RSV healthcare encounter cases, proportionally adjusted for the population served by the wastewater treatment plant.

The estimated number of cases closely follows the observed cases in both magnitude and pattern. The observed cases peak two weeks before the estimated cases during both periods of increased disease prevalence. During the second peak, wastewater levels did not rise as much as expected, leading to a lower case estimate, specifically I_t .

The mismatch in peak levels between the estimated cases and observed cases, could be due to several choices in our modeling framework. The first is selecting a random walk for the transmission process. A modification that includes two weeks of history instead of one may be more appropriate. Additionally, the disease transmission weights and shedding weights for the renewal equations were derived from scientific literature rather than being specifically fitted to the RSV case data for this study area. This was an intentional decision for the study. Our aim was to evaluate how well we could model the case counts and reproduction number relying solely on wastewater.



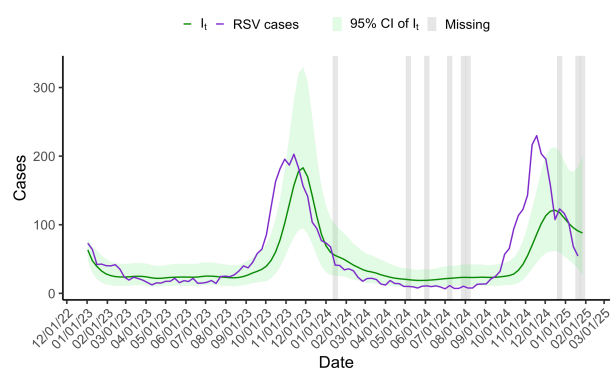


Fig. 3 Time series plot of estimated infections from wastewater, I_t , with 95% credible intervals (green). The observed health care encounter count (violet) is included for comparison. Grey vertical bars indicate when WW measurements are missing.

Conclusions

Through this study we have demonstrated a practical and scientifically valid approach to estimate the effective reproduction number and the number of infections in a population where wastewater RNA copies per liter are measured. The focus of our application is Respiratory Syncytial Virus in a large sewershed serving over 500,000 people residing in Houston, TX. The two-stage approach separates estimating the wastewater viral trend from the downstream estimate of the infection counts and dynamics. This analytic approach simplifies the analysis pipeline, bringing actionable science to real-world applications in a timely fashion.

Further, the two-stage approach allows the analyst to focus on two very different modeling challenges. By capturing the trend dynamics in wastewater, the analyst has a path forward for nowcasting and short-term forecasts of wastewater virus levels. The use of the SSM results in a clear separation of the underlying trend from the observation noise driven by both measurement and sampling uncertainty. We have specified a simple SSM model to track RSV RNA levels for our study area, and it worked well. However, the SSM technology is very flexible and can handle more complexity as needed for a specific situation. Further, the SSM technology supports missing or irregularly sampled wastewater time series. By handling these data issues early, we streamline the next step of the analysis, understanding the disease dynamics.

The stage-two Bayesian renewal model to estimate the number of individuals with the disease and the effective reproduction number can be implemented in an on-line fashion providing real-time actionable public health information. It is simpler than a Bayesian methodology that tries to address both modeling steps one and two simultaneously. Again, the input to the Bayesian renewal system is the estimated wastewater viral trend, and the uncertainty in that estimate.

An additional input to the stage-two Bayesian renewal model is the initial number of infected persons, representing the starting point of the disease. In our study, we used a proportional allocation of the infected population within Houston at the beginning of the study. This approach worked well, allowing us to

accurately capture the levels and disease dynamics of RSV. It is important to note that this starting point serves as an anchor for estimating the number of infected individuals by the Bayesian renewal model. Essentially, we define the population with this setting. Since we relied on healthcare encounter data, our estimated number at later times reflects the same population. However, this estimate is lower than the actual number of infected individuals, as it does not include those who are sick but do not seek medical care. If there is reliable scientific data on the degree of undercounting, this information could be used to refine the initial estimate of infected individuals. We also implicitly assume that the disease dynamics, for example the shedding distribution, does not change over time. If more scientific knowledge is available about the disease that information can be incorporated into the modeling. Beyond this initial point, our disease trajectory is driven entirely by the wastewater trend estimate and its uncertainty.

The fact that the estimated number of infections measured from wastewater peaks after the healthcare encounter counts, reflects the lag in observed wastewater levels. This scientific feature is worth further study to elevate the public health use of wastewater for understanding RSV spread in the community. The lag may be reflective of infected individuals in the population who are not seeking medical care. Advances in scientific knowledge about the disease and measurement dynamics can be incorporated into the stage two model to improve the public health utility of the wastewater estimated infections present in the community.

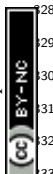
Wastewater-based epidemiology (WBE) is a proven effective tool for tracking community circulation of respiratory viruses. We contribute to the scientific gap that takes measured wastewater viral load of Respiratory Syncytial Virus (RSV) and estimate the trajectory of the effective reproduction number and the number of infections in the population. Future statistical research is warranted to understand the mismatch in the timing of peak levels in number of RSV health care encounters. Specific areas to address include the assumption that disease transmission is a random walk, and the weights selected from the scientific literature for both the shedding and transmission distributions. Another hypothesis for the timing mismatch is that clinical data captures the onset of the disease by documenting severe cases, while cases estimated through wastewater reflect the spread of RSV to the larger and potentially less vulnerable population. These are questions that require further scientific investigation to answer.

Conflicts of interest

There are no conflicts to declare.

Notes on contributor(s)

Ensor, Rice Co-PI of the Houston Wastewater Epidemiology CDC National Wastewater Surveillance System Center of Excellence, is the corresponding author. She led the research and wrote the initial draft of the manuscript. Palacio researched, developed, and implemented the statistical methodology for disease transmission, producing all results in the paper and is lead author on a deeper statistical paper on this same topic. Keller collaborated on the two-stage implementation with specific attention to uncertainty quantification. Schneider oversees the day-to-day im-



plementation of the HHD WBE system and represents the team in
national wastewater surveillance data analysis conversations. Do-
makonda serves as HHD manager for this project and all wastewa-
ter surveillance programs for HHD. Hopkins, serves as HHD PI for
this project, and ensures expert science reaches decision makers
and that the team is addressing the right questions. Stadler, Rice
Co-PI of this project, oversees all aspects of the laboratory anal-
yses for the Center, and brings the essential scientific leadership
and expertise in WBE including quantitative areas. All authors
contributed to the scientific discussion and edited the manuscript.

Data availability

Wastewater RSV levels and health care encounter counts used in
this manuscript are hosted on the Kinder Institute Urban Data
Platform (<https://www.kinderudp.org>).²⁴ R code for modeling
is available through <https://github.com/hou-wastewater-epi-org>.

Acknowledgements

The authors disclose receipt of the following financial support
for the research, authorship, and/or publication of this article:
This work was supported by the Centers for Disease Control and
Prevention (ELC-ED grant no. 6NU50CK000557-01-05 and ELC-
CORE grant no. NU50CK000557). The authors acknowledge
Houston Public Works for their contributions to the HHD WBE
system. The authors would like to acknowledge the CDC National
Wastewater Surveillance System (NWSS) scientific community.
For more information on the Houston Wastewater Epidemiology
Center of Excellence, see <https://www.hou-wastewater-epi.org>.

Notes and references

- M. B. Diamond, A. Keshaviah, A. I. Bento, O. Conroy-Ben,
E. M. Driver, K. B. Ensor, R. U. Halden, L. P. Hopkins, K. G.
Kuhn, C. L. Moe, E. C. Rouchka, T. Smith, B. S. Stevenson,
Z. Susswein, J. R. Vogel, M. K. Wolfe, L. B. Stadler and S. V.
Scarpino, Wastewater Surveillance of Pathogens Can Inform
Public Health Responses, *Nature Medicine*, **28**, 1992–1995.
- B. Hughes, D. Duong, B. J. White, K. R. Wigginton, E. M. G.
Chan, M. K. Wolfe and A. B. Boehm, *Respiratory syncytial virus
(RSV) RNA in wastewater settled solids reflects RSV clinical
positivity rates*, medRxiv preprint, 2021, <https://doi.org/10.1101/2021.12.01.21267014>, Posted December 5, 2021,
DOI:.
- A. Zulli, M. R. Varkila, J. Parsonnet, M. K. Wolfe and A. B.
Boehm, Observations of Respiratory Syncytial Virus (RSV)
Nucleic Acids in Wastewater Solids Across the United States in
the 2022–2023 Season: Relationships with RSV Infection Pos-
itivity and Hospitalization Rates, *ACS ES&T Water*, **4**, 1657–
1667.
- D. M. Allen, M. I. Reyne, P. Allingham, A. Levickas, S. H. Bell,
J. Lock, J. D. Coey, S. Carson, A. J. Lee, C. McSparron, B. F.
Nejad, J. McKenna, M. Shannon, K. Li, T. Curran, L. J. Broad-
bent, D. G. Downey, U. F. Power, H. E. Groves, J. M. McKinley,
J. W. McGrath, C. G. G. Bamford and D. F. Gilpin, Genomic
Analysis and Surveillance of Respiratory Syncytial Virus Us-

ing Wastewater-Based Epidemiology, *The Journal of Infectious
Diseases*, **230**, e895–e904.

- M. Beach, R. Corchis-Scott, Q. Geng, A. M. Podadera Gonza-
lez, O. Corchis-Scott, E. Harrop, J. Norton, A. Busch, R. A.
Faust, B. Irwin, M. Aloosh, K. K. S. Ng and R. M. McKay,
Wastewater-Based Surveillance of Respiratory Syncytial Virus
Reveals a Temporal Disconnect in Disease Trajectory across
an Active International Land Border, *Environment & Health*,
2025, **3**, 425–435.
- L. Hopkins, D. Persse, K. Caton, K. Ensor, R. Schneider, C. Mc-
Call and L. B. Stadler, Citywide Wastewater SARS-CoV-2 Lev-
els Strongly Correlated with Multiple Disease Surveillance In-
dicators and Outcomes over Three COVID-19 Waves, *Science
of The Total Environment*, **855**, 158967, DOI:.
- J. Wu, K. B. Ensor, L. Hopkins and L. B. Stadler, Assessment
and application of GeneXpert rapid testing for respiratory
viruses in school wastewater, *Environ. Sci.: Water Res. Tech-
nol.*, 2025, **11**, 64–76.
- J. Wu, M. X. Wang, T. J. Treangen, K. Ensor, L. Hopkins and
L. B. Stadler, Sensitivity, Throughput, and Cost Analysis of
Concentration Methods for Multitarget Pathogen Wastewater
Monitoring, *ACS ES&T Water*, **5**, 5531–5541.
- Houston Health Department, *Respiratory Diseases Surveillance
Report, 2024–2025 Influenza Season (MMWR Week 2025-02)*,
Houston health department technical report, 2025, DOI:.
- K. B. Ensor, J. C. Schedler, J. Palacio, R. Schneider, K. Do-
makonda, L. Hopkins and L. B. Stadler, A Nonlinear Hierar-
chical Time Series Approach to Citywide Trend Assessment
of Viruses, Hot Spot Signals, and Right-Sizing System, *Data
Science in Science*, 2025, **4**, year.
- R. H. Shumway and D. S. Stoffer, *Time Series Analysis and
Its Applications: With R Examples*, Springer Nature, 5th edn,
2025, DOI:.
- D. Ouyang, L. Chung, M. Osborn, T. W. Schacker and C. R.
Doss, Dynamic Linear Models for Wastewater-Based Epi-
demiology with Missing Values: an Application to Covid-19
Surveillance, *Data Science in Science*, 2025, **4**, 2562199.
- C. Fraser, Estimating individual and household reproduction
numbers in an emerging epidemic, *PLoS ONE*, 2007, **2**, e758.
- J. S. Huisman, J. Scire, L. Caduff, X. Fernandez-Cassi, P. Gane-
sanandamoorthy, A. Kull, A. Scheidegger, E. Stachler, A. B.
Boehm, B. Hughes, A. Knudson, A. Topol, K. R. Wigginton,
M. K. Wolfe, T. Kohn, C. Ort, T. Stadler and T. R. Ju-
lian, Wastewater-based estimation of the effective reproduc-
tive number of SARS-CoV-2, *Environmental Health Perspec-
tives*, 2022, **130**, 057011.
- D. Champredon, I. Papst and W. Yusuf, ern: An R package
to estimate the effective reproduction number using clini-
cal and wastewater surveillance data, *PLOS ONE*, 2024, **19**,
e0305550.
- J. Palacio, K. B. Ensor, S. A. Keller, R. Schneider, K. Do-
makonda, L. Hopkins and L. B. Stadler, Inferring Transmis-
sion Dynamics of Respiratory Syncytial Virus from Houston
Wastewater, *submitted to Data Science in Science*, TBA, TBA.
- M. A. Vink, M. C. J. Bootsma and J. Wallinga, Serial intervals



- 514 of respiratory infectious diseases: A systematic review and
515 analysis, *American Journal of Epidemiology*, 2014, **180**, 865–
516 875.
- 517 18 Houston Health Department, *Respiratory Diseases Surveillance*
518 *Report: 2024–2025 Influenza Season, MMWR Week 2025-*
519 *02, 2025*, [https://www.houstonhealth.org/media/11526/
520 download?inline=](https://www.houstonhealth.org/media/11526/download?inline=), Accessed 2025-09-30, DOI:.
- 521 19 M. Cevik, P. A. Piedra, V. Demicheli, C. Cantarelli, D. Manno,
522 S. Vergnano, S. Pullan and R. Harwood, Virology, transmis-
523 sion, and clinical features of respiratory syncytial virus in
524 adults, *BMJ*, 2023, **380**, e072292.
- 525 20 C. B. Hall, C. E. Long and K. C. Schnabel, Respiratory syncytial
526 virus infections in previously healthy working adults, *Clinical*
527 *Infectious Diseases*, 2001, **33**, 792–796.
- 528 21 J. P. DeVincenzo, C. M. El Saleeby and A. J. Bush, Respi-
529 ratory syncytial virus load predicts disease severity in pre-
530 viously healthy infants, *Journal of Infectious Diseases*, 2005,
531 **191**, 1861–1868.
- 532 22 E. E. Holmes, E. J. Ward and K. Wills, MARSS: Multivariate
533 autoregressive state-space models for analyzing time-series
534 data, *The R Journal*, 2012, **4**, 11–19, DOI:.
- 535 23 P. de Valpine, D. Turek, C. Paciorek, C. Anderson-Bergman,
536 D. Temple Lang and R. Bodik, Programming with models:
537 writing statistical algorithms for general model structures
538 with NIMBLE, *Journal of Computational and Graphical Statis-*
539 *tics*, 2017, **26**, 403–413.
- 540 24 L. Hopkins, K. Domakonda, K. Caton, L. Schneider, R.
541 abd Stadler and K. Ensor, *City of Houston Respiratory Syncytial*
542 *Virus Wastewater Monitoring and Health Care Encounter Data*
543 *(Version 1) [Dataset]*, 2026, <https://www.kinderudp.org>.



DAS

Wastewater RSV levels and health care encounter counts used in this manuscript are hosted on the Kinder Institute Urban Data Platform (kinderudp.org) with citation:

Hopkins, L., Domakonda, K., Caton, K., Schneider, R., Stadler, L., & Ensor, K. (2026). City of Houston Respiratory Syncytial Virus Wastewater Monitoring and Health Care Encounter Data (Version 1) [Dataset]. Rice University-Kinder Institute:

UDP. <https://doi.org/10.25612/837.VN2JPLW82ZDV>

R code for modeling is available through <https://github.com/hou-wastewater-epi-org>.

