SARS-CoV-2 RNA concentrations in primary municipal sewage sludge as a leading indicator of COVID-19 outbreak dynamics

Jordan Peccia[1,‡,*], Alessandro Zulli[1,‡], Doug E. Brackney[2,‡], Nathan D. Grubaugh[3], Edward H. Kaplan[8,4,1], Arnau Casanovas-Massana[3], Albert I. Ko[3], Amyn A. Malik[5,6], Dennis Wang[5], Mike Wang[5], Daniel M. Weinberger[3], Saad B. Omer[3,5,6,7,*]

[1]*Department of Chemical and Environmental Engineering, School of Engineering and Applied Science, Yale University, New Haven, CT, USA*

[2]*Connecticut Agricultural Experimental Station, State of Connecticut, New Haven, CT, USA*

[3]*Department of Epidemiology of Microbial Disease, School of Public Health, Yale University, New Haven, CT, USA*

[4]*School of Management, Yale University, New Haven, CT, USA*

[5]*Yale School of Medicine, New Haven, CT, USA*

[6]*Yale Institute for Global Health, New Haven, CT, USA*

[7]*Yale School of Nursing, Orange, CT, USA*

[8]*Public Health Modeling, Yale School of Public Health, New Haven, CT, USA*

[‡]Contributed equally to this work

[*]Corresponding authors: Jordan Peccia, Tel: +1 (203) 432-4385; email: Jordan.Peccia@Yale.edu; Saad Omer, Tel +1 (203) 432-3656; Saad.Omer@Yale.edu

Abstract

We report a time course of SARS-CoV-2 RNA concentrations in primary sewage sludge during the Spring COVID-19 outbreak in a northeastern U.S. metropolitan area. SARS-CoV-2 RNA was detected in all environmental samples and, when adjusted for the time lag, the virus RNA concentrations were highly correlated with the COVID-19 epidemiological curve (R^2=0.99) and local hospital admissions (R^2=0.99). SARS-CoV-2 RNA concentrations were a seven-day leading indicator ahead of compiled COVID-19 testing data and led local hospital admissions data by three days. Decisions to implement or relax public health measures and restrictions require timely information on outbreak dynamics in a community.

Introduction

The most common metric followed to track the progression of the COVID-19 epidemic within communities is derived from testing symptomatic cases and evaluating the number of positive tests over time.[1] However, tracking positive tests is a lagging indicator for the epidemic progression.[2, 3] Testing is largely prompted by symptoms, which may take up to five days to present[4], and individuals can shed virus prior to exhibiting symptoms. There is a pressing need for additional methods for early sentinel surveillance and real-time estimations of community disease burden so that public health authorities may modulate and plan epidemic responses accordingly.

SARS-CoV-2 RNA is present in the stool of COVID-19 patients[5-7] and has recently been documented in raw wastewater.[8-10] Thus, monitoring raw wastewater (sewage) within a community's collection system can potentially provide information on the prevalence and dynamics of infection for entire populations.[11] When municipal raw wastewater discharges into

treatment facilities, solids are settled and collected into a matrix called (primary) sewage sludge, which has been shown to contain a broad diversity of human viruses including commonly circulating coronavirus strains.[12] Primary sludge provides a well-mixed and concentrated sample that may be advantageous for monitoring SARS-CoV-2. As viral shedding can occur before cases are detected, we hypothesize that the time course of SARS-CoV-2 RNA concentrations in primary sewage sludge is a leading indicator of outbreak dynamics within a community served by the treatment plant.

Results

During the COVID-19 outbreak from March 19, 2020 to May 1, 2020 in the New Haven, Connecticut (CT), USA metropolitan area **(Figure 1A)**, daily primary sludge samples were collected from the wastewater treatment facility which serves approximately 200,000 residents. SARS-CoV-2 viral RNA concentrations were quantitatively compared with local hospital admission data and community COVID-19 compiled testing data. SARS-CoV-2 viral RNA was detectable in all samples tested and ranged from 1.7×10^3 virus RNA copies mL^{-1} to 4.6×10^5 virus RNA copies mL^{-1}. The lower concentration in this range corresponds to a qRT-PCR cycle threshold (CT) value of 38.75 and can be considered a detection threshold for this method and sludge matrix. Overall, 96.5% of all CT values were less than 38 and values between 38 and 40 were reported as positive if detection occurred with virus nucleocapsid N1 and N2 primer sets and both replicates. Replicated samples demonstrated similar SARS-CoV-2 RNA concentration values and comparisons between SARS-CoV-2 primer values with the human ribonuclease P (RP) gene primer values indicate that dynamic temporal concentration changes in SARS-CoV-2 RNA were from actual change in virus concentration **(Figure 1B)**. Concentration comparisons

between replicates produced an average slope of 0.98 ± 0.17 st. dev., while replicates comparing N and RP primer values result in a flat line, with average slope of 0.05 ± 0.05 st. dev. (**Figure 1B)**. Overall, the SARS-CoV-2 RNA concentration time series results consistently reveal a curve that increases over times, peaks, and then decreases to May 1 (**Figure 1C**).

New Haven COVID-19 epidemic suggest that these data may provide useful epidemiological insights **(Figure 1C)**. SARS-CoV-2 RNA sludge concentrations were quantitatively compared with data that are commonly used to track the community progression of COVID-19 including hospital admissions **(Figure 2A,B)** and COVID-19 compiled testing data for the four municipalities (New Haven, East Haven, Hamden, and Woodbridge, CT) served by the ESWPAF **(Figure 2C)**. All three measures traced the initial wave of the SARS-CoV-2 outbreak in the New Haven metropolitan area. Applying Locally Weighted Scatterplot Smoothing (LOWESS) to the data and rescaling enables comparison **(Figure 2A,B)**. The virus RNA time course peaked 3 days earlier than hospital admissions (April 9 versus April 12) and a cross correlation analysis revealed a correlation coefficient (R=0.996) between smoothed RNA and hospital data when the latter was shifted 3 days forward.

Comparing the LOWESS-smoothed daily virus RNA and COVID-19 testing data reveals that the peak in COVID-19 cases (April 16) occurred 7 days following the peak in SARS-CoV-2 RNA concentration (April 9) **(Figure 2C)**. A cross correlation analysis resulted in a maximum correlation coefficient (R=0.994) when the COVID-19 testing data is adjusted 7 days forward. Regressing the adjusted COVID-19 testing data and SARS-CoV-2 virus RNA concentrations yield a value of 1,305 virus RNA copies/mL per new COVID-19 case report as the outbreak ascended (**Figure 2D**), and 1,240 virus RNA copies/mL per new COVID-19 case report as the outbreak descended (**Figure 2E**).

Discussion

We produced a SARS-CoV-2 RNA concentration time course in primary sewage sludge during a COVID-19 outbreak in the New Haven, CT metropolitan area. Approximately 200,000 people are served by the treatment facility and COVID-19 total documented cases by testing rose from less than 29 to 2,609 during the March 19 to May 1, 2020 surveillance period. Our results demonstrate: (1) the utility of SARS-CoV-2 primary sludge monitoring to accurately track outbreaks in a community and (2) primary sludge SARS-CoV-2 RNA concentrations can be a leading indicator over other commonly used epidemiology approaches including summarized COVID-19 test results and hospital admissions.

This study uniquely utilized primary sewage sludge instead of raw wastewater for virus RNA measurements. Sewage sludge is comprised of solids that are removed during primary sedimentation steps and typically gravity thickened. As a result, primary sludge has a greater solids content (2 to 5%) than raw wastewater (0.01 to 0.05%). Due to the elevated solids content and the high case load observed during the outbreak (~1,200 per 100,000 population), the concentrations of SARS-CoV-2 RNA reported here ranged from two to three orders of magnitude greater than raw wastewater SARS-CoV-2 values previously reported[8, 10, 13]. Monitoring primary sludge is broadly applicable. Wastewater treatment plants with primary and secondary treatment are standard in many regions of the world, and treatment facilities are rapidly expanding in urban areas of lower and middle-income countries.[14] Within the US, approximately 16,000 treatment plants serve more than 250,000,000 people.[15]

SARS-CoV-2 RNA concentrations in sewage sludge were a leading indicator of community outbreak dynamics over hospitalization and compiled COVID-19 testing data. SARS-CoV-2 RNA concentrations led hospital admissions by 3 days and COVID-19 cases by 7 days. Hospital admissions to Yale New Haven Hospital from the four towns served by the wastewater treatment facility both rose and fell more slowly than the observed RNA concentrations. The rates of SARS-CoV-2 RNA increases and decreases were similar to the COVID-19 testing data, and reflect the very concentrated outbreak that took place within New Haven and its neighboring towns. The strong correlation between the virus RNA and COVID-19 case data provides an approach to independently evaluate the local SARS-CoV-2 testing rate and to estimate the number of new cases. While the results here demonstrate the potential utility of such an approach, we note that these relationships could be treatment plant specific, as primary sludge process trains are not uniform, and that compiled COVID-19 testing is prompted by symptoms, thus underestimating the true number of cases.[16]

In conclusion, the SARS-CoV-2 RNA concentrations in primary sewage sludge were tracked throughout a COVID-19 epidemic and compared with traditional outbreak epidemiological indicators. SARS-CoV-2 RNA concentration in primary sludge closely followed the epidemiology curves established by compiled COVID-19 testing data and hospital admissions, but was a leading indicator by seven and three days, respectively. Our study could have substantial policy implications. Jurisdictions can use primary sludge SARS-CoV-2 concentrations to preempt community outbreak dynamics or provide an additional basis for easing restrictions, especially when there are limitations in clinical testing. Raw wastewater and sludge-based surveillance is particularly useful for low and middle-income countries where clinical testing capacity is limited.

Methods

Sample collection. Primary sewage sludge was collected from the East Shore Water Pollution Abatement Facility (ESWPAF) in New Haven, CT, USA. Samples were taken daily from March 19 to May 1, 2020 between 8 and 10 am EDT and stored at -80^{o}C prior to analysis. The first sampling dates were prior to widespread individual testing in the region, and prior to the start of stay at home restrictions implemented throughout the State of Connecticut, USA **(Figure 1A)**. From the sampling start and end dates, cities served by the ESWPAF experienced greater than 85 times increase in confirmed COVID-19 cases (by testing) from less than 29 cases to 2,609.[17] The plant serves an estimated population of 200,000 people with average treated flows of 1.75 m^3/s. Sludge collected from ESWPAF is primary sludge, sampled at the outlet of a gravity thickener, ranging in solids content from 2.6% to 5%.

Viral RNA quantitative testing. To quantify SARS-CoV-2 virus RNA concentrations in primary sludge, 2.5 mL of well mixed sludge were added directly to a commercial kit optimized for isolation of total RNA from soil (RNeasey PowerSoil Total RNA kit, Qiagen). Isolated RNA pellets were dissolved in 50 µL of ribonuclease free water and total RNA measured by spectrophotometry (NanoDrop, ThermoFisher Scientific). SARS-CoV-2 RNA was quantified by one-step quantitative reverse transcriptase-polymerase chain reaction (qRT-PCR) using the U.S. Center for Disease Control N1 and N2 primers sets.[18] For control, analysis was also conducted for the human Ribonuclease P (RP) gene. Samples were analyzed using the Bio-Rad iTaq Universal Probes One-Step kit in 20 µL reactions run at 50°C for 10 min, 95°C for 1 min

followed by 40 cycles of 95°C for 10 s and 60°C for 30 s per the manufacturer's recommendations. SARS-CoV-2 RNA concentrations were determined using a standard curve as previously described[18] and presented as virus RNA copies. The SARS-CoV-2 concentration results were normalized to the total RNA extracted to reduce day to day variations in treatment plant flow, sludge solids content, and RNA extraction efficiency. All samples were diluted 5x for use as template to ensure the removal of inhibition. Performing qRT-PCR on undiluted and 5x diluted samples typically resulted in the expected 5x decrease in concentration. Sewage sludge from 2018 was used as a control and no SARS-CoV-2 detection was observed from either N1 or N2 primers. These control sludges were consistently positive for the human RP gene.

Epidemiological Analysis. Daily COVID-19 admissions to the Yale New Haven Hospital were compiled from hospital records and confirmed by laboratory testing. Numbers of laboratory-confirmed COVID-19 cases in the towns served by the ESWPAF (New Haven, East Haven, Hamden, and Woodbridge CT) were compiled from daily reports published by the CT Department of Public Health.[17] SARS-CoV-2 RNA, COVID-19 cases, and hospital admissions time series were smoothed using Locally Weighted Scatterplot Smoothing (LOWESS) to better visualize the time trend in the data.[19]

Acknowledgements

We wish to thank the East Shore Water Pollution Abatement Facility, New Haven, CT, USA for primary sludge sampling assistance. AZ is supported by the Yale Institute of Global Health and NDG is supported by a gift from the Huffman Family Donor Advised Fund.

References

1. Dong, E.; Du, H.; Gardner, L., An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* **2020,** *20* (5), 533-534.
2. Mizumoto, K.; Kagaya, K.; Zarebski, A.; Chowell, G., Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Euro Surveill* **2020,** *25* (10), 2000180.
3. García-Basteiro, A. L.; Chaccour, C.; Guinovart, C.; Llupià, A.; Brew, J.; Trilla, A.; Plasencia, A., Monitoring the COVID-19 epidemic in the context of widespread local transmission. *Lancet Respir Med* **2020,** *8* (5), 440-442.
4. He, X.; Lau, E. H. Y.; Wu, P.; Deng, X.; Wang, J.; Hao, X.; Lau, Y. C.; Wong, J. Y.; Guan, Y.; Tan, X.; Mo, X.; Chen, Y.; Liao, B.; Chen, W.; Hu, F.; Zhang, Q.; Zhong, M.; Wu, Y.; Zhao, L.; Zhang, F.; Cowling, B. J.; Li, F.; Leung, G. M., Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med* **2020,** *26* (5), 672-675.
5. Wang, W.; Xu, Y.; Gao, R.; Lu, R.; Han, K.; Wu, G.; Tan, W., Detection of SARS-CoV-2 in different types of clinical specimens. *JAMA* **2020**, *323* (18), 1843-1844.
6. Zhang, J.; Wang, S.; Xue, Y., Fecal specimen diagnosis 2019 novel coronavirus-infected pneumonia. *J Med Virol* **2020**, doi: 10.1002/jmv.25742.
7. Xu, Y.; Li, X.; Zhu, B.; Liang, H.; Fang, C.; Gong, Y.; Guo, Q.; Sun, X.; Zhao, D.; Shen, J.; Zhang, H.; Liu, H.; Xia, H.; Tang, J.; Zhang, K.; Gong, S., Characteristics of pediatric SARS-CoV-2 infection and potential evidence for persistent fecal viral shedding. *Nat Med* **2020,** *26* (4), 502-505.
8. Ahmed, W.; Angel, N.; Edson, J.; Bibby, K.; Bivins, A.; O'Brien, J. W.; Choi, P. M.; Kitajima, M.; Simpson, S. L.; Li, J.; Tscharke, B.; Verhagen, R.; Smith, W. J. M.; Zaugg, J.; Dierens, L.; Hugenholtz, P.; Thomas, K. V.; Mueller, J. F., First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: A proof of concept for the wastewater surveillance of COVID-19 in the community. *Science of The Total Environment* **2020,** *728*, 138764.
9. Medema, G.; Heijnen, L.; Elsinga, G.; Italiaander, R.; Brouwer, A., Presence of SARS-Coronavirus-2 in sewage. *medRxiv* **2020**, 2020.03.29.20045880.
10. Wu, F.; Xiao, A.; Zhang, J.; Gu, X.; Lee, W. L.; Kauffman, K.; Hanage, W.; Matus, M.; Ghaeli, N.; Endo, N.; Duvallet, C.; Moniz, K.; Erickson, T.; Chai, P.; Thompson, J.; Alm, E., SARS-CoV-2 titers in wastewater are higher than expected from clinically confirmed cases. *medRxiv* **2020**, 2020.04.05.20051540.
11. Pöyry, T.; Stenvik, M.; Hovi, T., Viruses in sewage waters during and after a poliomyelitis outbreak and subsequent nationwide oral poliovirus vaccination campaign in Finland. *Appl Environ Microbiol* **1988,** *54* (2), 371.
12. Bibby, K.; Peccia, J., Identification of viral pathogen diversity in sewage sludge by metagenome analysis. *Environ Sci Technol* **2013,** *47* (4), 1945-1951.

13. Wurtzer, S.; Marechal, V.; Mouchel, J.-M.; Maday, Y.; Teyssou, R.; Richard, E.; Almayrac, J. L.; Moulin, L., Evaluation of lockdown impact on SARS-CoV-2 dynamics through viral genome quantification in Paris wastewaters. *medRxiv* **2020**, 2020.04.12.20062679.
14. Zhang, Q. H.; Yang, W. N.; Ngo, H. H.; Guo, W. S.; Jin, P. K.; Dzakpasu, M.; Yang, S. J.; Wang, Q.; Wang, X. C.; Ao, D., Current status of urban wastewater treatment plants in China. *Environ Int* **2016,** *92-93*, 11-22.
15. U.S. Cybersecurity and Cyber Intrastructure. Water and Wastewater Systems Sector: Critical Infrastructure Sectors. Washington, D.C.: U.S. Department of Homeland Security. 2016-01-08. 2018.
16. Li, R.; Pei, S.; Chen, B.; Song, Y.; Zhang, T.; Yang, W.; Shaman, J., Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **2020,** *368* (6490), 489.
17. State of Connecticut, CT Data: Confirmed COVID-19 Cases. Connecticut Department of Health, https://data.ct.gov/Health-and-Human-Services/COVID-19-confirmed-cases-by-town-/28fr-iqnx Access, 5/3/2020. **2020**.
18. Vogels, C. B. F.; Brito, A. F.; Wyllie, A. L.; Fauver, J. R.; Ott, I. M.; Kalinich, C. C.; Petrone, M. E.; Casanovas-Massana, A.; Muenker, M. C.; Moore, A. J.; Klein, J.; Lu, P.; Lu-Culligan, A.; Jiang, X.; Kim, D. J.; Kudo, E.; Mao, T.; Moriyama, M.; Oh, J. E.; Park, A.; Silva, J.; Song, E.; Takehashi, T.; Taura, M.; Tokuyama, M.; Venkataraman, A.; Weizman, O.-E.; Wong, P.; Yang, Y.; Cheemarla, N. R.; White, E.; Lapidus, S.; Earnest, R.; Geng, B.; Vijayakumar, P.; Odio, C.; Fournier, J.; Bermejo, S.; Farhadian, S.; Dela Cruz, C.; Iwasaki, A.; Ko, A. I.; Landry, M.-L.; Foxman, E. F.; Grubaugh, N. D., Analytical sensitivity and efficiency comparisons of SARS-COV-2 qRT-PCR primer-probe sets. *medRxiv* **2020**, 2020.03.30.20048108.
19. Cleveland, W. S., Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* **1979,** *74* (368), 829-836.

Figures

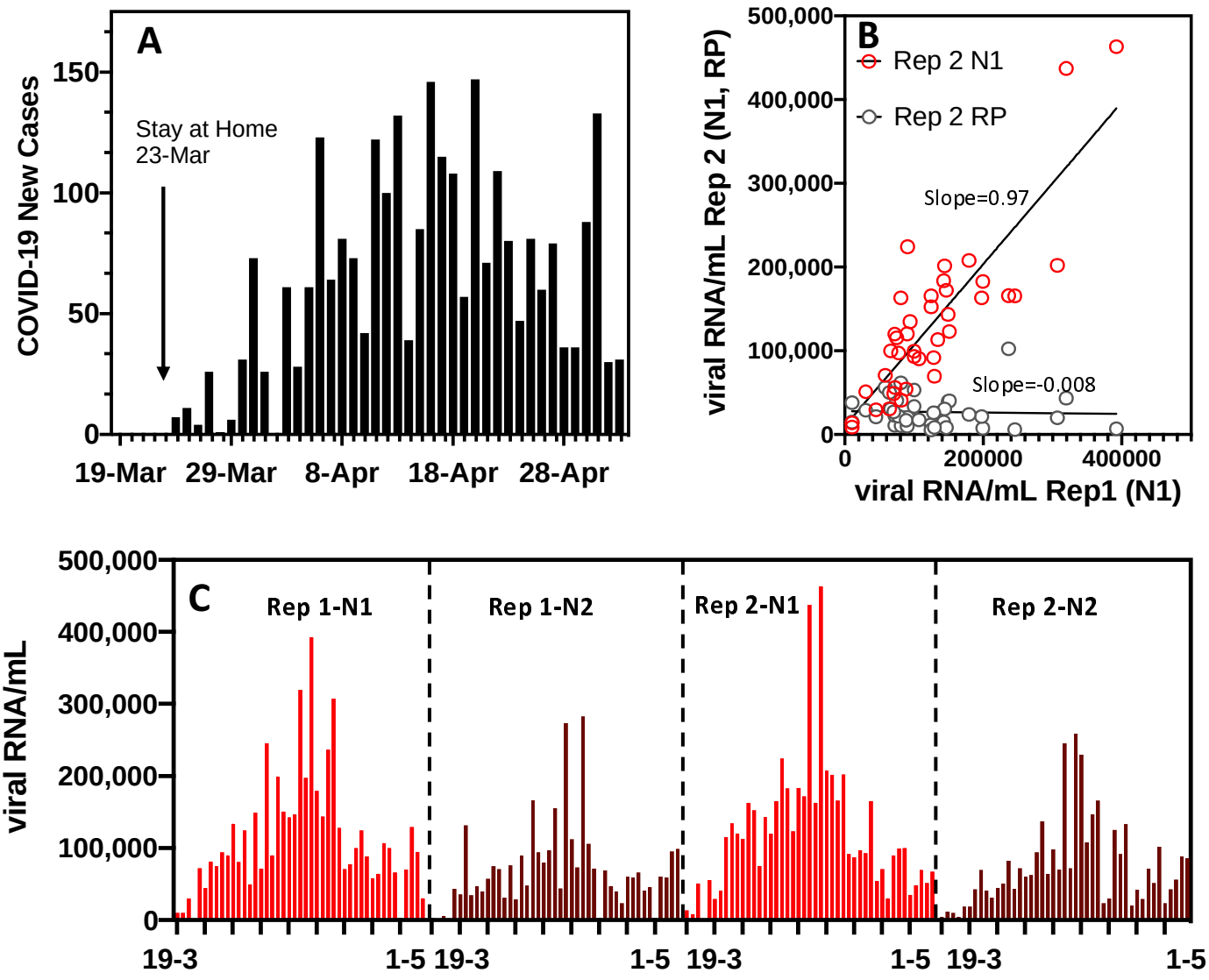

Figure 1. (A) Epidemiology curve for COVID-19 new cases (based on testing) from the cities in the wastewater catchment area (New Haven, East Haven, and Hamden, CT) served by the East Shore Water Pollution Abatement Facility (ESWPAF)[1]; **(B)** Example comparison of SARS-CoV-2 virus RNA replicates (N1 primer set) and comparison between N1 primer concentrations versus human RP concentrations; **(C)** SARS-CoV-2 virus RNA concentration time course for all replicates and primers considered.

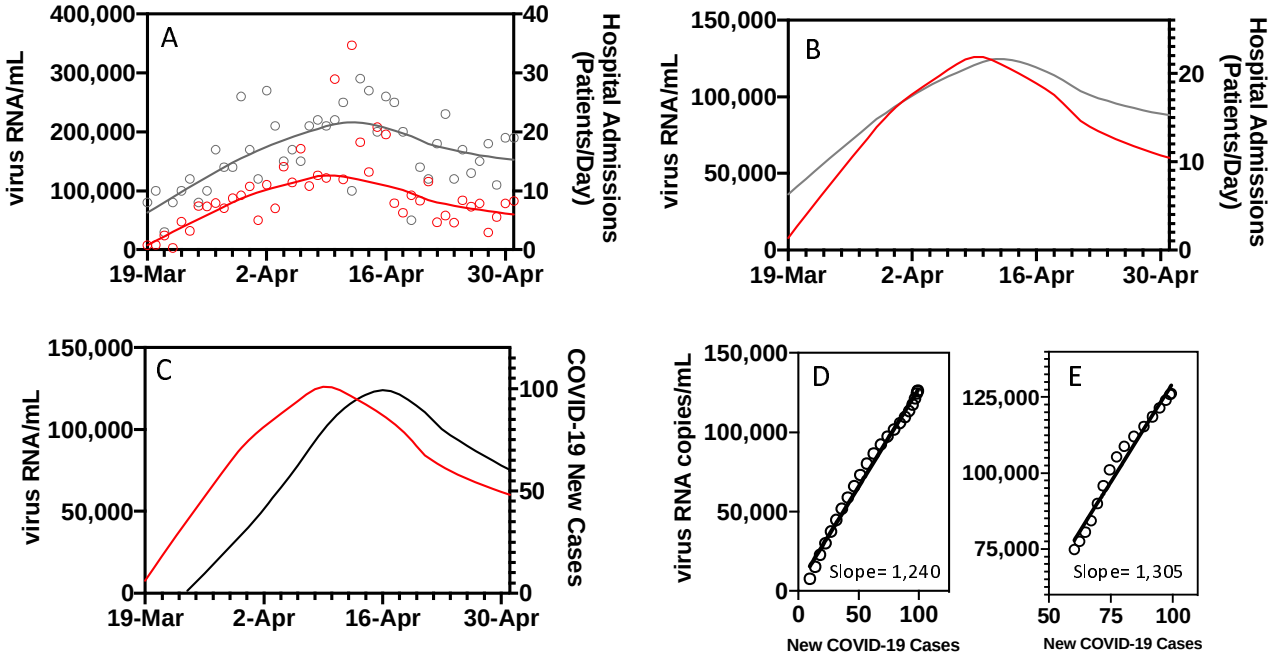

Figure 2. (A) Average sludge SARS-CoV-2 RNA concentration time course data (o) and average hospital admissions data (o) with LOWESS smoothing; (**B**) rescaled smoothed SARS-CoV-2 virus RNA concentrations (**---**) and hospital admissions (**---**); **(C)** smoothed sludge SARS-CoV-2 virus RNA concentration (**---**) with smoothed COVID-19 epidemiology curve (**---**); (**D**) regression between smoothed virus RNA and new COVID-19 cases (ascending), slope=1,240 virus RNA copies/new case, R^2=0.99; (**E**) regression between smoothed virus RNA and new COVID-19 cases (descending), slope=1,305 virus RNA copies/new case, R^2=0.97.