City of Houston
Wastewater Epidemiology for SARS-CoV-2
Houston Wastewater Epidemiology (HWE)
Best Practices

Authors: HWE Executive Team

Loren Hopkins, PhD:
City of Houston
Team Lead/PI

Kathy Ensor, PhD:
Rice University Analytics
Team Lead/PI

Lauren B. Stadler, PhD:
Rice University Wastewater
Team Lead/PI

Acknowledgements

The Houston Wastewater Epidemiology Executive Team would like to thank the following professionals who contributed their time and expertise to the development of these best practices.

HWE Core Team Members:

City of Houston Team Members: C. Hundley, MPH; K. Domakonda; K. Sheth, MPH; L. Mojica, MS; R. Schneider, MStat

City Public Works: A. Ali; C. LaBreche, PE; C. Lepadatu; C. Living; C. Mallillin; D. Chee; D. Nguyen; J. Conatto; J. Farrell; J. Gilleland; J. Moor; J. Nguyen; J. Tian; K. Nguyen; M. Siddiqui; N. Rahan; P. Munive; P. Shitabule; P. Zappi, PE; R. Hernandez; R. Mills, PE; W. Samarneh, PE; Y. Rodriguez

City Lab: A. Lawal, PhD; J. Myers; P. Brown; R. Penn, MS

Rice University Wastewater Team Members: C. McCall, PhD.; E. Lou; K. Palmer; L. Bauhs; M. Wolken; P. Kalvapalle; R. Carlson-Stadler; W. Rich

Rice University Analytics Team Members: F. El-Dahdah, PhD; T. J. Treangen, PhD; Rice's Spatial Studies Lab staff; Rice statistics students, faculty, and staff

Table of Figures

Table of Contents

Overview

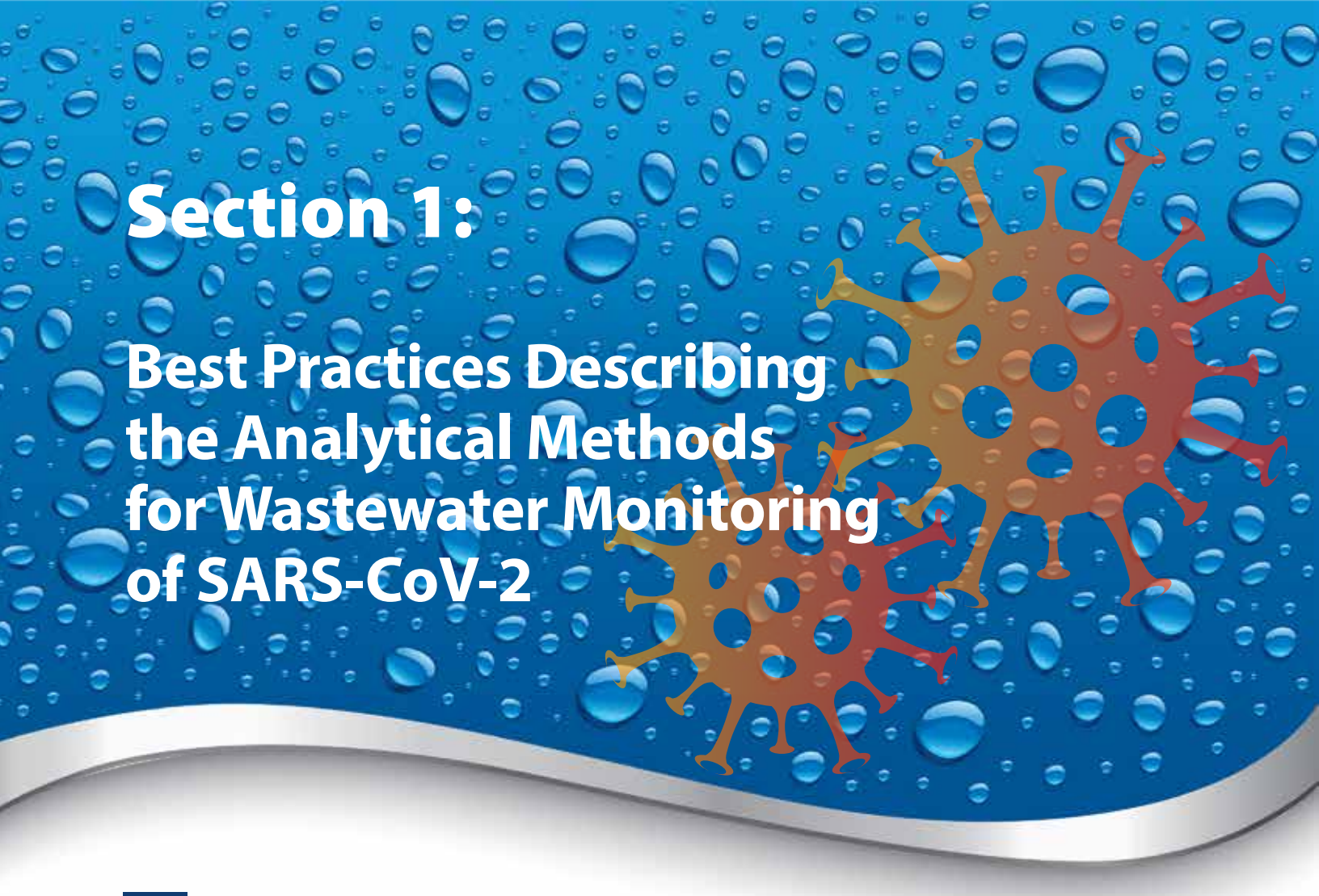
Individuals infected with COVID-19, the disease caused by the SARS-CoV-2 virus, shed viral particles in their feces, making wastewater a valuable source of information about disease prevalence in a given community. Since May 2020, the Houston Health Department (the HHD) has led a coalition of municipal and academic partners to develop a SARS-CoV-2 wastewater surveillance and reporting program for the city of Houston, Texas. The HHD and its collaborators analyze wastewater collected from wastewater treatment plants and other locations throughout the city to measure the amount of SARS-CoV-2 viral fragments in community-wide samples and determine whether levels of the virus in a particular community are increasing, decreasing, or stable. Wastewater data is analyzed using a statistical system that generates relevant public health information, which is then communicated to the appropriate government and health authorities.

The HHD acknowledges and is grateful for the scientific leadership and dedication of its partners—the Stadler, Ensor, and Spatial Studies labs at Rice University (the Rice Labs); the Taylor Lab at Baylor College of Medicine; and Houston Water (a division of the Houston Public Works Department)—which have worked closely with the HHD to establish Houston's wastewater surveillance system. The current day-to-day operations of the system are managed by the HHD, Houston Water, and the Rice Labs.

Using Houston's SARS-CoV-2 wastewater surveillance system as a case study, this best practices document is designed to help municipalities develop and implement their own wastewater epidemiology systems. The document is organized into three parts: (1) establishing and implementing a system and protocols for wastewater sampling; (2) establishing and implementing a statistical system for assessing data from wastewater samples for pertinent public health information; and (3) developing and implementing a strategy for communicating information generated by the wastewater monitoring and statistical assessment systems.

Section 1:

Best Practices Describing the Analytical Methods for Wastewater Monitoring of SARS-CoV-2



1.1. Introduction

Wastewater-based epidemiology (WBE) is a comprehensive and resource-efficient method for real-time monitoring of community infection dynamics through the analysis of pooled wastewater samples (Gonzalez et al. 2020). During the COVID-19 pandemic, WBE has exploded as a time- and cost-effective means to monitor community infection burden and track the progression of the disease in communities ranging from college dormitories (Betancourt et al. 2021) to major metropolises (Gonzalez et al. 2020; Stadler et al. 2020).

High-throughput, sensitive, and reproducible methods for quantifying viral RNA in wastewater samples are critical to a successful WBE system. This section describes our methods for concentrating SARS-CoV-2 in wastewater samples, RNA extraction, and SARS-CoV-2 quantification using reverse transcription droplet digital polymerase chain reaction (RT-ddPCR).

1.2. Aim and Scope

Section 1 of this best practices document provides recommendations for implementing and validating a city-wide wastewater monitoring system for SARS-CoV-2. We outline the quality control and assessment protocols required for sample collection through RNA quantification to establish a successful, high-throughput WBE system. The methodologies discussed here can also be used to monitor specific variants of SARS-CoV-2, as well as other pathogenic viruses, such as poliovirus, rotavirus, hepatitis A virus, hepatitis E virus, norovirus, enterovirus, and adenoviruses (LaTurner et al. 2021).

1.3. Wastewater Sample Collection

1.3.1. Autosampler Selection and Installation

Autosamplers that collect 24-hour time-weighted composite samples are used to collect samples of raw (untreated) wastewater. The type of sampler used depends on the availability of an alternating current power outlet and on whether the sampler can be secured to the wastewater source with a chain. The HACH AS950 refrigerated sampler and the Teledyne ISCO Avalanche sampler can connect to a power outlet (see Figure 1). The HACH AS950 portable sampler and the regular ISCO water sampler cannot be connected to a power source and require a battery (see Figure 2). The HACH AS950 refrigerated sampler and the Avalanche sampler are preferred because non-refrigerated samplers require an employee to surround the sample collection container with ice before scheduled sample collection begins.

Figure 1. HACH AS950 all-weather, refrigerated automatic water sampler installed at nursing home

Figure 2. HACH AS950 portable automatic water sampler installed at elementary school

The first step in the sampler installation process is to program the sampler to take the desired number of samples at a set frequency and volume. Use a tape measure to measure the length of tubing required to reach from the sampler to the desired wastewater stream. Note the length, which will be used for subsequent tubes as sampling continues. Program the length and width of the tubing into the sampler before each use. Install a strainer at one end of the tubing to prevent solid debris from obstructing it. Once the strainer is installed and the other end of tubing is connected to the sampler, lower the tubing into the wastewater stream until it is completely submerged. Failure to completely submerge the strainer may result in the sampler failing to collect all scheduled samples.

To prevent equipment from being stolen, samplers that are installed in areas the public has access to must be chained to a stationary object (e.g., a fence, pole, tree, etc.) with a lock. Samplers located in lift stations do not have to be chained since these locations are not accessible to the public. If it is not possible to keep a manhole sampler permanently chained, the sampler can either be: (1) installed on a bracket inside the manhole or (2) chained to a nearby object, and then temporarily unchained and moved near the manhole for scheduled collections. While collecting samples, place a cone near a temporarily moved sampler and cover any exposed tubing with a floor cover or mat to prevent pedestrians from tripping on it.

Only the HACH AS950 portable sampler and the ISCO sampler can be placed inside a manhole. Two people are required to install samplers inside manholes and to collect wastewater samples from samplers mounted inside manholes. For this reason, it is best to avoid installing samplers in manholes if possible. However, if necessary, the following steps will help guide sampler installation (and subsequent removal) in a manhole:

Use a j-hook to open the manhole cover. After removing the cover, allow the manhole air out for a few minutes (if it has not been opened recently). Once opened, a two-person team should attach the sampler to the manhole lid using the four arms of the bracket. Use a flashlight to ensure that the strainer end of the tubing is submerged in the desired wastewater stream before lowering the sampler into the manhole and replacing the cover. The four arms of the bracket should contact the inner lip of the manhole when installed. If a manhole is located on or near a street, use trucks and cones to navigate traffic around the manhole while installing a sampler or collecting samples.

To remove a sampler installed in a manhole, use a j-hook to open the manhole cover. After removing the cover, allow the manhole to air out for a few minutes (if it has not been opened recently). Working slowly, a two-person team should remove the sampler while holding onto the four arms of the bracket.

1.3.2. Sample Collection and Delivery

Always wear thick gloves when collecting samples. Once a sampler has been accessed, open the sampler and pour the wastewater into the required bottles. Dump any leftover wastewater and ice into the nearest sanitary sewer. Schedule the sampler to start at the next desired day. Lock the sampler. Wipe down sample bottles with disinfectant wipes and place them in an ice chest for delivery.

Once all scheduled samples are collected for the day, they must be delivered promptly to the required laboratories. All samples must be kept refrigerated and delivered to the laboratories on ice. A chain of custody sheet should be reviewed and signed by each relevant party at every point of transfer to ensure that all samples are properly accounted for.

1.3.3. Scheduling

Once a sample collection and delivery schedule is established, a lab schedule should be set up to track the tasks required for sample processing. The wet lab workflow for routine wastewater monitoring can be divided into five steps: (1) inventory and preparation, (2) receiving and aliquoting samples, (3) concentration, (4) RNA extraction, and (5) RT-ddPCR quantification. Generally, when receiving 8-10 sample sets per week, it takes 2-3 days (approximately 15-20 hours of wet lab labor) to process each sample set and review results.

It is not uncommon for unexpected issues (e.g., weather, autosampler issues, power outages, etc.) to cause delays in the scheduled delivery and processing of samples. Lab teams should be aware of and prepare for potential interruptions in their work schedule(s). For example, some project tasks (such as data analysis) can be performed outside of the lab if a remote desktop is enabled on the computer that controls the RT-ddPCR instrument.

1.4. Laboratory and Personnel Recommendations

1.4.1. Laboratory Safet

1.4.1.1. Engineering Controls

Because the risks associated with exposure to wastewater containing SARS-CoV-2 particles are not yet fully understood and the process for concentrating wastewater involves generating bioaerosols, we follow Centers for Disease Control and Prevention (CDC) recommendations and perform this work in a Biosafety Level 2 (BSL-2) facility with unidirectional airflow and Biosafety Level 3 precautions, including the use of respiratory protection (see Section 1.4.1.3) and a designated area for donning and doffing personal protective equipment (PPE).

Interim laboratory biosafety guidelines for the handling and processing of COVID-19 specimens can be found on the CDC's website.

All lab spaces and equipment should be cleaned and disinfected with 70% ethanol and 10% bleach before and after each use. Separate preparation areas and equipment should be designated for each step in the testing process to limit cross-contamination between lab spaces and personnel.

One room in the lab should be designated as the "concentration room." The air pressure in this room should be kept at a lower level than the rest of the lab to prevent wastewater aerosols or spills from contaminating surrounding areas. All steps involving raw wastewater should be carried out inside the designated concentration room, using biosafety cabinets (BSCs). After raw samples are concentrated and lysis buffer has been added to the concentrate, the virus in the samples is considered inactive and can be handled outside of a BSC and without respiratory protection.

1.4.1.2. Training

Personnel should be trained in general laboratory safety, biosafety, bloodborne pathogen safety, and lab-specific safety. Laboratory-specific safety training should include best-practices for BSL-2 and higher facilities, including PPE usage, waste management practices, disinfection processes, and BSC training. Team members should observe all relevant protocols in depth at least three times prior to performing tasks themselves to ensure understanding of testing procedures and safety guidelines. Each team member can then perform the protocols—including donning and doffing appropriate PPE—while under observation by a trained expert as many times as is deemed necessary to ensure proficiency. Safety data sheets for all reagents and a first aid kit should be kept up to date and easily accessible.

1.4.1.3. Personal Protective Equipment

To minimize aerosol exposure, disposable or autoclavable lab coats, gloves, respiratory protection (particularly when concentrating virus in wastewater samples), and eye protection should be worn when handling raw samples. Gloves, eye protection, and regular lab coats should be worn when handling samples after concentration. Multiple lab coats should be provided to each team member and designated for use in specific lab rooms to avoid cross-contamination between personnel and samples.

1.4.2. Facilities and Equipment

1.4.2.1. Instrumentation

The major instruments required for high-throughput wastewater monitoring include barcode/label printers, barcode scanners, small and large centrifuges, vacuum pumps, vacuum manifolds, bead beaters, an automated RNA extraction system (e.g., the PerkinElmer chemagic 360, or chemagic 360, a 96-well extraction instrument), a liquid handling robot for PCR plate preparation (e.g., JANUS G3 Automation Workstation, or JANUS robot), quantitative reverse transcription PCR or RT-ddPCR instruments and associated computers, refrigerators (4°C), freezers (-20°C and -80°C), pipettes, repeater pipettes, and multichannel pipettes. Most instruments require regular maintenance and consumables for use.

1.4.2.2. Consumables

In addition to general lab needs, certain instrument-specific consumables must be purchased on a regular basis to maintain a reliable, high-throughput wastewater testing system. Regular check-ins with technicians and expense-tracking logs are necessary to ensure that inventory needs and orders are kept up to date. Once a testing system is established, weekly consumable needs can be estimated based on sample quantities, enabling bulk purchases at regular intervals. When estimating procurement and delivery times, potential supply chain issues should be considered. The lab schedule should also take into account the fact that some consumables require extra preparation before they are ready for use (e.g., barcode labels for centrifuge tubes and bead tubes need to be printed and applied; glass beads and lysis buffer need to be pre-loaded into bead tubes; primers and probes need to be aliquoted for usage and storage; mastermix reagents need to be combined for extraction and quantification; and other reagents need to be aliquoted into smaller tubes to avoid contamination of larger stocks).

1.4.2.3. Data Management

Samples should be analyzed in replicate for precision and to assess variability. Once sampling locations are known, abbreviations should be created and assigned to each location for use in a laboratory information management system (LIMS) and analysis pipelines. We recommend including the date of sample collection in each sample abbreviation (e.g., 0923A1 and 0923A2). These abbreviations are used to track samples and note any issues as they move through each step of the testing process. Once samples are aliquoted into pre-labeled centrifuge tubes, the abbreviations can be scanned with barcode readers to input sample records into an Excel spreadsheet or other LIMS. This data is used to create electronic plate layouts to track grid positions on the 96-well extraction and ddPCR plates. It is important to account for how liquid handling robots fill and pull from 96-well plates. For example, in our protocol, the RNA extraction instrument we use (the chemagic 360) fills by row, whereas the droplet reader software operates by column. As a result, the electronic layouts and extract plates need to be transposed in between use of these instruments.

1.4.2.4. Waste Management

Each step of the testing process generates biohazard and materials packaging wastes that require regular disposal. Raw wastewater concentration waste is to be treated with 10% bleach and allowed to sit for a minimum of 30 minutes prior to disposal via the drain. Concentration waste can also be autoclaved for 15 minutes (liquid cycle) and dumped down the drain. Pipette tips and centrifuge tubes used for all steps should be treated as biohazardous waste. Instruments that produce liquid chemical wastes, such as the chemagic 360, JANUS robots, and ddPCR readers, require regular emptying and disposal. Biohazard and chemical wastes should be disposed of in accordance with Environmental Health and Safety protocols and any established standard operating procedures for the laboratory.

1.4.3. Personnel Needs

The number of staff (research technicians, research scientists, and/or graduate students and postdoctoral researchers) needed for wastewater analysis will depend on the number of samples to be processed each week and required turnaround time. The most time-consuming step of the analytical process is concentration of wastewater samples, but there are a number of related laboratory tasks that should be delegated to personnel to ensure smooth sample processing: keeping laboratory spaces, refrigerators, and freezers clean and organized; maintaining the necessary supply of consumable materials and preparing them for weekly use; disposing of biohazard and packaging material wastes; maintaining laboratory safety standards; and maintaining and cleaning equipment on a regular basis.

1.5. Assays for Quantification of SARS-CoV-2 RNA in Wastewater

There are numerous testing methods and laboratory workflows for measuring SARS-CoV-2 RNA in wastewater samples. After comparing five different methods (LaTurner et al. 2021), our lab developed the workflow described here, which includes procedures for sample storage, sample concentration, RNA extraction, and quantification of viral RNA targets.

Detailed instructions for our specific protocols can be found in our Standard Operating Procedures.

1.5.1. Sample Storage

Samples are transported in a cooler with ice and stored in a 4°C refrigerator. Samples are aliquoted into smaller tubes that are stored at 4°C until concentration. Samples are typically stored for no longer than 24 hours before concentration occurs, but can be re-tested up to a week after receipt with minimal loss in viral signal. It is good practice to retain all raw samples until analysis is complete in case retesting is required.

1.5.2. Concentration

Because SARS-CoV-2 viral particles are typically present at very dilute concentrations in wastewater, a concentration step is performed to improve detection of the virus. First, samples are clarified via centrifugation to remove large solids and aid in the throughput of the filtration-based concentration process. After centrifugation, the pellet and any remaining solids are discarded and the supernatant is saved for sample concentration via filtration. Some portion of the virus that is sorbed to the solids is lost during this step. Next, magnesium chloride ($MgCl_2$) is added to the sample, followed by electronegative filtration using HA filters. The $MgCl_2$ ensures that the viral particles bind to the negatively charged filter (LaTurner et al. 2021). $MgCl_2$ is added to the sample to achieve a final concentration of 25 mM. Once the sample and concentrated $MgCl_2$ solution have been mixed and allowed to stand for five minutes, the sample is filtered through either a magnetic filtration manifold unit or a disposable filtration unit until no liquid remains. If the sample cannot be completely pulled through the filtration unit, the excess liquid should be removed and measured so that the amount filtered can be adjusted during analysis calculations. The filters are folded and placed into bead beating tubes that have been pre-loaded with glass beads and lysis buffer to prepare them for RNA extraction. The bead tubes are then stored at 4°C to await further processing.

1.5.3. RNA Extraction

We perform RNA extraction using the chemagic 360. We follow the manufacturer's recommended protocols, with the following modifications in the sample preparation process prior to loading: We add the manufacturer's lysis buffer to each bead tube containing the filter generated from concentration. The addition of lysis buffer to the filter ruptures the cells, allowing the release of their contents. We then perform bead beating to agitate the samples and release nucleic acids from cells and viral particles. After bead beating, the tubes are centrifuged to pellet the beads and shredded filters. The supernatant from each sample is loaded into a 96-deep well plate according to a predetermined plate layout. In accordance with chemagic 360 protocol guidelines, each sample well is diluted with more lysis buffer and then a mixture of Proteinase K and Poly(A)RNA is added to each sample. Proteinase K digests proteins that could cause contamination and also inactivates nucleases that could potentially degrade the RNA. Poly(A)RNA increases the stability of the RNA molecules. It is important that extraction begin within 10 minutes of adding the Proteinase K and poly(A)RNA mixture to the lysate plate because the solutions will begin to react with one another once combined. After the chemagic 360 completes its protocol, the elution plate contains the RNA extracts that will be used for further processing and analysis. Extracts are stored at 4°C for no longer than 24 hours to await ddPCR.

Detailed instructions for the chemagic 360 RNA extraction protocol can be acquired from the manufacturer (PerkinElmer).

1.5.4. Quantification of SARS-CoV-2 Using RT-ddPCR

We perform quantification of the SARS-CoV-2 N gene (N1 and N2 targets as published by the CDC) using the Bio-Rad QX200 Droplet Digital PCR system (the QX200). Our applications require use of the Bio-Rad 1-step RT-ddPCR Advanced Kit for Probes, which contains supermix, reverse transcriptase, and dithiothreitol as a stabilizer. These three reagents are combined with primers and probes for the targets in our duplex N1/N2 assay. Using a JANUS robot, mastermix and 10 µL of RNA extract (or water for no-template controls; see Section 1.5.5) is added to each reaction well of a PCR plate according to a predetermined sample plate layout. The plate is then sealed using a Bio-Rad heat sealer with pierceable aluminum foil seals. To ensure adequate mixing of the mastermix and RNA extracts before droplet generation, the plate is vigorously mixed for 1 minute using a vortex mixer, then centrifuged for 5 minutes at 4,100 xg and 4°C. Once all reagents are at the bottom of the wells, we carefully place the plate into the droplet generator along with the necessary consumables. After droplet generation, samples are placed into a thermocycler. Finally, we place the plate in the QX200. The template must be set up correctly for the channels to read N1 and N2 properly. This is based on the corresponding probe's dye setting (FAM/ch.1 & VIC/ch.2 for this method). Once processed, any remaining RNA extracts are stored at -80°C. Although it is possible to thaw samples for further testing, degradation can occur when there are multiple freeze-thaws.

1.5.5. Quality Controls

To reduce the risk of contamination, the following quality controls should be incorporated into each step of sample processing: deionized water replicates during concentration; empty bead tubes during extraction; and RNase-free water in place of extracts during quantification. If there are three or more positive droplets in any of the controls, the possibility of sample contamination during processing must be evaluated. In addition, during quantification, a positive control containing SARS-CoV-2 RNA or DNA (gBlocks) should be added to each ddPCR plate. Positive controls ensure that the ddPCR assay can efficiently amplify SARS-CoV-2 N1/N2 genes. In addition, we periodically perform two different assays as checks of sample quality and recovery controls: (1) we spike a surrogate virus (bovine coronavirus) into the samples prior to concentration as a recovery control and (2) we quantify the presence of a fecal indicator virus (pepper mild motile virus) to check if there are issues with the collection of a representative sample.

1.5.6. Limit of Detection

We use a custom R script to calculate the limit of detection (LOD) for each ddPCR plate analyzed. The LOD for SARS-CoV-2 N1/N2 is calculated as follows: 3 droplets are set as the threshold. The R script searches the data for 3 droplets and assumes that copy number concentration to be the threshold. If samples with 3 droplets are detected, then the concentrations for all samples with 3 droplets are averaged. If there are no 3-droplet samples on the plate, the script assumes a value of 0.7 as the concentration (3 droplets per 10,000 total droplets, the most conservative estimate). The limit of blank (LOB) is then calculated. The LOB is the mean concentration of all negative control samples on the plate plus 1.6 times the standard deviation of the negative controls. The 3-droplet concentration is added to the LOB to determine the LOD. A template volume of 10 µL RNA extract is used to ensure adequate detection of low abundant viruses (human viruses are typically of low abundance in environmental samples compared to viruses that infect plants, animals, and bacteria). When dealing with very high concentrations, dilution may be needed if there is no clear separation between droplet clusters. Excessive rain in the droplet data (see Section 1.5.7) can also result from target concentrations that are too high.

1.5.7. RT-ddPCR Analysis

All sample thresholds should be set manually. Accepted droplet counts should be checked to ensure that each sample has at least 10,000 positive droplets. According to the manufacturer (Bio-Rad), a droplet count of 10,000 is necessary for robust statistical calculations when using QuantaSoft software. If samples display any of the following issues, they should be reprocessed: (1) low droplet counts, (2) evidence of contamination (e.g., negative controls with amplification, high LOD), (3) unclear droplet separation, and/or (4) significant differences in replicate concentrations.

RT-ddPCR should generate a clear separation between positive and negative droplet clusters. However, scattered droplets (rain) may appear in a sample. Rain droplets emit intermediate fluorescence and are difficult to classify as either positive or negative detection events. There is no single cause of rain; it can result from sample impurities, DNA degradation, insufficient amplification, or other issues (Kokkoris et al. 2021). One way to reduce the incidence of rain is to optimize the annealing temperature, which contributes to the specificity of PCR reactions. The proper annealing temperature range should be based on validated methods for the selected primer/probe set and can be adjusted to improve droplet separation between positive and negative clusters.

1.5.8. Precision and Variability

The use of automated laboratory equipment (i.e., the chemagic 360, JANUS robots, and repeater pipettes) and strict adherence to standard operating procedures facilitate adequate repeatability and reproducibility. During the concentration process, at least two replicates should be used to increase confidence in the final data and to assess variability in analytical methods. Quality controls should be used in each step of the sampling process to identify systematic errors such as contamination. If a change in method is implemented, the same set of samples should be quantified pre- and post-method change to assess the impact of the change on the results. In addition to relying on replication to assess variability within a single lab, multiple labs can process samples in parallel to assess variability across locations.

1.6. Screening for Variants of Concern or Interest Using Targeted Amplicon Sequencing

We screen for variants of concern in wastewater by amplifying and then sequencing the SARS-CoV-2 genome. RNA extracts are used to generate cDNA and amplified using the ARTIC v3 protocol. Sequencing is performed on an Illumina MiSeq System, as described in Lou et al. (2022). Sequences are analyzed using a custom bioinformatics pipeline described in Sapoval et al. (2021). The allele frequency of each quasi-unique mutation—defined as mutations present in at least 50% of the available genomes of the lineage of interest and not present in 50% or more of any other lineage—is summarized and reported to the HHD each week for variants of concern or interest. In addition to quasi-unique mutations, we report the number of characteristic mutations detected for each variant of concern or interest. A standard operating procedure for routine sequencing and sequencing analysis is being developed.

1.7. Other Factors that Could Impact Comparisons between Viral Concentrations and Disease Incidence

Sampling design, system characteristics, and environmental parameters are all factors that can impact measured viral concentrations in wastewater and associated relationships with disease incidence in communities. One critical consideration is the stability of the viral genetic material in the sewershed. Degradation of viral genetic material during transport through the sewer system can reduce the concentration in a sample collected at a downstream point, such as at the influent of a wastewater treatment plant. We performed a study in which we modeled SARS-CoV-2 RNA degradation in sewersheds across Houston that vary in service population and geographic area (McCall et al. 2022). In the study, we used published and experimentally derived first-order decay rates of SARS-CoV-2 RNA and looked at the impact of wastewater temperature and wastewater travel times on viral RNA decay. We found that wastewater travel time had a greater influence on RNA degradation across the sewershed than temperature. Caution should be exercised in directly comparing viral concentrations in samples collected from sewersheds in widely different geographic areas. In general, we recommend comparing viral loads (concentration * flow) from a single sewershed over time, as opposed to comparing viral loads from multiple locations over space at a single timepoint, for decision-making (discussed in further detail in Section 2).

Another factor that can impact viral concentrations in wastewater samples is dilution due to infiltration and inflow of stormwater into the sewer system. Houston has separate sewers, but many sewer systems across the United States are combined, meaning they convey both wastewater and stormwater. Wet weather, particularly in combined sewers, can substantially dilute viral concentrations in wastewater samples and introduce constituents that can inhibit PCR. We recommend using flow-based normalization (also referred to as load-based analysis) because the approach computes the concentration * flow, and thus theoretically should not be impacted by wet weather as any dilution due to increased flow would be accompanied by a concomitant increase in the measured flow rate. Flow-based normalization is described in further detail in Section 2.2.3. Reverse transcription and PCR inhibition should be checked using positive controls of spiked synthetic RNA.

Section 2:

Best Practices to Develop a Statistical System to Analyze Wastewater Data

2.1. Introduction

The statistical system we use to analyze data from Houston's SARS-CoV-2 wastewater monitoring system is divided into two components: (1) analysis of data from wastewater treatment plants (WWTPs) and (2) analysis of data from lift stations and manholes associated with congregate living facilities and schools. The system is versatile and allows us to generate significant public health information at different spatial and temporal scales.

Sampling locations for the wastewater monitoring system include:

- 39 WWTPs (see Figure 3)
- 63 lift stations
- 2 jails
- 11 shelters
- 9 nursing homes
- 51 schools

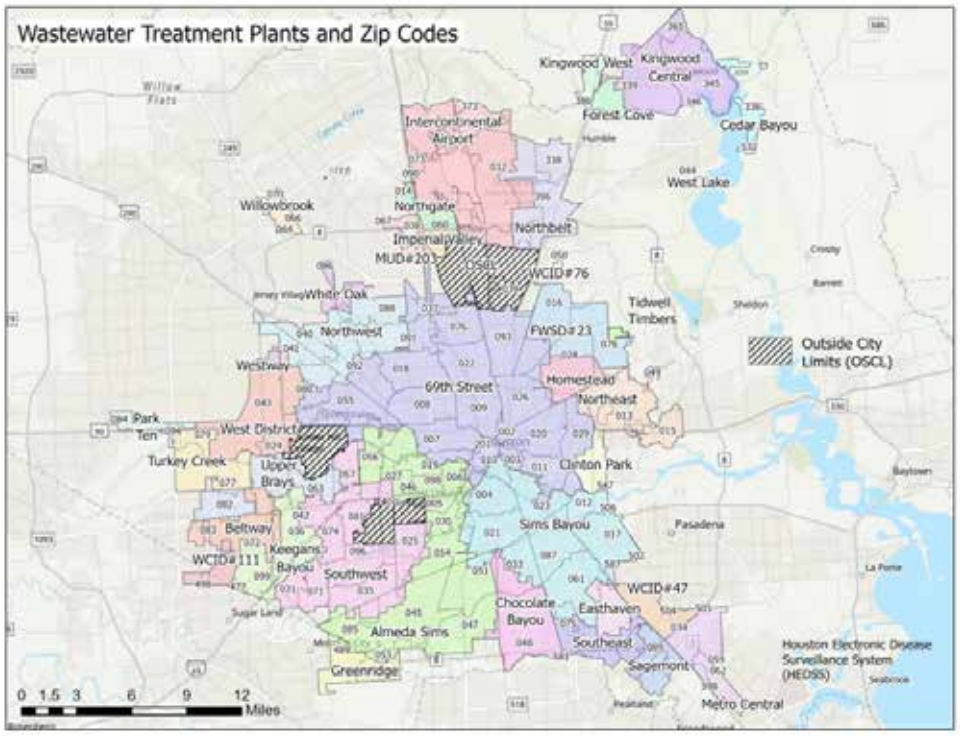


Figure 3. WWTP map

The 39 WWTPs are sampled weekly. Data from these samples forms the analytical foundation for assessing SARS-CoV-2 levels in Houston. The remaining locations are sampled either weekly or on an as-indicated basis. For example, there is no reason to sample school manholes when school is not in session.

Section 2.2 describes our methods for analyzing wastewater data from the WWTPs. Our statistical system can generate a temporal analysis of information from a single WWTP or a spatial-temporal analysis that combines data from multiple WWTPs to obtain an overall estimate for the city. Note, flow information is available for normalization across WWTPs serving a wide range of population sizes. The statistical system was built in the programming language R. It is not automated, but it is streamlined so that once data is received, an analyst can produce a weekly report in approximately one hour.

Section 2.3 describes the analytical methods we use to study wastewater data from lift stations and manholes, which provide a real-time, spatial view of findings with simple trend metrics to assess change. These analyses are performed in Excel with automated scripts.

Internal reports derived from our statistical analyses are used to inform the HHD and other government and health authorities about the spread of COVID-19 in Houston. Our summary analyses are also displayed on an interactive dashboard hosted on the HHD's COVID-19 website. For more detail on our communication strategy, see Section 3 below.

2.2. Statistical System for Wastewater Treatment Plant Data

Twenty-four hour composite samples from the 39 WWTPs are collected between Monday morning and Tuesday morning of each week. Each composite sample is divided in half. One half-set is delivered to the HHD lab and the other half-set is delivered to the Rice Labs. Each lab splits the samples it receives, resulting in two quantification measurements per sample per lab.

A breadth of information is available for each WWTP, including geographic coverage, population served, and measured flow rates. The populations served by the WWTPs range from approximately 10,000 to 500,000 people.

Our statistical system for analyzing wastewater data from the WWTPs addresses the following key issues:

- Measurements falling below the LOD
- Technical replicates and calibration across labs
- Normalization across WWTPs
- Missing measurements
- Temporal modeling in individual WWTPs
- Combination across WWTPs to create an overall city estimate
- Reference date and communication of results
- Uncertainty quantification
- Visualization of results
- Reflections on analyses

2.2.1. Limit of Detection

Depending on the lab method used, the number of measurements falling below the LOD can be small or large. The Stadler team has piloted quantification methods that limit the percentage of samples falling below the LOD. Further, after an extensive sensitivity analysis, we have adopted a simple rule to address measurements falling below the LOD: measurements that fall below the LOD are replaced by the maximum of the measurement and a random value that is between zero and one-half of the LOD.

2.2.2. Technical Replicates and Calibration Across Labs

In the WWTP analysis, replicates are grouped but not immediately averaged. The rational for this is that variability in the technical replicates contributes to the uncertainty quantification when producing a temporal profile for each WWTP.

Because the HHD lab provides a total quantification of the N1 and N2 genes, the corresponding individual measurements from the Rice Labs are combined. If separate quantification is available for N1 and N2 genes, this additional information can be incorporated into the analysis by foregoing the aggregation.

Calibration across labs is performed through regression. Measurements from the Rice Labs are considered the "truth" and measurements from the HHD lab are adjusted based on the appropriate regression model. We currently use a cubic-polynomial regression model (see Figure 4).

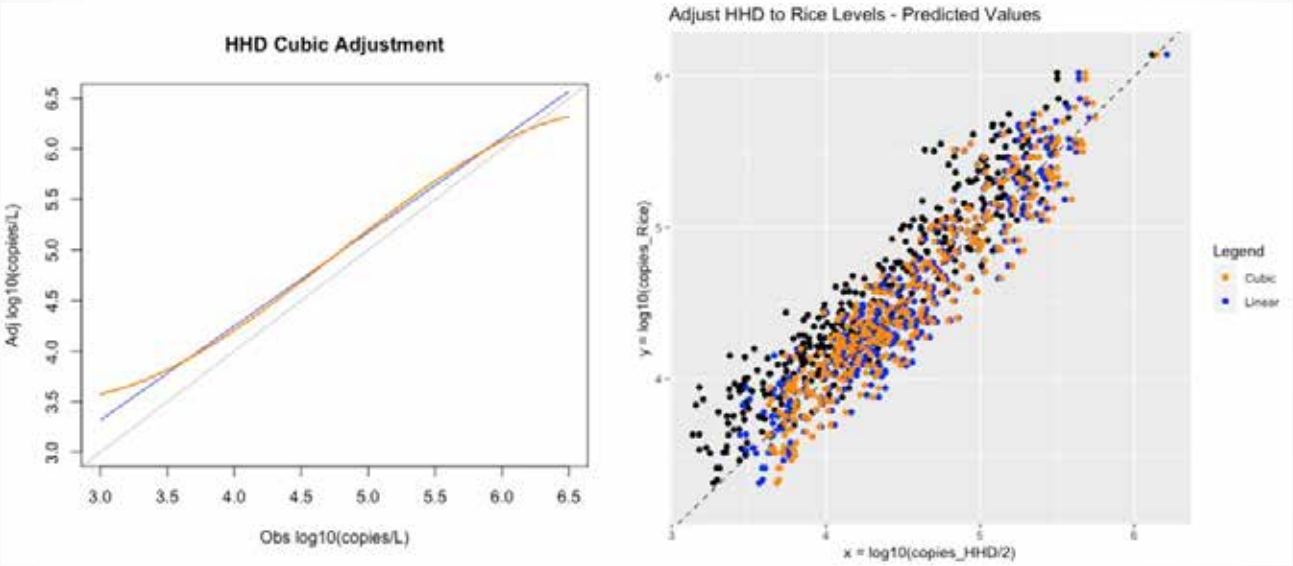


Figure 4. Cubic regression model

2.2.3. Normalization Across WWTPs

The 39 WWTPs in the monitoring system vary widely in size, serving between approximately 10,000 and 500,000 people (see Figure 3). Measurements are reported in copies/liter and are normalized by flow, which is measured in liters/day. The final measurements are reported as copies/day and our analysis is based on copies/day.

To date, we have not investigated variability and uncertainty in flow measurements, but are doing so in a separate project. We have conducted sensitivity tests related to this normalization methodology and put it forward as best for our current purposes.

2.2.4. Missing Measurements

Up to two consecutive missing flow measurements are imputed based on linear temporal imputation for each WWTP.

Up to two missing copies/day measurements may be imputed based on spline temporal imputation for each WWTP. Uncertainty quantification is currently not performed.

When more than two consecutive copies/day measurements for a WWTP are missing, we consider whether imputation is reasonable. If there are more than four weeks of missing values in the time series, imputation is not performed and the impact of the affected WWTP on the overall estimate for the city is minimized. For a moderate number, three to four, of missing values in the time series, the imputation method is spline-based, but is reviewed by an analysist to ensure the validity of the imputation.

2.2.5. Temporal Modeling in Individual WWTPs

A regression spline model is used to fit the log10 copies/day time series for each WWTP. The knots in the spline are chosen optimally based on the quantiles of the copies/day. The number of knots or degrees of freedom is a sensitivity parameter that is checked each week.

The transformation of log10 resolves the issue of extreme values and skewness of the copies/day measurements.

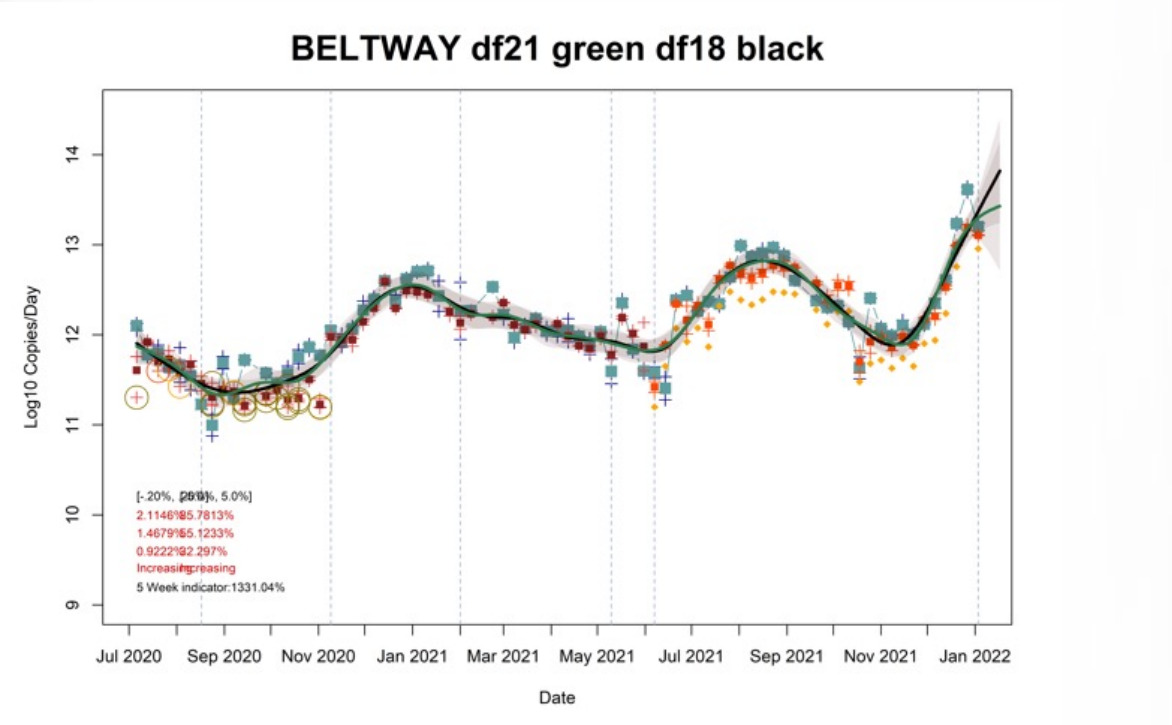


Figure 5. Diagnostic plot

Initially, the degrees of freedom is optimized to ensure the accuracy of our estimates. A visual sensitivity check is then performed by varying the degrees of freedom around the optimal. We are concerned with over fitting and unusual dynamics at the end of the series. WWTP data can be highly variable due to lab or sampling issues, and any model may be inappropriately impacted by incorrect measurements. We developed diagnostic plots (see Figure 5) and review the fit each week before finalizing our estimates. Even as we refine the science to account for anomalies in the data, we still highly recommend a visual sensitivity check.

95% confidence intervals are computed and displayed for each WWTP estimate of the log10 copies/day. The uncertainty incorporates lab variation and sampling variation.

We allow the spline to extend two weeks beyond our estimates. These are not the best predictions of the next values (as indicated by the 95% confidence bounds), but are a simple way to project and display the direction of the viral prevalence trend in a WWTP. We produce statistically sound projections of positivity rate in a separate analysis, but these are not included in our standard WWTP reports.

2.2.6. Combination Across WWTPs

The estimated copies/day for each WWTP are summed to produce the total copies/day for the entire city. This requires taking the log10 estimated values and raising them to the 10th power prior to summing. The total estimated copies/day is again transformed to log10.

The total estimated copies/day is an estimate of the median of the probability distribution for copies/day. We do not perform the necessary adjustment to obtain an estimate of the mean of this distribution. Week-to-week uncertainty can change dramatically, translating to very large and unstable estimates of the mean.

2.2.7. Reference Date and Communication of Results

Early in the development of our statistical analysis system, we struggled with how to communicate results from multiple WWTPs of different sizes. The copies/day measurement did not resonate with our public health officials, so we rescaled these measurements by dividing by the population size within a WWTP to create a common impact scale across all WWTPs—an impact metric comparable to the CDC's recommended per capita scaling. Again, however, this metric was not useful to public health officials. Finally, we decided to tie our interpretations to the best estimates of virus levels as of July 6, 2020 for each WWTP. Figure 8 provides an example of a time series plot of estimated virus levels relative to July 6, 2020 for an individual WWTP (Southeast).

We chose July 6, 2020 as our reference date for two reasons. First, our wastewater data was high quality by this date. Second, by July 6, 2020, virus levels had declined across the city from their peak in May 2020. All plots and communication use July 6, 2020 as their benchmark and are relative to our estimates on that date (individual WWTP estimates are compared to their July 6, 2020 results; the overall total is compared to the overall total as of July 6, 2020).

2.2.8. Uncertainty Quantification

The 95% confidence limit for each WWTP spline estimate of log10 copies/day is computed. The estimate for the total is obtained based on this confidence limit. We do not adjust for correlation across WWTP spline estimates of log10 copies/day, resulting in an overestimate of the uncertainty.

2.2.9. Visualization of Results

WWTP results are presented in three basic formats: (1) maps that display the spatial distribution of results (see Figures 6 and 7); (2) time series plots of data from individual WWTPs and all WWTPs, collectively (see Figures 8, 9, and 10); and (3) heatmaps that depict both spatial and temporal changes in viral loads (see Figures 11 and 12). Each format provides valuable insight to our understanding of the level of COVID-19 in communities across Houston.

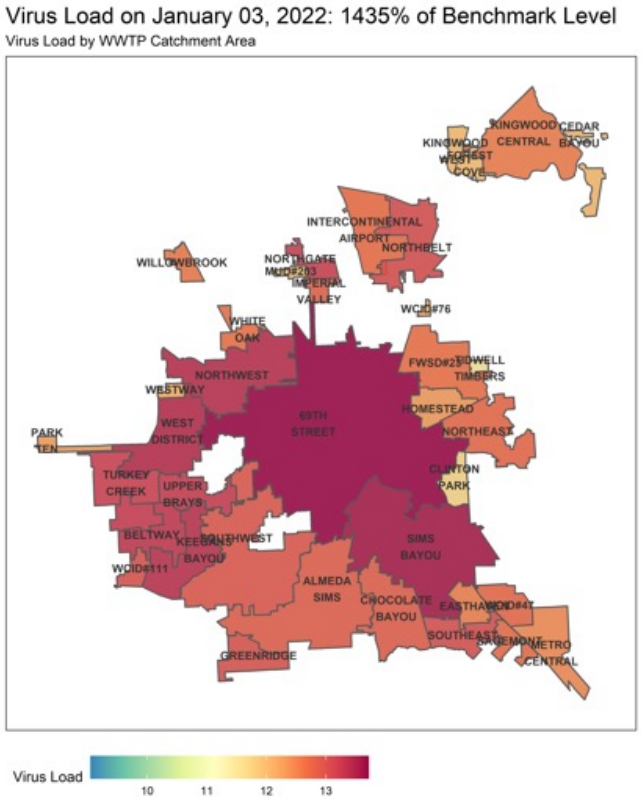


Figure 6. Map of estimated viral load in log10 copies/day for all WWTPs on a single day

The color scale range in Figure 6 is based on the estimated viral load for all WWTPs for all sample dates. The lowest value is the 1st quantile of the estimated viral load minus 1.5 times the interquartile range (IQR) of the estimated viral load. The highest value is the 3rd quantile of the estimated viral load plus 1.5 times the IQR of the estimated viral load. If a WWTP was not sampled on a certain date, its color on the map is gray. Scale calculations may be updated over time. The map is titled with the sampling date and the estimated viral load compared to the benchmark estimated viral load for the city. The benchmark date is July 6, 2020.

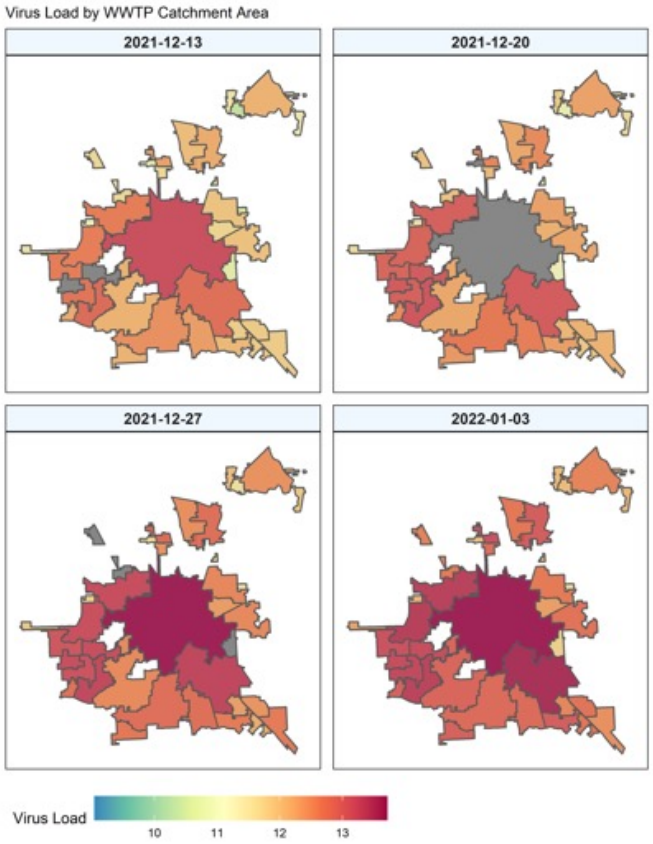


Figure 7 shows a series of Figure 6 maps plotted over a 4-week period. For legibility, WWTP names are removed. Each component map is titled with the appropriate sampling date.

The time series plot in Figure 8 has dual y-axes, one for wastewater viral load at the Southeast WWTP and one for PCR testing results in the associated sewershed.

The left y-axis is for the estimated viral load, in copies/day, for the WWTP, relative to the WWTP's benchmark estimated viral load. The time series for the relative estimated viral load is plotted as a green line. July 6, 2020 (the benchmark date) is the starting date for the time series and is therefore at 100%. A circle indicates the average relative viral load, in copies/day, for the sample date.

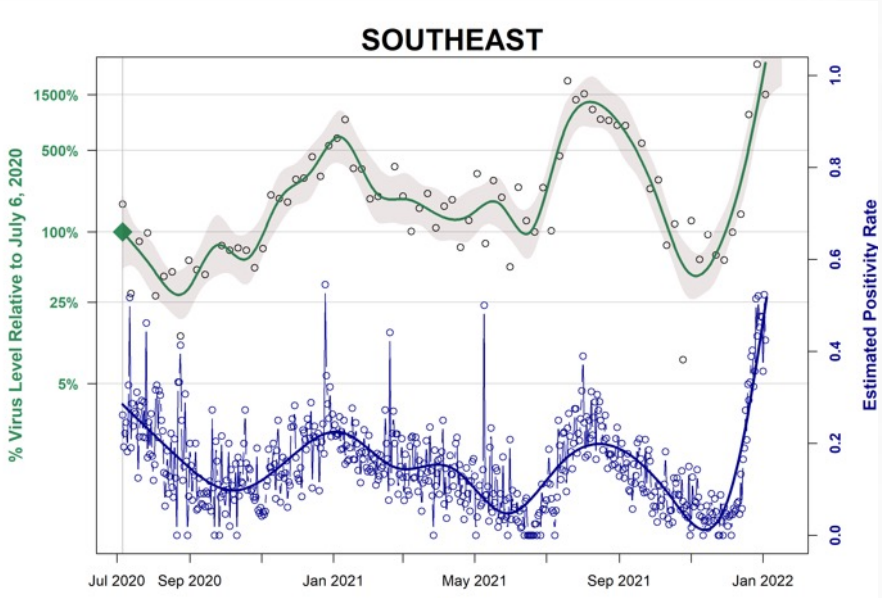


Figure 8. Time series plot of estimated WWTP viral load in copies/day for an individual WWTP and estimated positivity rate for the associated sewershed

The right y-axis is for the PCR testing positivity rate. A blue circle indicates the daily PCR positivity rate for residents in the WWTP's sewershed. This value is based on reported PCR test results, not on wastewater testing. The blue line is the estimated PCR positivity rate.

The estimated relative viral load in the WWTP and the PCR positivity rate are plotted up to the most recent wastewater sampling date.

A sample analysis of the Figure 8 time series plot: On July 6, 2020, the estimated viral load at Southeast was 3.817e+11 copies/day. On December 27, 2021, the estimated viral load was 4.465e+12 copies/day. Therefore, on December 27, 2021, the viral load was 1170% of the July 6, 2020 viral load.

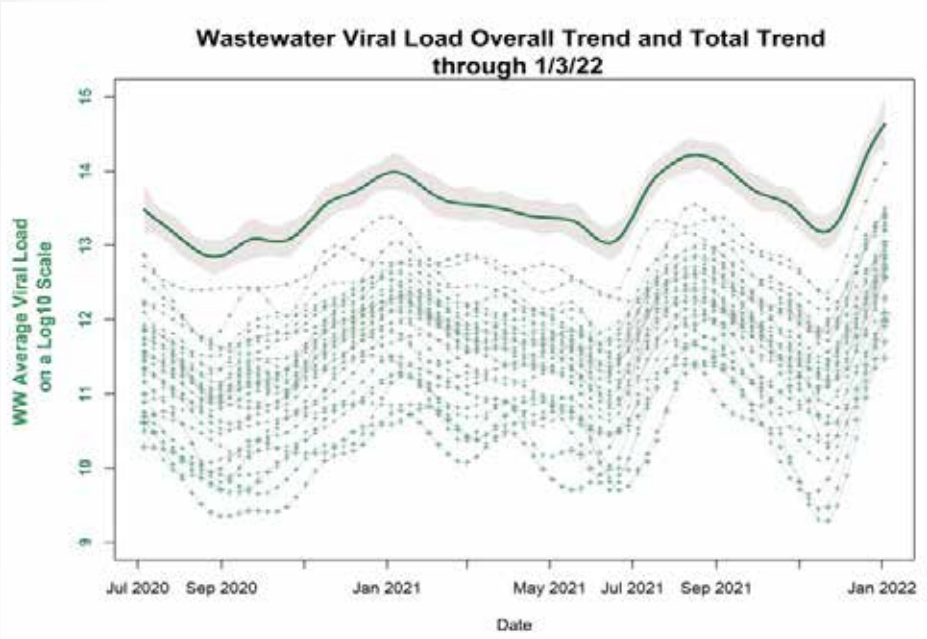


Figure 9. Time series plot of estimated viral load in log10 copies/day for all WWTPs and the City of Houston

In Figure 9, each WWTP's weekly estimated viral load is plotted as a green plus sign (+), the size of which is proportional to the size of the standard error of the viral load estimation. In other words, the larger the standard error, the larger the plus sign. A gray line connects each WWTP's weekly estimated viral load. The estimated total viral load for the city is plotted as a solid green line with the 95% confidence interval.

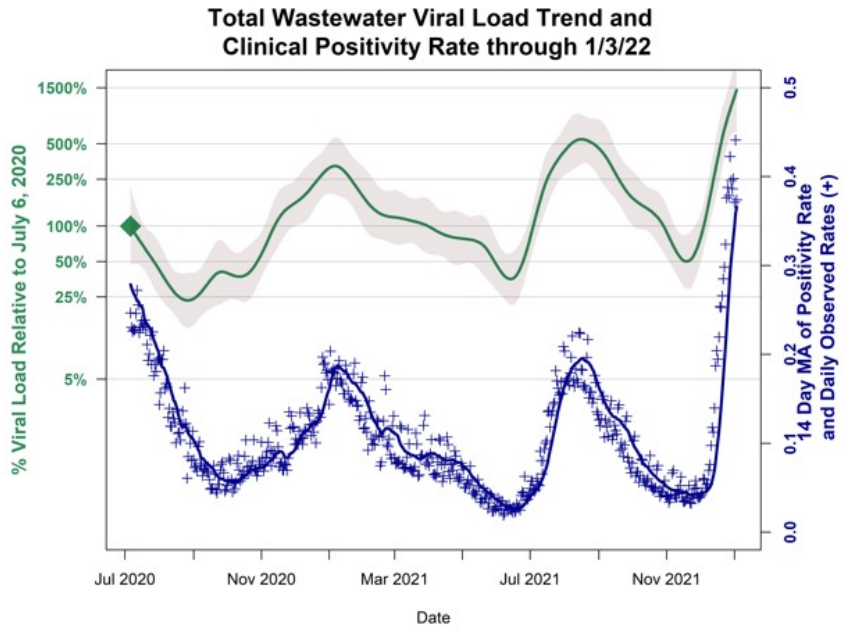


Figure 10. Time series plot of estimated viral load in copies/day for all WWTPs and estimated positivity rate for the City of Houston

The time series plot in Figure 10 has dual y-axes, one for estimated viral load at all WWTPs and one for PCR testing results for the entire City of Houston.

The left y-axis is for the estimated viral load for all WWTPs, in copies/day, relative to the benchmark estimated viral load. The time series for the relative estimated viral load is plotted as a green line. July 6, 2020 (the benchmark date) is the starting date for the wastewater time series and is therefore at 100%.

The right y-axis is for the PCR testing positivity rate for the city. A blue plus sign (+) indicates the daily PCR positivity rate. This value is based on reported PCR test results, not on wastewater testing. The blue line is the 14-day moving average of the daily PCR positivity rate.

The estimated relative viral load and the city-wide PCR positivity rate are plotted up to the most recent wastewater sampling date. The plot is titled with the latest sampling date.

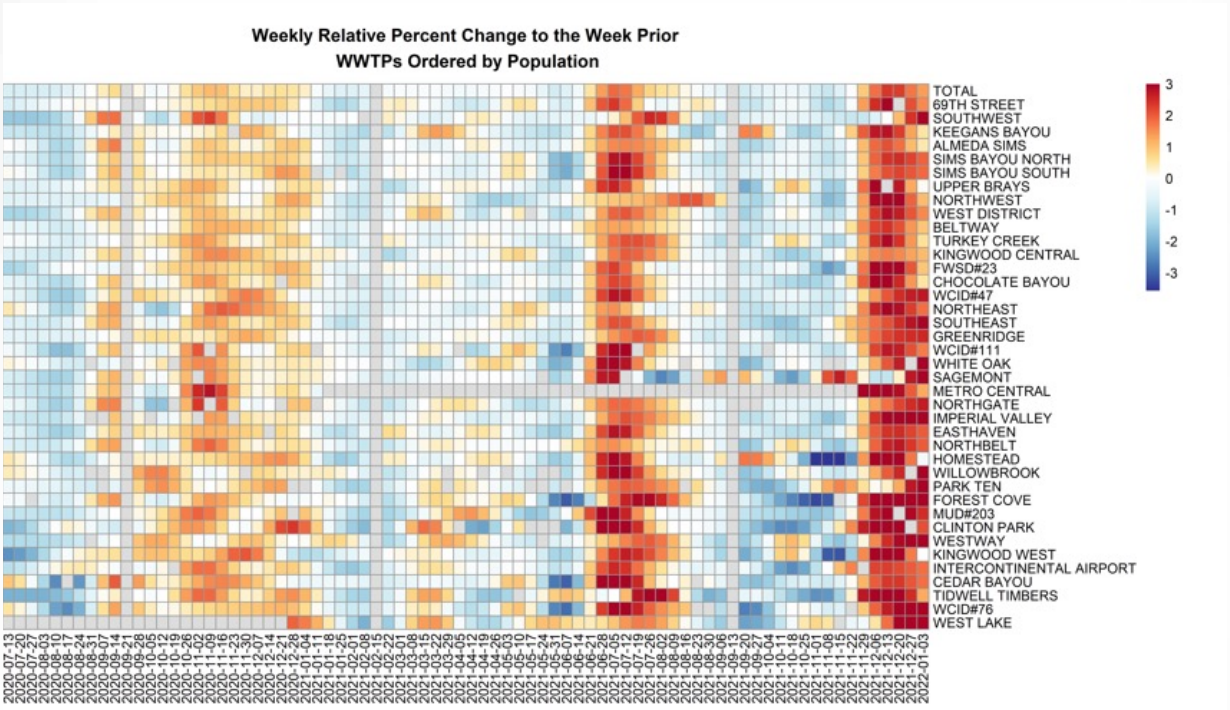


Figure 11. Heatmap of weekly change in estimated viral load at all WWTPs

The heatmap in Figure 11 shows the week-to-week percent change in the estimated viral load at all WWTPs in copies/day. Each square represents the percent change in estimated viral load for a specific WWTP on a given sampling date compared to the viral load at the same WWTP one week prior. Wastewater sampling dates are plotted on the x-axis; the WWTPs are listed on the y-axis, including an estimate for the entire city (labeled "TOTAL"). The ordering of the WWTP list is based on the size of the population each WWTP serves. Because there are no results from the week prior to July 6, 2020 (the benchmark date) upon which to base a comparison, the heatmap's starting date is July 13, 2020.

Estimated viral loads for this heatmap are converted to the log10 scale for plotting purposes. The values for the color scale range are based on historical week-to-week changes and are set at (-3.57 to 3). These values may be updated over time.

A sample analysis of the Figure 11 heatmap: On December 27, 2021, the estimated viral load at 69th Street was 13.88 log10 copies/day . One week later, on January 3, 2022, it was 14.1 log10 copies/day. The week-to-week difference was 0.22. Accordingly, the percent change in estimated viral load relative to the week prior was 1.6%.

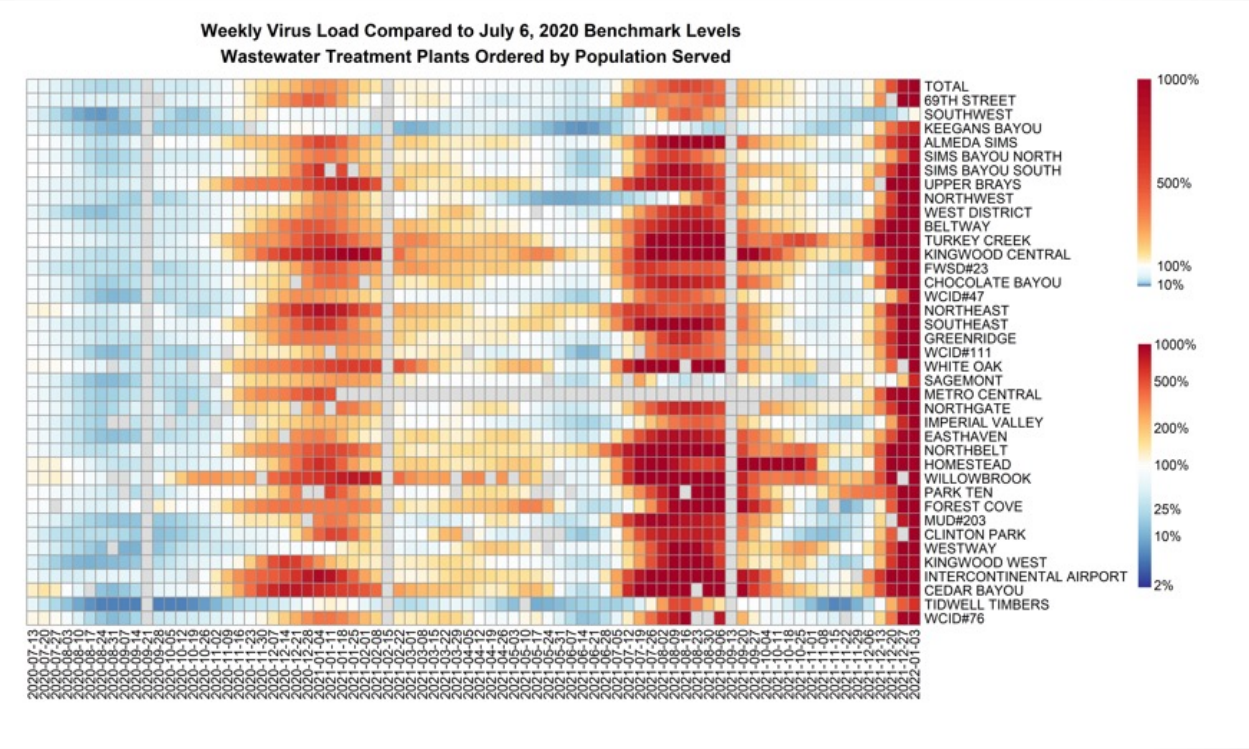


Figure 12. Heatmap of change in estimated viral load relative to the benchmark estimated viral load for all WWTPs

The heatmap in Figure 12 shows the weekly estimated viral load, in copies/day, relative to the benchmark estimated viral load for all WWTPs. Each square represents the estimated viral load at a specific WWTP for a given week divided by its estimated viral load on the benchmark date (July 6, 2020). Wastewater sampling dates are plotted on the x-axis; the WWTPs are listed on the y-axis, including an estimate for the entire city (labeled "TOTAL"). The ordering of the WWTP list is based on the size of the population each WWTP serves.

The color scale range on the heatmap is set at (2%, 1000%), and is in the log10 scale for plotting purposes. The two legends situated to the right of the heatmap provide different views of the same color scale. The top legend is automatically produced by the plotting function (pheatmap in r). Though the color scale values are distributed on the log10 scale, this legend displays the values as evenly distributed from 2% to 1000% (with 500% at the midpoint of the legend), making it difficult to see the color scale for values from 10% to 100%. To address this issue, the bottom legend displays the color scale values distributed on the log10 scale (with 100% at the midpoint of the legend), which defines the color scale for values from 10% to 100% more clearly.

Sample analysis of the Figure 12 heatmap: On July 6, 2020, the estimated viral load at Keegan's Bayou was 3.474e+12 copies/day. On January 3, 2022, the estimated viral load was 2.393e+13 copies/day. Therefore, on January 3, 2022, the viral load was 689% of the July 6, 2020 viral load.

2.2.10. Reflections on Analyses

There are multiple statistical approaches to analyzing WWTP copies/day measurements. This section has outlined a framework for conducting an analysis based on one observation per week at multiple locations, with technical replicates available across labs. We elected to use natural splines to estimate median copies/day of SARS-Cov-2. The spline approach is useful because it easily adapts to the changing circumstances of the virus, as represented by the weekly measurements at each WWTP. Further, this approach allows us to aggregate to the city level and quantify the uncertainty in our estimated viral load. We conducted extensive sensitivity

checks before implementing this methodology, and continue to conduct weekly sensitivity checks to ensure it is working as designed. This modeling choice has held up throughout our study period, resulting in strong statistical characterizations of estimated viral loads and accurate measures of uncertainty.

An important component of our communication strategy was the selection of a benchmark date for comparing results. Choosing a date just after the first peak of COVID-19 worked well for the city of Houston and SARS CoV-2. Another option would be to follow the CDC guidelines and communicate results using a per capita metric, which is computed as the estimated viral load divided by the number of people represented.

2.3. Statistical Summaries of Data from Lift Stations and Manholes

Composite samples from lift stations, congregate living manholes, and school manholes are collected weekly. Lift station and congregate living manhole samples are 24-hour composites and school manhole samples are 8-hour composites. Samples are collected according to predetermined routes. Each route is collected on the same day of the week, from Monday to Thursday. The Monday, Tuesday, and Wednesday samples are divided in half, with one half-set delivered to the HHD lab and one half-set delivered to the Rice labs. All Thursday composite samples are sent to the Rice Lab. Each lab splits the samples it receives, resulting in two quantification measurements per sample per lab.

We have a breadth of information about each lift station, including geographic coverage, ZIP code coverage, population served, and WWTP served. We are finalizing methods to measure flow rates at lift stations so that we can normalize results across all lift stations in the monitoring system. The populations served by the lift stations range from approximately 700 to 370,000 people.

Manhole information is more limited and includes only the facility's name and type.

Our statistical system for analyzing lift station and manhole data addresses the following key issues:

- Measurements falling below the LOD
- Classification of manhole results
- Classification of high positive manhole results
- Temporal monitoring of manholes
- Communication of manhole results
- Uncertainty quantification
- Visualization of results
- Current work: Lift station analysis
  - Technical replicates and calibration across labs
  - Normalization across lift stations
  - Temporal modeling in individual lift stations
  - Combining lift station data with WWTP data

2.3.1. Limit of Detection

Lift station measurements that fall below the LOD can be limited or extensive. Therefore, we apply the same simple rule to lift station data that we use in our WWTP analysis: measurements below the LOD are replaced by the maximum of the measurement and one-half of the LOD.

For manhole analysis, measurements below the LOD are left unchanged.

2.3.2. Classification of Manhole Results

Because we do not have flow information for manhole samples, measurements are not normalized. Instead of calculating copies/day, we calculate if SARS-CoV-2 was detected in the sample or not. Results are classified as positive, negative, or inconclusive.

The Rice Labs report separate N1 and N2 measurements for each replicate for each facility's manhole. Classification of these results is based on how many of the four measurements are above the LOD.

A facility is considered positive if both N1 and N2 on one replicate are above the LOD, and if N1 and/or N2 on the other replicate is above the LOD. In other words, if three or more of the four measurements are above the LOD, the facility's results are classified as positive.

A facility's results are classified as negative if none of the genes on either replicate is above the LOD.

A facility's results are classified as inconclusive in the following instances: (1) if both N1 and N2 on one replicate are above the LOD, but neither N1 or N2 on the other replicate are above the LOD; (2) if one gene, N1 or N2, is above the LOD on one replicate, and one gene, N1 or N2, is above the LOD on the other replicate; or (3) if one gene, N1 or N2, is above the LOD on one replicate, but neither N1 or N2 on the other replicate are above the LOD. In other words, if only one or two of the four measurements are above the LOD, the facility's results are classified as inconclusive.

2.3.3. Classification of High Positive Manhole Results

After analyzing the first few months of manhole data, we saw an opportunity to expand the definition of a positive classification by creating a high positive classification. If manhole measurements meet all of the requirements for a positive classification and at least one of the gene readings is above 10,000 copies/day, the facility is classified as high positive. The 10,000 copies/day benchmark was selected based on current measurements from our Houston-specific facilities and is not intended to be a generalized or universal cutoff point; we intend to change it as necessary.

Since the manhole analysis is not based on normalized measurements, we cannot compare an elementary school's positive to a county jail's high positive and interpret those results to mean that the county jail's manhole has a higher presence of the virus. Instead, we can use the distinction in positive classification to compare measurements over time only for a specific facility.

2.3.4. Temporal Monitoring of Manholes

After classifying each facility's manhole results, we count the number of consecutive weeks that current positive facilities have reported positive results. The count stops when a negative or inconclusive result is reported, or if there are more than 10 days between samples. Sampling for each facility usually occurs on the same day each week, but because of weather or holiday interruptions, sampling dates are not always consistent. The 10-day window allows for minor changes in the weekly sampling schedule not to affect the count of consecutive positive weeks for a facility.

2.3.5. Communication of Manhole Results

The city's public health officials initially requested that we report manhole sites that tested positive, inconclusive, or negative for SARS-CoV-2, as these are the most immediately actionable data points. To provide greater context, we now list sites according the following classifications: high positive, positive, inconclusive, negative, and not sampled. For each high positive and positive result, we include a number after the facility name, indicating the number of consecutive weeks the site has tested positive. This allows public health officials to quickly see which sites are newly positive. Several weeks after we began submitting reports, the "not sampled" classification was added to reduce confusion about why a particular facility might not appear on the list.

2.3.6. Uncertainty Quantification

Uncertainty quantification for manhole and lift station results is currently incorporated into the classification process. In other words, the level that signals a positive result takes into account potential randomness due to sampling and measurement.

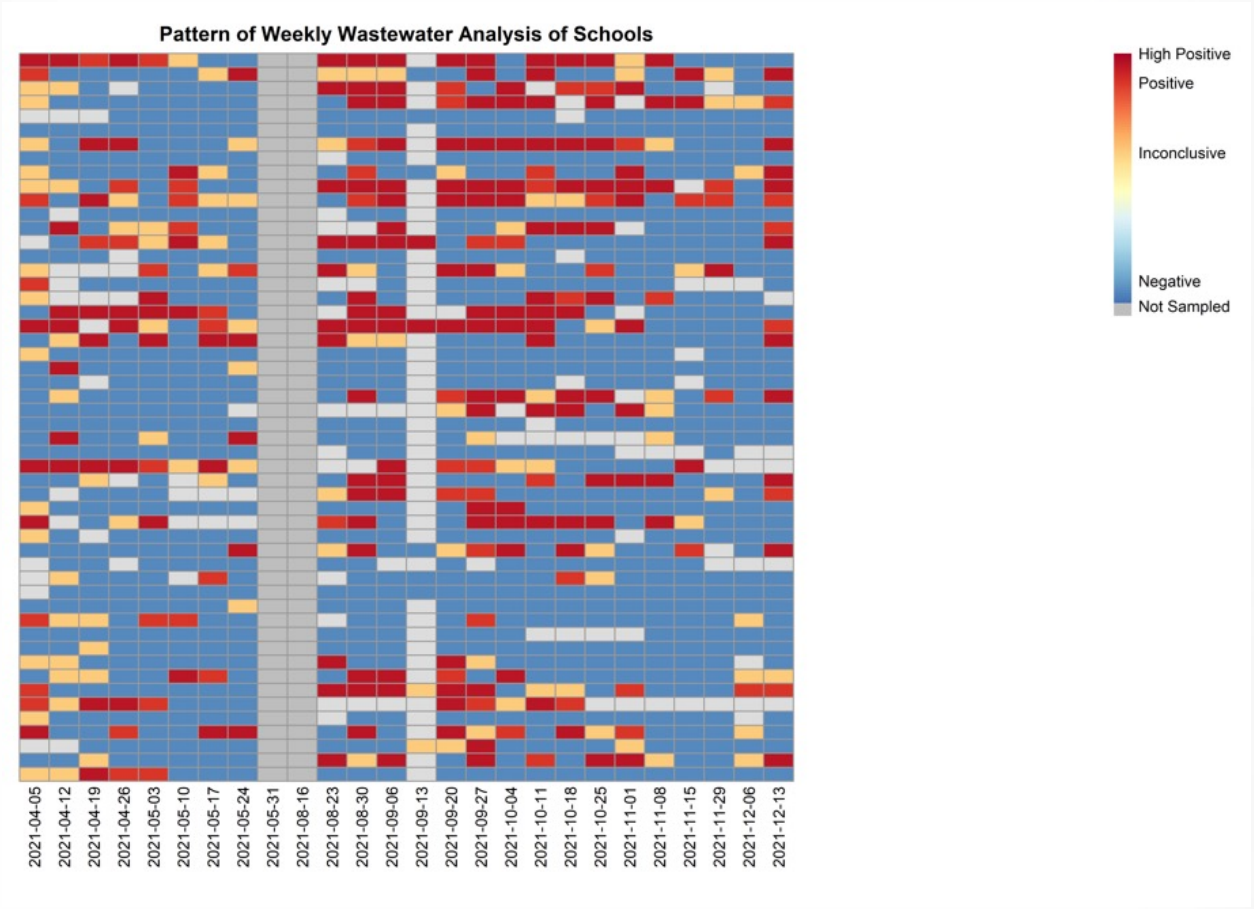
2.3.7. Visualization of Results

Manhole results are reported using a list of weekly results and heatmaps of all historical results. The list of weekly results provides easy-to-read, immediately actionable data for public health officials to share with relevant facilities (see Figure 13). The heatmaps provide additional, valuable insights about temporal changes in COVID-19 levels at schools and congregate living facilities (see Figure 14).

High Positives (9) total	**Lower Positives (6) total**	**Inconclusive (1) total**	SCHOOL 29
SCHOOL 1 [0] ++	SCHOOL 10 [0] +	SCHOOL 16	SCHOOL 30
SCHOOL 2 [0] ++	SCHOOL 11 [0] +		SCHOOL 31
SCHOOL 3 [0] ++	SCHOOL 12 [0] +	**Negatives (29) total**	SCHOOL 32
SCHOOL 4 [0] ++	SCHOOL 13 [0] +	SCHOOL 17	SCHOOL 33
SCHOOL 5 [0] ++	SCHOOL 14 [0] +	SCHOOL 18	SCHOOL 34
SCHOOL 6 [0] ++	SCHOOL 15 [1] +	SCHOOL 19	SCHOOL 35
SCHOOL 7 [0] ++		SCHOOL 20	SCHOOL 36
SCHOOL 8 [0] ++		SCHOOL 21	SCHOOL 37
SCHOOL 9 [0] ++		SCHOOL 22	SCHOOL 38
		SCHOOL 23	SCHOOL 39
		SCHOOL 24	SCHOOL 40
		SCHOOL 25	SCHOOL 41
		SCHOOL 26	SCHOOL 42
		SCHOOL 27	SCHOOL 43
		SCHOOL 28	SCHOOL 44

Number in bracket indicates the number of consecutive weeks a site has tested positive; ++ signifies high level of virus (>10,000 copies/L); + signifies lower level of virus (<10,000 copies/L)]

Figure 13. List of weekly manhole results (de-identified)



2.3.9. Current Work: Lift Station Analysis

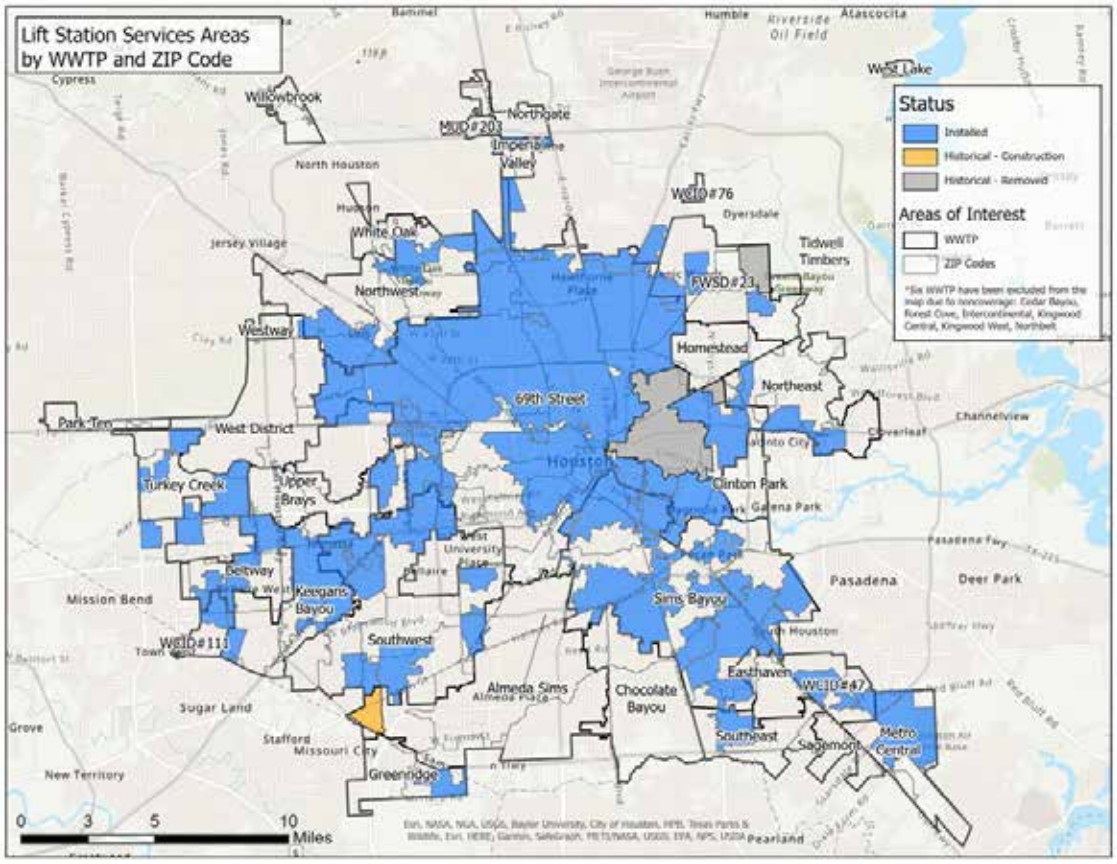
2.3.9.1. Technical Replicates and Calibration Across Labs

The same methodology for WWTP measurements is applied to lift station measurements. Analysis replicates are grouped but not immediately averaged. Because HHD provides a total quantification of N1 and N2, the corresponding individual measurements from the Rice Labs are combined.

We are working on calibration across labs. As in our WWTP analysis, calibration for lift station data will be performed through regression, with measurements from the Rice Labs considered the "truth" and HHD measurements adjusted based on the appropriate regression model.

2.3.9.2. Normalization Across Lift Stations

The lift stations we study vary widely in size, serving between approximately 700 and 370,000 people (see Figure 15). Unlike manholes, there is measurable flow at lift stations. Unlike WWTPs, however, flow information is not readily available for lift stations and must be estimated. We are working to capture flow at lift stations to normalize the copies/liter measurements, which will allow us to report results in copies/day.



2.3.9.3. Temporal Modeling in Individual Lift Stations

Once we can calculate copies/day measurements, we can begin to explore temporal modeling in individual lift stations. Each lift station flows into a WWTP, providing more geospatially targeted results. Temporal modeling of lift station results will provide targeted insights into spatial and temporal changes in COVID-19 levels in Houston's wastewater.

2.3.9.4. Combining Lift Station Data with WWTP Data

Once we have normalized data across lift stations, we plan to integrate the lift station analysis into the WWTP analysis to obtain a more comprehensive understanding of COVID-19 trends throughout the city.

Section 3:

Wastewater Epidemiology in Practice

3.1. Introduction

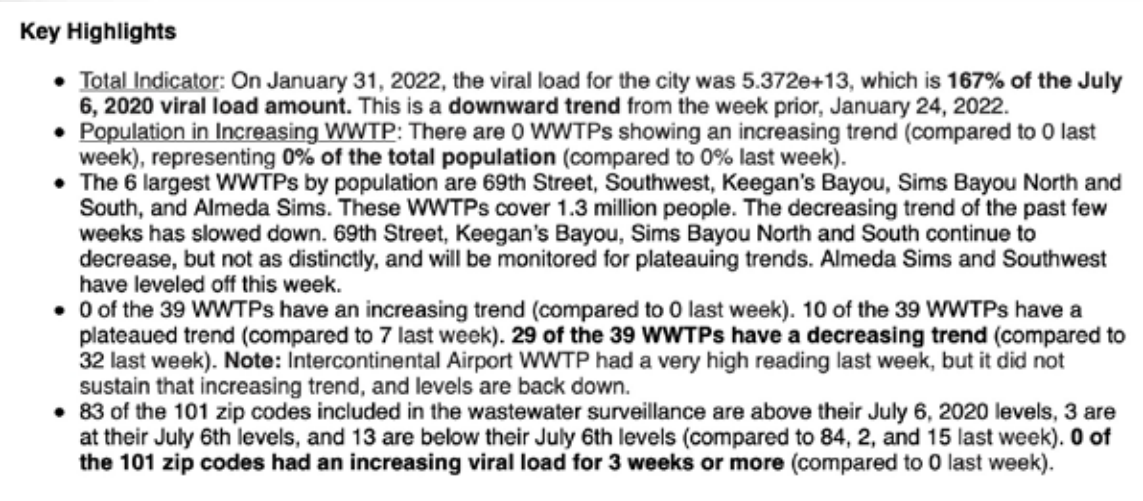
The statistical analysis of data from Houston's SARS-CoV-2 wastewater monitoring system is compiled in weekly reports that provide information about the virus at sewershed, ZIP code, and manhole levels. These reports are shared each week with the Chief Medical Officer for the City of Houston (the Health Authority) and the Director and Assistant Directors of the HHD (see Section 3.2), and are discussed in bi-weekly "Data to Action" meetings coordinated by the HHD (see Section 3.3).

The Health Authority communicates the information in the weekly reports to the Mayor of Houston, Texas Medical Center (TMC) leadership, and infectious disease doctors from major hospital systems in the area. When necessary, the public is notified of important developments via press conference. Data from the statistical analysis system is also shared with the public through an interactive dashboard hosted on the HHD's website and managed by the Spatial Studies Lab at Rice University (see Section 3.4).

The information in the weekly reports is used to track the spread of the virus, target interventions, and inform decision-making by government and health authorities. For example, TMC leadership compares trends in viral loads and positivity rates to data from TMC hospitals and uses the wastewater sewershed heatmap to determine where to allocate nursing staff during potential surges in COVID-19 cases.

3.2. Weekly Email to Houston Health Department Leadership

Each week, an email containing key highlights from the latest statistical analysis report (see Figure 16) and a list of results from the analysis of school manholes (see Figure 13) is sent to HHD leadership.



Key Highlights

- Total Indicator: On January 31, 2022, the viral load for the city was 5.372e+13, which is **167% of the July 6, 2020 viral load amount.** This is a **downward trend** from the week prior, January 24, 2022.
- Population in Increasing WWTP: There are 0 WWTPs showing an increasing trend (compared to 0 last week), representing **0% of the total population** (compared to 0% last week).
- The 6 largest WWTPs by population are 69th Street, Southwest, Keegan's Bayou, Sims Bayou North and South, and Almeda Sims. These WWTPs cover 1.3 million people. The decreasing trend of the past few weeks has slowed down. 69th Street, Keegan's Bayou, Sims Bayou North and South continue to decrease, but not as distinctly, and will be monitored for plateauing trends. Almeda Sims and Southwest have leveled off this week.
- 0 of the 39 WWTPs have an increasing trend (compared to 0 last week). 10 of the 39 WWTPs have a plateaued trend (compared to 7 last week). **29 of the 39 WWTPs have a decreasing trend** (compared to 32 last week). **Note:** Intercontinental Airport WWTP had a very high reading last week, but it did not sustain that increasing trend, and levels are back down.
- 83 of the 101 zip codes included in the wastewater surveillance are above their July 6, 2020 levels, 3 are at their July 6th levels, and 13 are below their July 6th levels (compared to 84, 2, and 15 last week). **0 of the 101 zip codes had an increasing viral load for 3 weeks or more** (compared to 0 last week).

Figure 16. Sample key highlights

The full report, which consists of the following maps, tables, and figures, is also attached to the email:

- figures showing the temporal trend of the viral load estimate and the positivity rate across the city, by sewershed and by ZIP code
- heatmaps of the sewershed viral load compared to benchmark and viral load relative to one-week percent change
- geographic information system maps of the sewershed viral load compared to the benchmark and of the viral load over the prior four weeks
- heatmaps showing detection of variants in the sewershed
- an excel file of ZIP code level information

3.3 Data to Action Meetings

The information in the weekly statistical analysis report is also presented and discussed at bi-weekly "Data to Action" meetings hosted by the HHD (DTA meetings). The purpose of the DTA meetings is to bring together program leads from the city's Covid-19 response network (e.g., school, nursing home, congregate living, jail, vaccination, outreach, testing, and reporting strike teams) so that interventions can be targeted based on the most current data.

3.3.1. ZIP Code Level Interventions

ZIP code level wastewater data is integrated with other pertinent information and analyzed to prioritize testing, education, outreach, and vaccination efforts.

The following key summary information is provided for each ZIP code in the city: COVID Community Vulnerability Index, cumulative positivity rate (historical infection), five-week positivity rate (current infection), unvaccinated rate (future infection), difference in positivity rate compared with one and two weeks prior, cumulative breakthrough count and rate, two-week breakthrough rate, estimated herd immunity, five-week positive cases per 1,000 people, percentage of the population that is Hispanic, percentage of the population that is Black, percentage of the population that is Asian, percentage of the population that is non-Hispanic white, and percentage of families below the poverty level (see Figure 17).

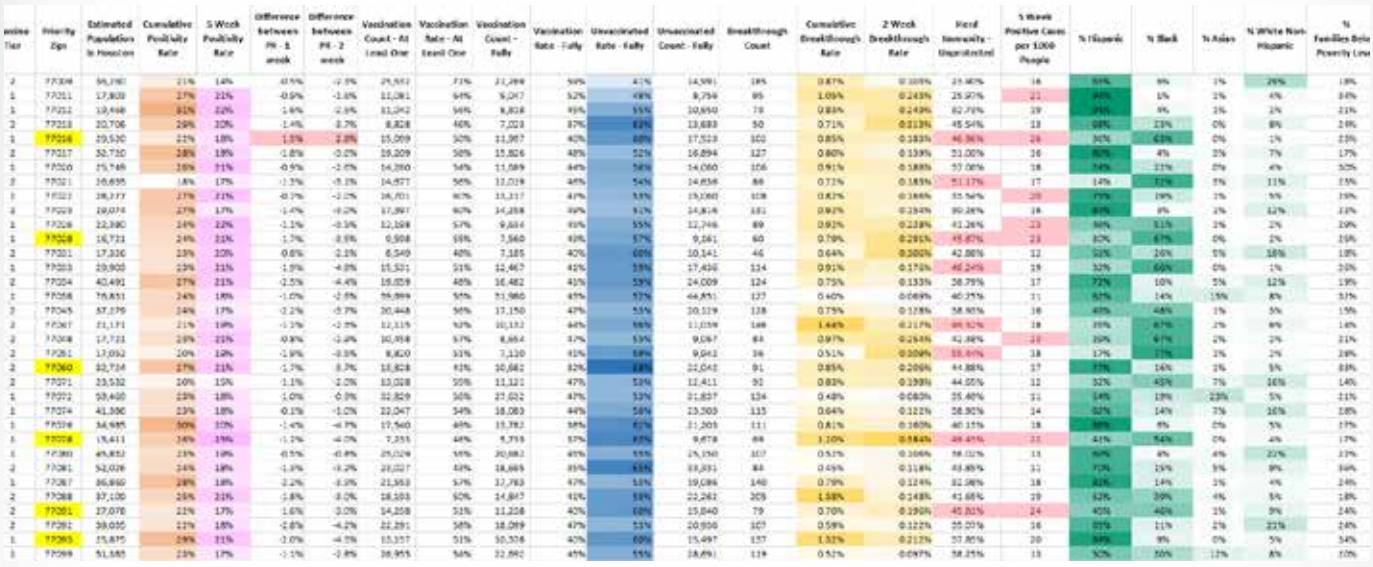
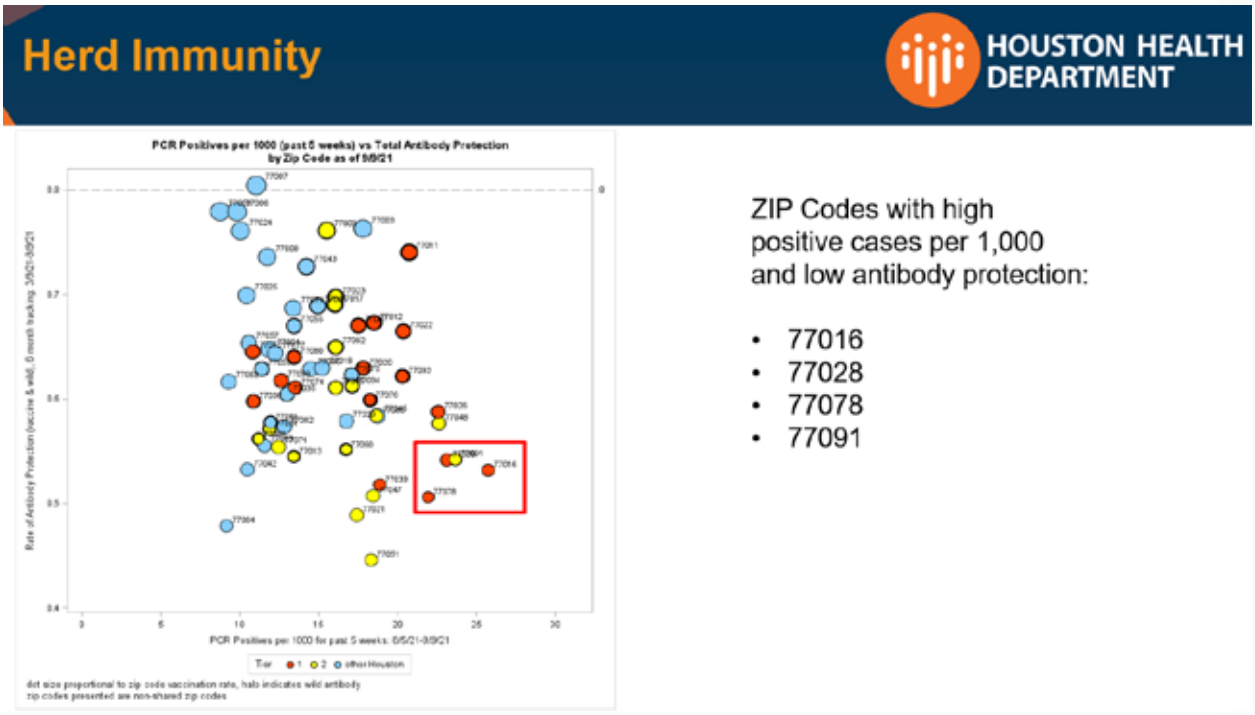


Figure 17. ZIP code summary information

Herd immunity (i.e., the percentage of residents in a particular ZIP code with COVID-19 antibodies) is calculated by combining the percentage of the population in a ZIP code that is vaccinated with the estimated percentage of the population that has natural immunity (using a model based on wastewater data, positive PCR counts, and seroprevalence study findings) (see Figure 18).



ZIP codes are identified for priority intervention using information from the ZIP code summary information or the herd immunity analysis chart:

1) high cumulative positivity rate (historical infection), high five-week positivity rate (current infection), and high unvaccinated rate (future infection); or

2) low total antibody protection (percentage of the population that is unprotected according to the herd immunity analysis) and high positive PCR counts per 1,000 residents.

Based on these criteria, priority ZIP codes are mapped to highlight areas of concern in the city and their spatial relationship to one other (see Figure 19).

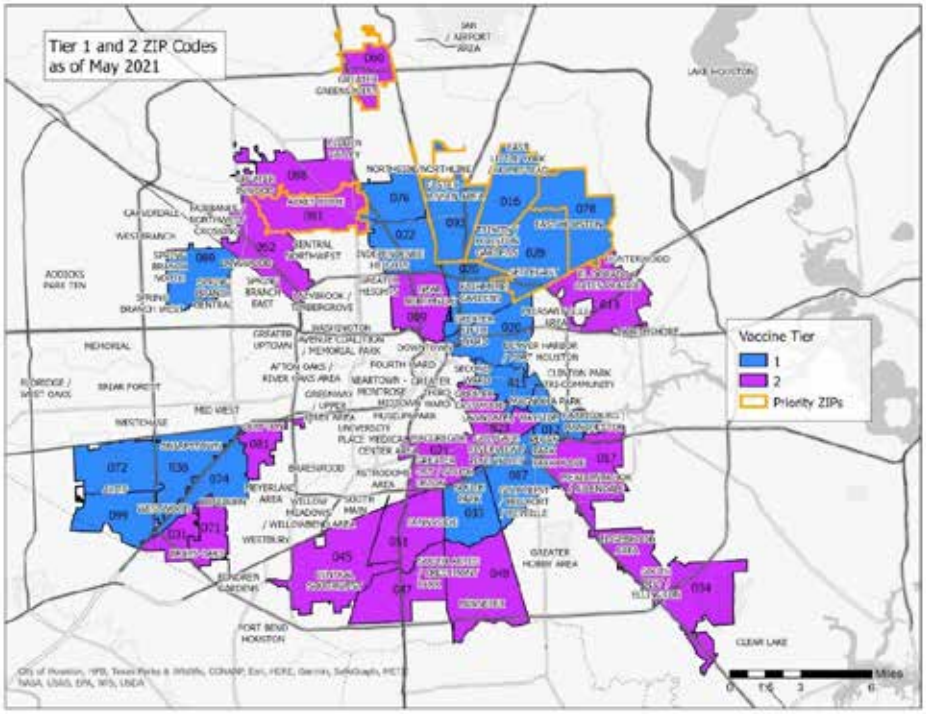


Figure 19. Map identifying priority ZIP codes

Detailed information about vaccination and positivity rates is summarized for each priority ZIP code: time series of the wastewater viral load relative to baseline and two-week prediction; time series of the estimated ZIP code positivity rate; time series of the one-dose vaccination rate; current one-dose vaccination rate and percent change since the previous week; current one-dose vaccination count; current fully vaccinated rate and percent change since the previous week; current fully vaccinated count; wastewater virus level (increasing, decreasing, or plateaued); cumulative positivity rate and percent change since the previous week; five-week positivity rate and percent change since the previous week; cumulative breakthrough rate and percent change from the previous week, and cumulative breakthrough count (see Figure 20).

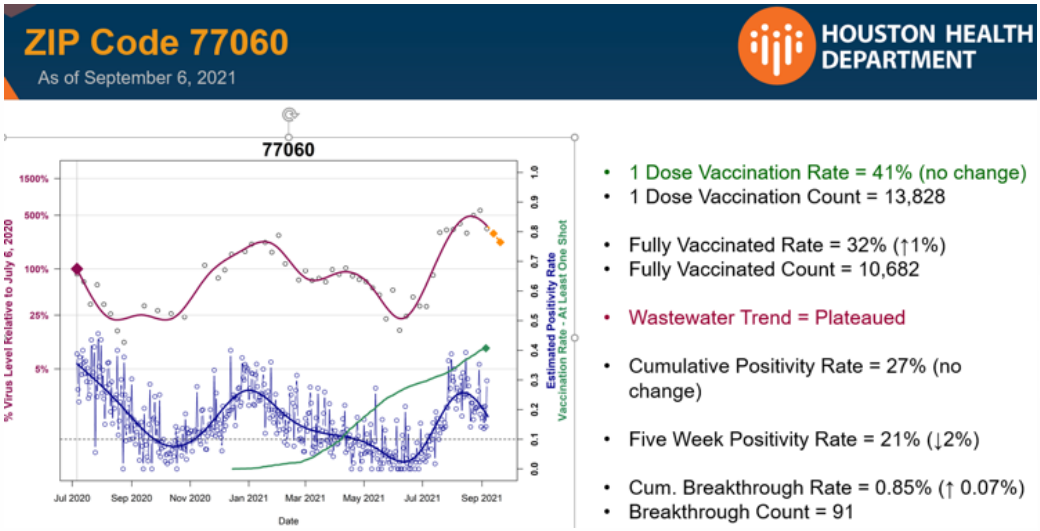
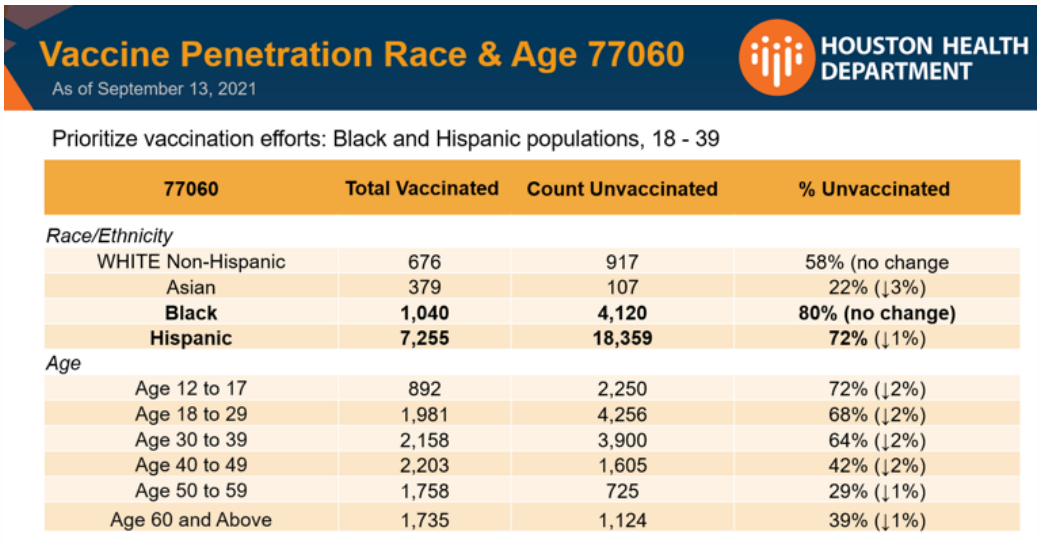


Figure 20. Priority ZIP code summary data

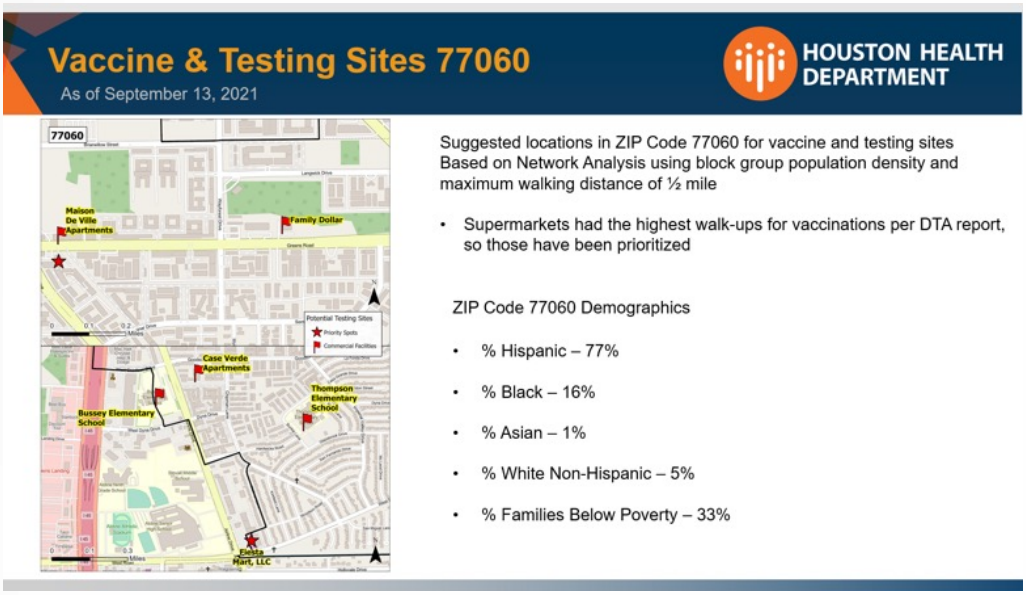
Information about vaccination penetration by race and age in each priority ZIP code is used to determine whether a particular segment of the population needs to be targeted for vaccine education and outreach (see Figure 21).

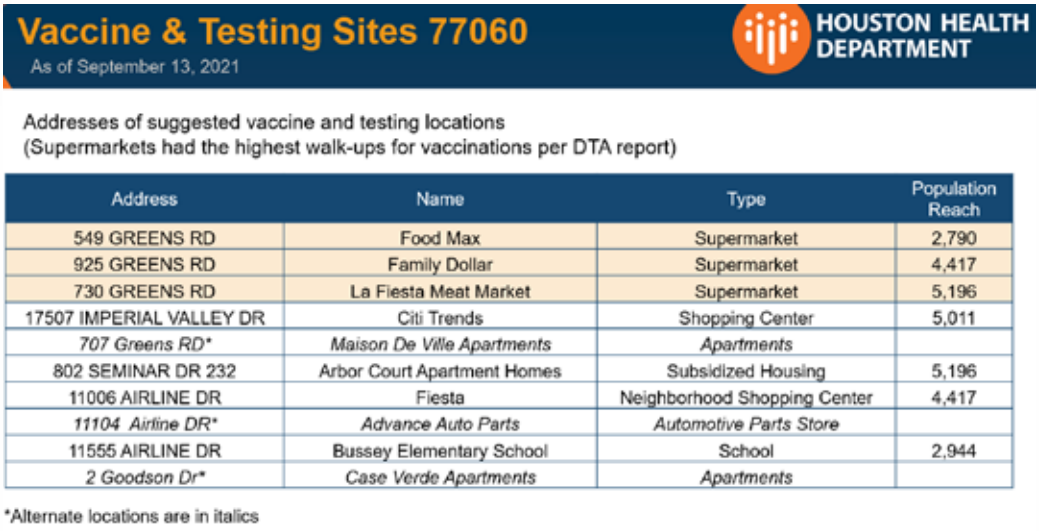


77060	Total Vaccinated	Count Unvaccinated	% Unvaccinated
Race/Ethnicity			
WHITE Non-Hispanic	676	917	58% (no change
Asian	379	107	22% (↓3%)
Black	**1,040**	**4,120**	**80% (no change)**
Hispanic	**7,255**	**18,359**	**72%** (↓1%)
Age			
Age 12 to 17	892	2,250	72% (↓2%)
Age 18 to 29	1,981	4,256	68% (↓2%)
Age 30 to 39	2,158	3,900	64% (↓2%)
Age 40 to 49	2,203	1,605	42% (↓2%)
Age 50 to 59	1,758	725	29% (↓1%)
Age 60 and Above	1,735	1,124	39% (↓1%)

Figure 21. Vaccine penetration by race and age

Finally, vaccine and testing sites, identified using a population density network analysis, are proposed for each priority ZIP code (see Figure 22).





Address	Name	Type	Population Reach
549 GREENS RD	Food Max	Supermarket	2,790
925 GREENS RD	Family Dollar	Supermarket	4,417
730 GREENS RD	La Fiesta Meat Market	Supermarket	5,196
17507 IMPERIAL VALLEY DR	Citi Trends	Shopping Center	5,011
*707 Greens RD**	*Maison De Ville Apartments*	*Apartments*	
802 SEMINAR DR 232	Arbor Court Apartment Homes	Subsidized Housing	5,196
11006 AIRLINE DR	Fiesta	Neighborhood Shopping Center	4,417
*11104 Airline DR**	*Advance Auto Parts*	*Automotive Parts Store*	
11555 AIRLINE DR	Bussey Elementary School	School	2,944
*2 Goodson Dr**	*Case Verde Apartments*	*Apartments*	

*Alternate locations are in italics

Figure 22. Proposed vaccine and testing sites

3.3.2. Manhole Level Interventions

Data indicating the detection of positive viral loads in specific manholes is sent to outbreak response teams for schools, nursing homes, congregate living facilities, and jails. Upon receipt of an alert that there has been a spike or increase in virus concentration in a wastewater sample from a specific facility's manhole, the appropriate outbreak response team contacts the facility's director and requests information to determine if a new, active outbreak has occurred. If there is sufficient data to confirm an active outbreak, an epidemiologist or subject matter expert conducts an onsite assessment and recommends a plan of action, which could include testing, lockdown of the facility, or isolation of infected individuals living in congregate settings.

For schools with positive manhole results, the appropriate outbreak team assesses the school's past and current COVID-19 status and works with the school to coordinate contact tracing and testing and to provide isolation and quarantine letters as needed. The outbreak response team maintains daily contact with the school's point of contact to track newly reported cases for a minimum of ten days after receiving an alert, or until wastewater samples from the school's manhole test negative.

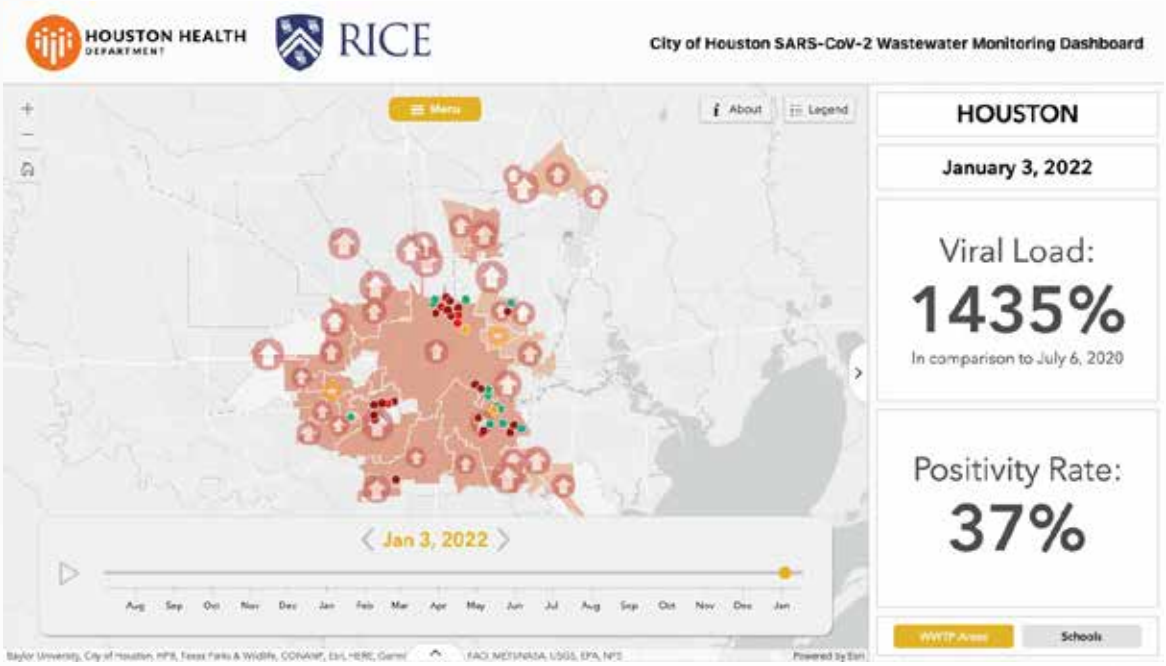


Figure 23. WWTP data from interactive dashboard

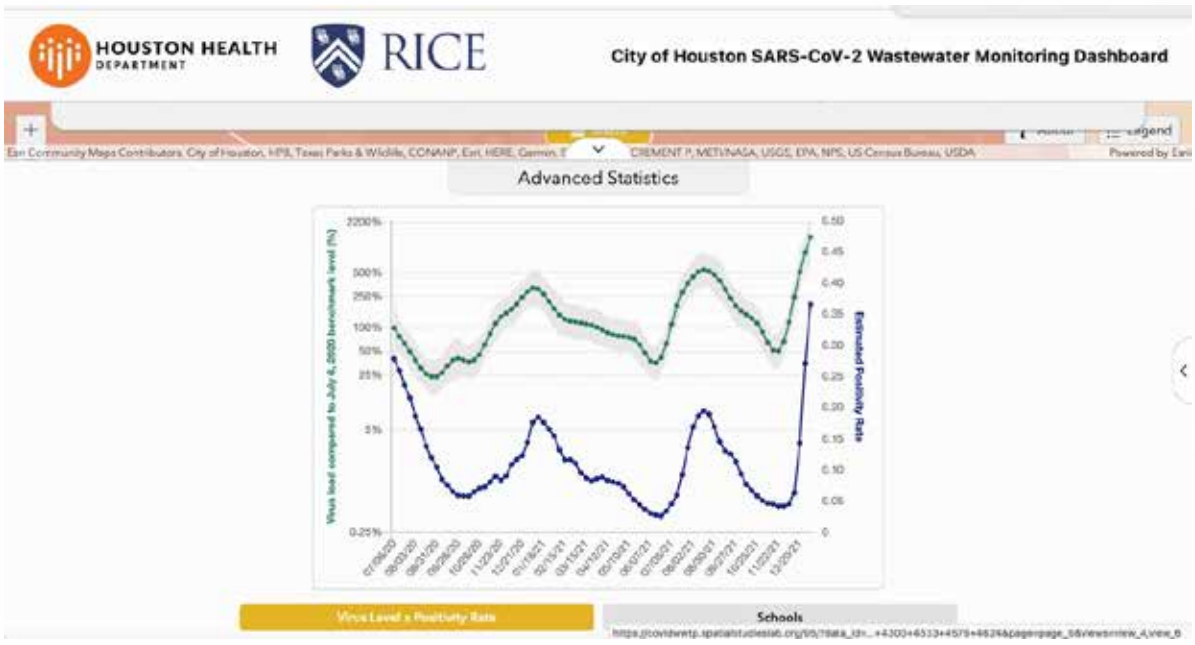


Figure 24. Historical plot of virus level against positivity rate on interactive dashboard

Schools are color-coded to indicate whether sufficient SARS-CoV-2 fragments have been detected in composite wastewater samples collected from manholes that carry wastewater only from those schools (see Figure 25). Viral loads from schools are designated as high positive, positive, negative, or inconclusive. The bottom of the dashboard expands to display historical testing data from individual schools (see Figure 26).

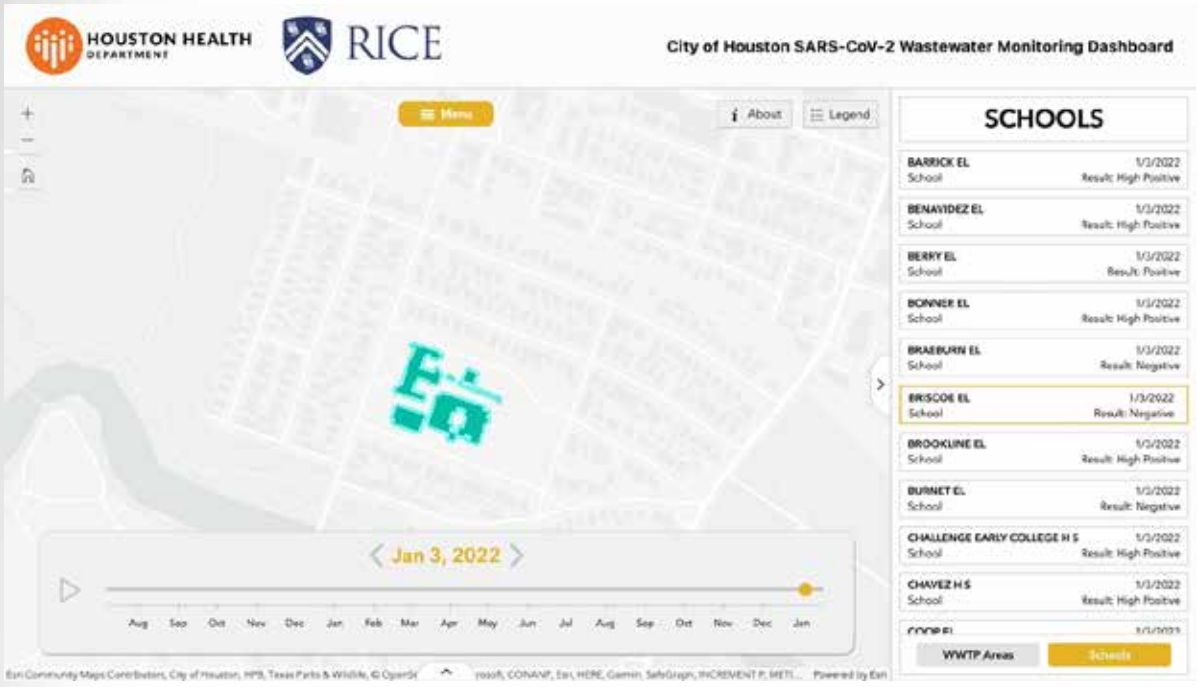


Figure 25. School data from interactive dashboard

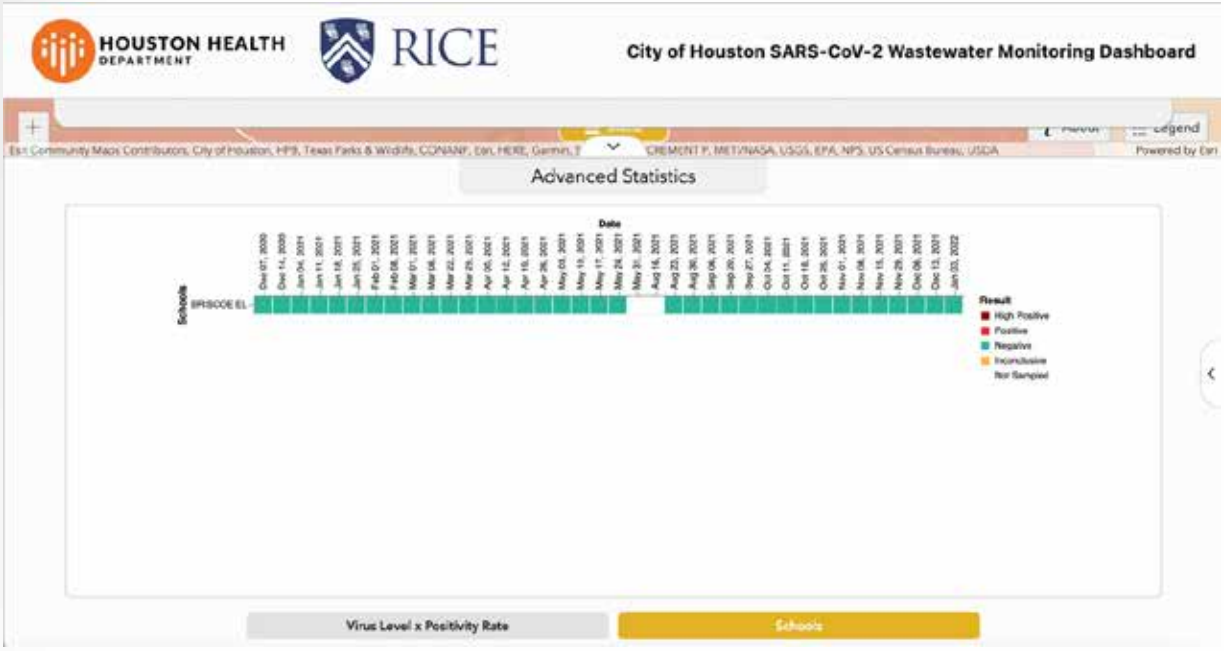


Figure 26. Historical testing data from individual school on interactive dashboard

The dashboard also includes data about the detection of COVID-19 variants in WWTPs (see Figure 27).

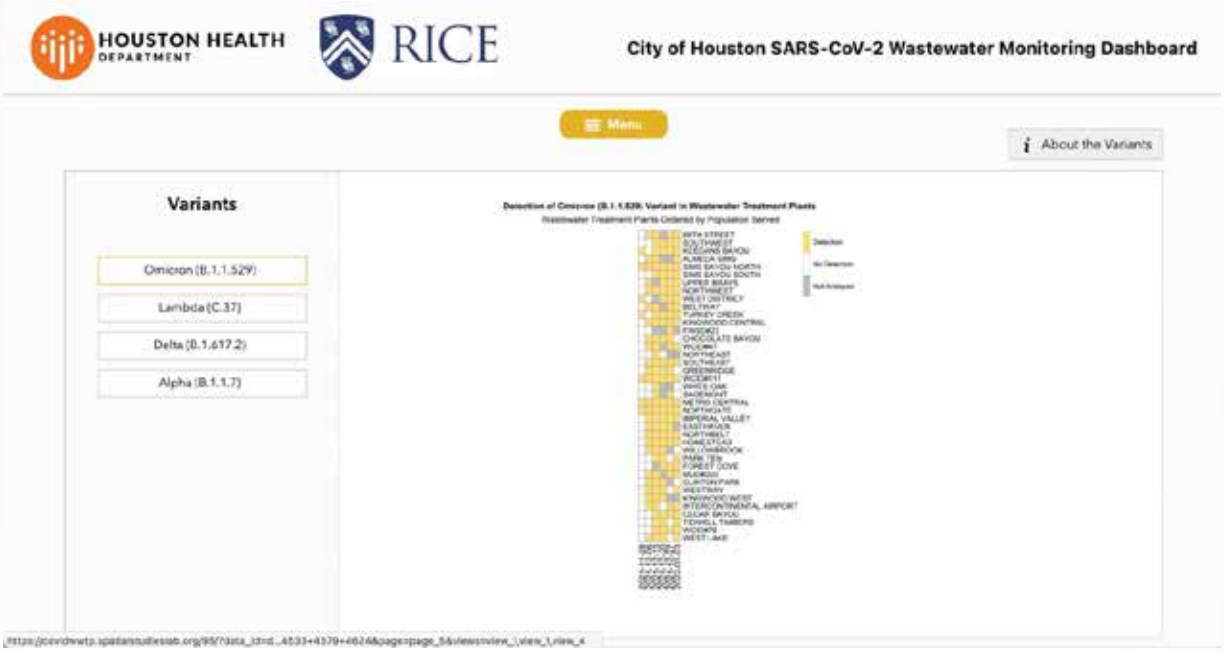


Figure 27. COVID-19 variant data on interactive dashboard

An ever-evolving source of information, the wastewater monitoring dashboard is an important component of the HHD's COVID-19 services and resources webpage.

Ethical Considerations

The benefits of WBE must be weighed against potential privacy concerns. Because wastewater samples in our surveillance system are pooled, results cannot be traced back to any single individual. The risk to privacy is therefore minimal compared to the public health benefits generated by our data.

Conclusion

WBE is an effective method for tracking community infections of SARS-CoV-2 and other viral pathogens. In the wake of the COVID-19 pandemic, it is anticipated that WBE will become a standard public health tool for monitoring infectious diseases. The recommendations contained in this document are intended to advise municipalities and other government entities on the implementation of comprehensive wastewater surveillance systems capable of monitoring viral levels in wastewater, analyzing wastewater data to generate valuable public health information, and communicating that information to the relevant authorities and the general public.

References

Betancourt, W. Q., Schmitz, B. W., Innes, G. K., Prasek, S. M., Pogreba Brown, K. M., Stark, E. R., Foster, A. R., Sprissler, R. S., Harris, D. T., Sherchan, S. P., Gerba, C. P., and Pepper, I. L. (2021). COVID-19 containment on a college campus via wastewater-based epidemiology, targeted clinical testing and an intervention. Science of the Total Environment 779. https://doi.org/10.1016/j.scitotenv.2021.146408.

"City of Houston SARS-CoV-2 Wastewater Monitoring Dashboard," Houston Health Department and Rice University's Ensor and Stadler Labs and Center for Research Computing's Spatial Studies Lab, accessed March 29, 2022, https://covidwwtp.spatialstudieslab.org/.

Gonzalez, R., Curtis, K., Bivins, A., Bibby, K., Weir, M. H., Yetka, K., Thompson, H., Keeling, D., Mitchell, J., and Gonzalez, D. (2020). COVID-19 surveillance in Southeastern Virginia using wastewater-based epidemiology. Water Research 186. https://doi.org/10.1016/j.watres.2020.116296.

"Interim Laboratory Biosafety Guidelines for Handling and Processing Specimens Associated with Coronavirus Disease 2019 (COVID-19)," Centers for Disease Control and Prevention, updated December 13, 2021, https://www.cdc.gov/coronavirus/2019-ncov/lab/lab-biosafety-guidelines.html.

Kokkoris, V., Vukicevich, E., Richards, A., Thomsen, C. and Hart, M. M. (2021). Challenges using droplet digital PCR for environmental samples. Applied Microbiology 1(1): 74-88. https://doi.org/10.3390/applmicrobiol1010007.

LaTurner, Z. W., Zong, D. M., Kalvapalle, P., Reyes Gamas, K., Terwilliger, A., Crosby, T., Ali, P., Avadhanula, V., Hernandez Santos, H., Weesner., K., Hopkins, L., Piedra, P. A., Maresso, A. W., and Stadler, L. B. (2021). Evaluating recovery, cost, and throughput of different concentration methods for SARS-CoV-2 wastewater-based epidemiology. Water Research 197. https://doi.org/10.1016/j.watres.2021.117043.

Lou, E., Sapoval, N., McCall, C., Bauhs, L., Carlson-Stadler, R., Kalvapalle, P, Lai, Y., Palmer, K., Penn, R., Rich, W., Wolken, M., Ensor, K. B., Hopkins, L., Treangen, T. J., and Stadler, L. B.

(2022). Direct comparison of RT-ddPCR and targeted amplicon sequencing for SARS-CoV-2 monitoring in wastewater. SSRN. https://dx.doi.org/10.2139/ssrn.4022373.

McCall, C., Fang, Z. N., Li, D., Czubai, A. J., Juan, A., LaTurner, Z. W., Ensor, K. B., Hopkins, L., Bedient, P. B., and Stadler, L. B. (2022). Modeling SARS-CoV-2 RNA degradation in small and large sewersheds. Environmental Science: Water Research and Technology 8: 290-300. https://doi.org/10.1039/D1EW00717C.

Rice University's Stadler Lab. (2021). Complete SARS-CoV-2 WBE Analysis Protocol. Retrieved from https://bf9f5227-500a-4928-a57e-69b0d8ada2fc.filesusr.com/ugd/4fb34c_a623c89f81364debaa29bcfe84746d06.pdf?index=true.

Sapoval, N., Lou, E., Hopkins, L., Ensor, K. B., Schneider, R., Treangen, T. J., and Stadler, L.B. (2021). Enhanced detection of recently emerged SARS-CoV-2 variants of concern in wastewater. medRxiv. https://doi.org/10.1101/2021.09.08.21263279.

Stadler, L. B., Ensor, K. B., Clark, J. R., Kalvapalle, P., LaTurner, Z. W., Mojica, L., Terwilliger, A., Zhuo, Y., Ali, P., Avadhanula, V., Bertolusso, R., Crosby, T., Hernandez, H., Hollstein, M., Weesner, K., Zong, D. M., Persse, D., Piedra, P. A., Maresso, A. W., and Hopkins, L. (2020). Wastewater analysis of SARS-CoV-2 as a predictive metric of positivity rate for a major metropolis. medRxiv. https://doi.org/10.1101/2020.11.04.20226191