SEWAGE SURVEILLANCE TOOL FOR TRACKING OF SARS-COV-2 IN URBAN BANGLADESH

A Research Paper submitted to the Department of Biomedical Engineering
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Biomedical Engineering

Lauren Hughlett, Claire Reagan, and Erin Wettstone

By

Claire Reagan

May 7, 2021

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISOR
Shannon Barker, Department of Biomedical Engineering

Sewage Surveillance Tool for Tracking of SARS-CoV-2 in Urban Bangladesh

Lauren Hughlett[a,1], Claire Reagan[a,2], Erin Wettstone[a,3], Mami Taniuchi[b], Isobel Blake[c]

[a] Biomedical Engineering Undergraduate at the University of Virginia
[b] Department of Medicine, Division of Infectious Diseases and International Health, at the University of Virginia
[c] Department of Infectious Disease Epidemiology, School of Public Health, at Imperial College London
[1] lh4aw@virginia.edu
[2] cr6ye@virginia.edu
[3] egw6vx@virginia.edu

Abstract

The coronavirus disease 2019 (COVID-19) emerged in China in December of 2019 and spread quickly, resulting in a global pandemic accompanied by economic, social, and political disruptions worldwide. Dhaka, Bangladesh is a densely populated city that hosts 21 million citizens, making COVID-19 difficult to track due to the effects of high transmission rates and asymptomatic infections. Wastewater-based epidemiology, also known as sewage surveillance, has been used in the past in low- and middle-income countries to track the spread of enteric diseases, such as poliovirus. The SARS-CoV-2 virus that causes COVID-19 has been found to shed in feces, making sewage surveillance a viable option for widespread passive testing in Dhaka. For our capstone project, we created a COVID-19 dashboard to track the spatiotemporal prevalence of COVID-19 in eight wards in Dhaka using sewage viral load and reported clinical case data. The dashboard provides mapping of sewage viral load and case data on a weekly basis, times series graphs of both sets of data, and an interpretation table to help public health officials determine the current COVID-19 situation in the study wards. The goal of our dashboard website is to provide actionable data to aid the Prime Minister's technical advisory and public health officials in Bangladesh in making actionable decisions to stop the spread of COVID-19 in the city, but the dashboard also has future implications for using wastewater-based epidemiology to track other diseases in low- and middle-income countries.

Keywords: SARS-CoV-2, Sewage Surveillance, COVID-19, Wastewater-based Epidemiology, Dhaka

---

Introduction

The context of our project is the global coronavirus disease 2019 (COVID-19) pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection.[1,2] First detected in China in the winter of 2019, COVID-19 quickly became cause for global concern. One great challenge of containing this virus has been how a person infected with SARS-CoV-2 can be asymptomatic while still unknowingly being able to spread COVID-19, and transmission of COVID-19 is very high between people.[1,3] Without widespread availability of the COVID-19 vaccine, clinical testing has proven to be the best way to track disease incidence and contain the spread of COVID-19. Widespread administration of COVID-19 tests, however, is not feasible given the lack of testing resources in low- to middle-income countries.

Dhaka, the capital of Bangladesh, is one of the most densely populated cities in the world, with approximately 21 million people.[4] Like many developing countries, the medical infrastructure in Dhaka is ill-equipped for such a large population, and hospitals are particularly underprepared for the patient influx caused by the SARS-CoV-2 pandemic.[5–7] The effects of asymptomatic infections, lack of testing, and weak access to healthcare make it difficult to determine the impact and extent of the COVID-19 disease.[8]

Since current methods of SARS-CoV-2 clinical surveillance in Bangladesh have not met the requirements of their high population density, as is possible in countries with greater health resources, a more cost effective solution to tracking the spatiotemporal prevalence of the virus is required.[5] Without a way to track the virus, it is difficult to quickly prioritize the limited resources, especially to those who need them most. High rates of COVID-19 cases have occurred in Bangladesh because of the inability to track the spread of the virus, overpopulation, and not having a means to allocate testing and other health resources to citizens.[5]

Fortunately, previous research has utilized sewage sampling to track other vaccine preventable diseases, such as poliovirus, norovirus, and rotavirus.[9] Recent studies have shown that these same sewage surveillance techniques can track the SARS-CoV-2 virus as well.[8,10–14] Sewage data reveals much about the health of a community. When humans defecate, traces of the viral pathogens are excreted and enter the sewage system, gathering with sewage from all the other humans in that particular sewage catchment area.[15] Wastewater samples can be collected from different catchment areas, and then the amount of viral pathogen in the system can be quantified through the process of reverse transcription quantitative polymerase chain reaction (qRT-PCR).[16] The amount of pathogen associated with one collection site is then associated with its watershed through sewage and drainage line tracing.[16,17] The data resulting from this methodology will represent the sewage viral load prevalence per catchment area in our dashboard.

In addition to its accuracy, sewage sampling is cheaper and requires fewer health professionals for data collection in comparison to traditional prevalence tests and contact tracing. Once the sewage data is collected, the sewage viral load is put into our dashboard to provide public health officials with a visual tool for tracking the spatiotemporal prevalence of the SARS-CoV-2 virus in their community. With this knowledge, public health officials can visualize which regions are having higher instances of the virus and can allocate more of their testing resources to those areas or implement interventions such as a lockdown to prevent further transmission.

Many communities, ranging from countries to college campuses, have successfully implemented environmental surveillance to track the cases of the SARS-CoV-2 virus in their communities. A study in the Netherlands determined that there was a strong correlation between the amount of viral RNA detected by qRT-PCR and the reported COVID-19 case prevalence.[18] To determine this, researchers gathered sewage data from five cities, quantified the amount of virus presented over the course of several weeks and determined that it was

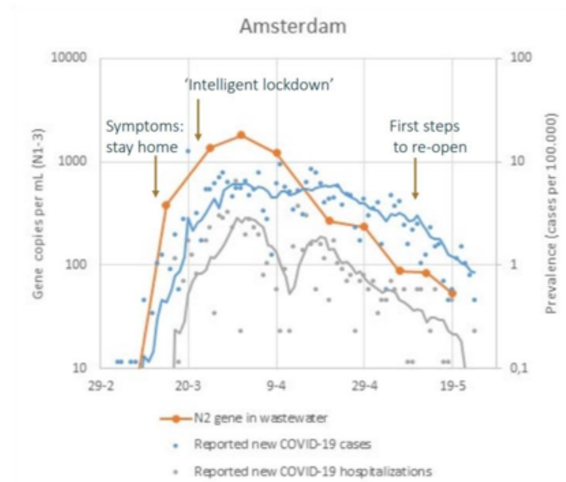


Figure 1: A graph showing how the gene copies of the SARS-CoV-2 virus detected in sewage corresponds to the case prevalence for a particular area, as well as public health reactions to this data.[18]

possible to predict outbreaks of COVID-19 cases days before cases were identified. By drawing upon these conclusions, the Dutch were able to produce a model to best contain the spread of the virus (*Figure 1*).[19] Sewage surveillance methods have also been used on college campuses in the United States to prevent the spread of the virus amongst students. Sewage data is collected from student dorms, and quarantine procedures are enforced if the concentration of viral pathogens reaches a particular threshold.[20] These methods have proven to be very effective at tracking outbreaks, giving public health officials time to react and minimizing transmission rates.

Sewage surveillance is particularly pertinent for developing countries like Bangladesh, as they do not have a strong medical infrastructure and cannot properly track or contain the SARS-CoV-2 virus with their current resources. By creating a dashboard that can visually represent the spatiotemporal prevalence of the virus, public health officials in densely populated areas will be able to identify areas of high transmission (hot spots), focus their testing resources on those areas, and implement restrictions to keep the citizens of the community as safe as possible. The largest challenges in creating this dashboard are implementing an interactive mapping system, creating accurate representations of existing data, and designing easy-to-use components that will be helpful to health officials.[17,21] Overcoming these challenges will prove that specific and distinct representation of SARS-CoV-2 is possible in heavily populated countries and could justify efforts to increase equitable representation of the disease in developing regions. By creating well documented code and methodology, the dashboard has the potential to represent spatiotemporal data for other low-resource countries and play a role in decreasing disease transmission in communities that may need it most.

Figure 2: Current Bangladesh trackers report the total number of COVID-19 cases for the country without providing spatial information like surrounding countries do.[22]

There have been about 750,000 reported cases of COVID-19 in Bangladesh, but there was previously no method for presenting distinct spatiotemporal data of the virus' prevalence in this country (*Figure 2)*.[22,23] Additionally, citizens of Bangladesh are experiencing general lack of awareness about the virus, decreased testing, and are given no daily updates about the progression of COVID-19.[6] The design of a dashboard that displays spatiotemporal prevalence of SARS-CoV-2 and other collected data obtained from environmental surveillance has the potential to spread awareness of the virus to citizens and allow increased monitoring and purposeful action by public health officials. Our COVID-19 dashboard is one of a kind in the world. There are currently no other dashboards that combine and spatiotemporally display both sewage viral load and case data for a specific country or area.

The first aim of our project was to analyze qPCR data to assess the sewage viral load of SARS-CoV-2 in Dhaka sewage lines in order to determine the spatiotemporal prevalence of COVID-19. Our second aim was to develop a dashboard tool to visualize and interpret environmental surveillance data in Bangladesh, which will provide critical information for actionable public health decisions, location-specific prevention, and tracking of SARS-CoV-2 infection. Accomplishing these specific aims will allow us to analyze and visualize the environmental surveillance data collected in the wards of Bangladesh and provide public health officials with an easily navigable means to understand the SARS-CoV-2 prevalence in the overpopulated city and act in a manner that helps to stop the spread of further viral infections.

Materials/Methods

Preliminary Plotting and Mapping in R

When planning how to accomplish our aims for our dashboard project, we determined that R software would be the best software to use to reach our goals. R software was chosen because the data analysis was being done in R by our advisor, so we decided to keep with one software system for ease of incorporating data. Our capstone team started our capstone project by completing R tutorials to familiarize ourselves with the software. Our first task to complete was to find sources that provide rainfall data for Dhaka and compare those sources. Rainfall data was obtained to determine if there is a correlation between sewage viral load and rainfall, as Dhaka City has open canals so rainfall could potentially affect the dilution of the sewage. To compare rainfall data we learned and used ggplot and other preliminary plotting methods in R that we were able to use later for our times series graphs in our dashboard.

Then, we moved into preliminary mapping methods and visualization styles. Our goal was to figure out how to best display our different shape files and coordinates in order to represent Dhaka and COVID-19 prevalence. We started investigating tmap and leaflet packages to map our study areas, and used animate to see how the shading of the areas would look over time. We also investigated different color palettes to determine the best and most understandable way to display sewage viral load data and reported clinical cases. We determined that a traffic light color scheme was the most effective color scheme to use for cases and sewage viral load, as red easily tells users that COVID-19 cases are bad or high, while green signifies low prevalence. We also concluded that leaflet packages were our best option for mapping of the study wards and catchment areas and we would create our dashboard in an R Shiny application. R Shiny was chosen because it is the most interactive dashboard that can be used and updated seamlessly through an online website.

Identifying Trends

In order to identify trends over time in our data, we plotted sewage viral load and case data versus date to determine if we could automate and identify trends in our data. To do this, we gathered data from each catchment site and plotted these values over time individually. After plotting, it became clear that the jumps in case data caused by weekly sampling was in need of data smoothing. Data was smoothed in R using a log10 average across three weeks, resulting in data that was more representative of the actual sewage viral load trends. We initially began to find correlations by analyzing the trends in data across a six week period. We found, however, that the data fluctuates on a shorter time scale than six weeks even with data smoothing. In our final dashboard, we have automated trends in sewage and case data over a three week period in our "Current Situation Tab", which allows us to identify and present information about the increasing or decreasing trends in prevalence of the SARS-CoV-2 virus.

R Shiny

Our dashboard was created using an R Shiny Web App. This program base allows the designer to incorporate different user inputs and to publish directly to the web through hosting platforms such as shinyapps.io. The user input (UI) portion of the app was constructed with a bootstrapPage layout and written in HTML to broaden our scope of visual design choices. The server logic portion of the app contains many reactive variables that respond to user input and

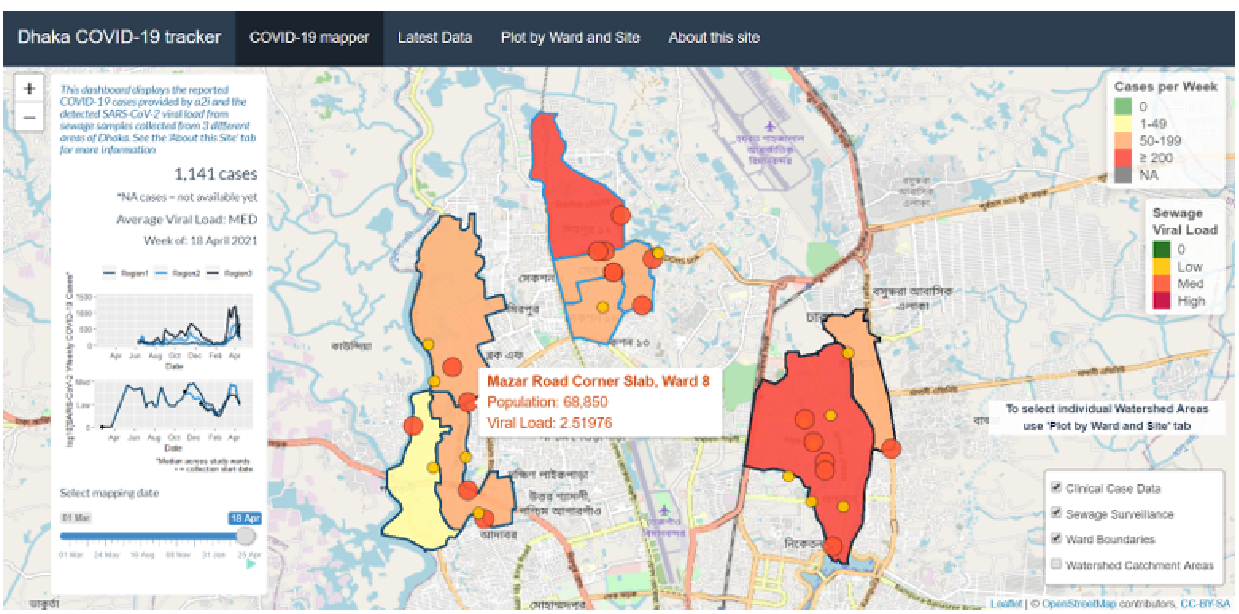


Figure 3: A snapshot of the first tab of our dashboard that spatiotemporally displays clinical case data and sewage viral load in eight wards in Dhaka.

calls functions that are written outside of the UI and server logic, to limit computational repetition. These functions are only called when reactive variables update, instead of constantly being run inside the UI or server.

Collectively, our dashboard is made up of four main tabs: 1) COVID-19 Mapper, 2) Latest Data, 3) Plot by Ward and Site and 4) About this Site. The first tab, COVID-19 Mapper is shown in *Figure 3* and displays the main map of the 8 wards and 33 sewage sites of our study in Dhaka. It also contains a panel bar, which displays the total clinical case incidence, viral load level, and date for that week, as well as linear plots of clinical case and viral load trends by region over time. Near the bottom of the panel, the user is able to drag the date slider according to the week of data they wish to view, and the map updates to show the weekly incidence of COVID-19 cases and SARS-CoV-2 sewage viral load for that week. Clinical case incidence is represented by the shading of the eight wards, and follows a traffic light scheme - green, yellow, orange and red - to mark zero to high levels of cases. Similarly, sewage viral load is represented by the corresponding traffic light scheme colors and changing radius size, depending on the zero, low, medium, or high viral load for a site that week. Legends are posted on the map to help guide the user's visual understanding of the data. Lastly, the user is able to hover over a ward or site and receive more specific information. Hovering over a sewage collection site will display the site name, population of the collection catchment area, and sewage viral load for that week. Hovering over a ward will show the ward name and number of cases that week. Lastly, the map has several layers that the user is able to overlay, including the sewage catchment areas.

The second tab, Latest Data, is shown in *Figure S1*, and provides the dashboard user with a snapshot summary of the most recent clinical case and sewage viral load data. The interpretation table reports the sewage viral load and case incidence levels for each site - either high, medium, or low. Additionally, a case forecast is provided for each site, which is based on the slope of the linear regression from the past three weeks of data. The case forecast column contains information that we hope will be the most pertinent to public health officials when making actionable decisions, because it creates a connection between the sewage trends and how they may reflect on case trends and COVID-19 incidence in the following weeks. The interpretation table is partnered with a static map that displays only the data from the most recent week.

The third tab, Plot by Ward and Site, is shown in *Figure S2 and S3*, and allows the user to view linear trends in the clinical case incidence and sewage viral load over time. The user is able to select whether they want to view the trendlines by ward or by site, and can choose which wards or sites they'd like to view from a drop-down menu. The user can also hover over a data point and view more specific information, such as the ward name, week, and viral load or case incidence for that week. When the user selects a ward from the drop-down menu, there is a map below that updates by plotting the corresponding sewage sites and catchment areas in that ward. The map provides the user with a geographical connection to the trends they're viewing.

The fourth and final tab, About this Site, is shown in *Figure S4* and contains all the background information behind the study and a brief explanation of how wastewater surveillance works. This includes descriptive sections such as Background, Sewage Surveillance, Code, Data Sources, Authors, Advisors, Special Thanks, and contact information.

Data Representation

Data for the dashboard is updated on a weekly basis and is sourced directly from organizations in Bangladesh. Every week, our advisor's research team in Dhaka conducts wastewater collection, processing, and qRT-PCR analysis of the samples. The viral load is recorded then pre-processed by taking the common logarithm of the concentration of the N1 region copies of the SARS-CoV-2 nucleocapsid gene per liter of sewage. Additionally, the clinical case data is sent from a2i, a digital public service organization supported by the Bangladesh government, every week. We then upload the new data sets to the dashboard code and republish to the web each week. Geographical coordinates for wards, collection sites, and catchment areas were provided to us by our advisors. Site name identifiers were developed by the lead research officer that oversees the project in Dhaka and then were given to us so that the data would be more relevant to Bangladeshi dashboard users.

Specific modes for data representation occur in the dashboard code as well. One major decision was to smooth the sewage viral load trends. Smoothing begins once weekly wastewater testing was established in October 2020 and we smooth the sewage data by taking a log10 three-week running average of the sewage viral load. Smoothing the data removes the noise from the weekly data and allows users to better see trends in the sewage viral load. For the sewage surveillance trends by ward, we take the average viral load across all sites in a ward, as opposed to the sum, in order to produce a fair comparison between wards.

Results

Sewage Surveillance

Wastewater is collected weekly from 33 catchment areas in Dhaka and qRT-PCR is conducted to find the sewage viral load of SARS-CoV-2 in the collected samples. qRT-PCR amplifies the viral RNA and indicates the amount of viral pathogen in the sample.[9] The viral load of the N1 region of the SARS-CoV-2 nucleocapsid gene associated with each collection site is then geospatially connected to its watershed through sewage and drainage line tracing.[24] The sewage viral load is then displayed by site in our dashboard to track the watershed areas that are experiencing increasing levels of SARS-CoV-2 RNA and provide a visual of the spatiotemporal prevalence of COVID-19. The dashboard also displays reported clinical case data alongside the sewage viral load, both in a times series plot and spatiotemporally on an open street map, so that public health officials can compare the data sets, investigate discrepancies, and draw conclusions. The sewage surveillance data is acquired a week ahead of case data, providing health officials with an up-to-date way to monitor SARS-CoV-2 and our 'Latest Data' tab provides an expected case forecast based on the linear regression trends of the sewage viral load in each catchment area over the past three weeks. With this knowledge, public health officials can visualize which regions have a higher prevalence of COVID-19 and which regions are expecting an increase in SARS-CoV-2 infections

Figure 4: A snapshot of our website that hosts the R Shiny Dashboard and is linked to IEDCR to be accessed by the Prime Minister's technical advisory, public health officials, and the general public.

and, therefore, can allocate more of their testing resources to those areas to prevent further transmission.

Website Creation

After creating a finalized version of our dashboard, we worked with UVA's Custom Applications & Consulting Services (CACS) to determine how to host our R Shiny dashboard on a website. The goal of the website was to host our dashboard on a public domain for easy and unrestricted public access and to include our collaborators on this study. The UVA CACS team helped us draw up a test site on a UVA template webpage, which hosts our dashboard front and center with our collaborators logos all beneath the dashboard, as seen in *Figure 4*. The website with our dashboard can be found here. This webpage allows our dashboard to be linked to the IEDCR website so that public health officials and citizens can monitor COVID-19 prevalence in their communities.

Product Validation

To validate the efficacy of our dashboard in meeting its goals, we implemented walk-through usability testing. We created a form with ten questions for users to answer, with an example user form seen in *Figure S5*, requiring users to navigate through all four tabs of our dashboard. The usability test questions were designed to test for comprehensibility and ease of use of our product for a new, random user. The creators of the dashboard filled out a sheet detailing the length of time each user spent on each question and what difficulties they encountered, as seen in *Figure S6*. Our walk-through usability testing was helpful in revealing areas in the dashboard that new users found confusing. In response to usability testing, we implemented a few changes to our dashboard to make the dashboard more user friendly. We created an opaque panel on the main tab, as the transparency of the panel for the graphs proved difficult for users to recognize. We also added titles to our time series plots because we realized that users were having trouble interpreting the axis labels. Lastly, we added ward numbers to the site descriptors on the map and on the drop downs so people could better locate each catchment site. After about ten minutes of navigating the site for the first time, user comprehension and navigation of the site improved greatly. If we had more time, it would have been interesting to do a second round of usability testing with the users from the first set of testing and new users, to see how previous experience with the website and the new changes we implemented affected the use of our dashboard. Overall, usability testing brought up some interesting product changes we had not considered but were able to implement, and we determined our dashboard met its goals.

Discussion

Our dashboard is novel because it represents sewage surveillance data in addition to clinical case data and provides spatial information as to which wards and watershed areas in Dhaka have the highest prevalence of SARS-CoV-2 RNA in the wastewater. Our environmental surveillance map will guide public health officials in mitigating the spread of COVID-19 early on, since the SARS-CoV-2 RNA can be found in the sewage before clinical reports of COVID-19 cases. Therefore, being able to visualize the spatiotemporal spread of COVID-19 in Dhaka will help citizens monitor disease transmission in their communities and aid public health officials in making informed decisions. Early detection of SARS-CoV-2 in wastewater will help prevent overloading of the already overburdened medical infrastructure in Dhaka and will reveal which wards should be focused on for testing. By creating a dashboard that visually represents the spatiotemporal prevalence of COVID-19 in the eight study wards, public health officials will be able to predict outbreaks in densely populated areas, focus their testing resources on those regions, and implement restrictions to keep the citizens of the community as safe as possible.

Tracking the spatiotemporal prevalence of SARS-CoV-2 in Dhaka, Bangladesh will provide public health officials with a tool that will give them greater ease in detecting the "hot spots" for COVID-19 in their country. Being able to identify where sewage viral load is high will clue officials into areas that would benefit from more clinical testing, which is important for Bangladesh because of the scarcity of clinical testing resources. Ability to predict surges in cases before symptoms present themselves in large populations can also be critical information. With it, public health officials can make actionable decisions for their communities to minimize the spread of the SARS-CoV-2 virus.

Since Bangladesh does not have free clinical case testing, sewage surveillance can act as a mediator for determining where clinical testing is most dire. Health officials with this information can allocate testing resources to this area, eliminating the need for patients to travel to hospitals and undergo expensive testing if it is not absolutely necessary. With knowledge of where testing is required, these officials can identify where need is high and potentially lower or eliminate the cost of superfluous clinical testing by using sewage surveillance results, which does not cost anything to individual citizens.

Despite strong evidence that the dashboard will be able to provide pertinent information to those in charge of public health, limitations do exist. For instance, Dhaka has an open canal system, meaning their sewage system is vulnerable to potential contamination and dilution. Excessive rainfall or contamination from unknown sources has the potential to dilute viral load detected. We have been unable to determine a normalization factor to account for this disturbance, as the potential for dilution and effect on sewage viral load has been difficult to correlate with any measurable factors, such as rainfall and variable population size. Without being able to normalize the sewage viral load, the viral load for each catchment will not necessarily be comparable to other catchments. While this is not ideal, it does not detract from the information provided by the sewage sampling, even without normalization. Along with this, some catchment areas are outside of the wards which they represent, which could present misrepresentation issues if the catchments are receiving sewage from other wards as well. Future work could be dedicated to determining how to normalize the sewage viral load of N1 and correlate the sewage viral load to a certain number of cases in a catchment area. The ability to normalize the factors that affect sewage viral load and determine what the sewage viral load means in terms of COVID-19 case count would enable public health officials to better understand the severity of cases in each watershed.

The dashboard is also limited by the manual input of new collected data. Sewage data is only collected on a weekly basis. Since sewage surveillance is only taken on a weekly basis, then it is not as useful in its ability to detect SARS-CoV-2 earlier than clinical testing. Daily collection would be the most comprehensive method of data collection, but this is not feasible given the budget and resources of the project. Similarly, sewage surveillance is only conducted in 33 catchment areas and 8 wards, so our dashboard and the data presented are not representative of all of Dhaka, but only tell public health officials of the situation in certain areas. With more funding and resources, sewage surveillance could be expanded to additional wards in Dhaka for a more comprehensive view of COVID-19 in the city. Additionally, while case data is collected daily, there is an approximate one week lag between when case data appears in the dashboard and when the cases were reported, due to the method by which the organization that pools the data delivers it. This means that the most current data on the dashboard only shows sewage data, which limits the ability to perform real time comparisons.

In our plots to identify trends in our data, we found that the plots showed some correlation visually that would indicate in certain catchment areas there is an increasing or decreasing trend of sewage data and case data over a six week period. Despite this, there is still a lot of uncertainty surrounding this data, as the data fluctuates over a shorter time period than six weeks even with data smoothing. To minimize the effects of the fluctuations, we shortened the time period observed to the last three weeks. Though we use this information in the “Current Situation” tab of the dashboard to indicate whether cases are increasing or decreasing in a catchment site, only using three weeks of data points limits the statistical power we have to draw conclusions about the transmission pattern. Further statistical analysis will be required to understand real-time changes in transmission trends.

Regardless of the slight limitations of the dashboard, it is the first dashboard of its kind, in both high-income and low- to middle-income countries, that compares real-time sewage data to clinical case data. This method of disease surveillance can serve as a model to be used in other cities or countries in the COVID-19 pandemic. Beyond the current pandemic, this dashboard can serve as a model and proof of concept for disease surveillance of all kinds of known diseases in low-, middle-, and high- income countries alike. With this knowledge, health officials will be able to respond to increases in disease prevalence at a quicker rate than if just presented with clinical case data, thus being able to reduce the prevalence of disease. Code that has been developed for the purposes of this dashboard can be cleaned and posted publicly on GitHub to serve as a template for similar sewage surveillance dashboards to be created.

This sewage surveillance pilot study in Dhaka will hopefully prove that specific and distinct representation of SARS-CoV-2 is possible in heavily populated countries and will justify efforts to increase equitable representation of COVID-19 in developing regions. By creating well documented code and methodology, the dashboard has the potential to be extrapolated to represent spatiotemporal data for other low- and middle-income countries and play a role in decreasing disease transmission in communities that need it most.

End Matter

Author Contributions and Notes

M.T. and I.B. performed research and collected and analyzed data. L.H., E.W., and C.R. performed research, designed the dashboard, and wrote this paper. The authors declare no conflicts of interest.

Acknowledgments

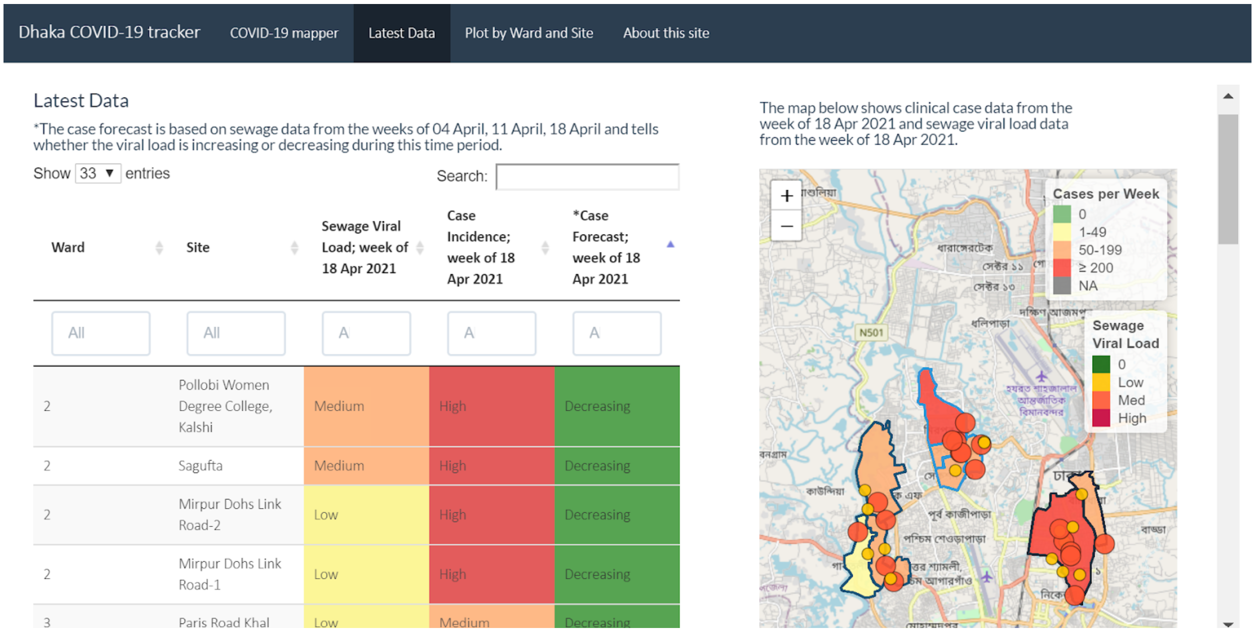
We would like to thank Mami Taniuchi and Isobel Blake for their advice and technical support throughout this whole study and project. We would also like to thank the research team in Dhaka and everyone involved in the BSESS project for their correspondence and allowing us to show their data, with a special thank you to Md Ohedul Islam, Rashidul Haque, Elizabeth Rogawski McQuade, and Stephanie Brennhofer for their work on this study. Thank you also to the icddr,b and IEDCR for sharing our dashboard, a2i in Bangladesh for data information, and the UVA CACS team for website development.

References

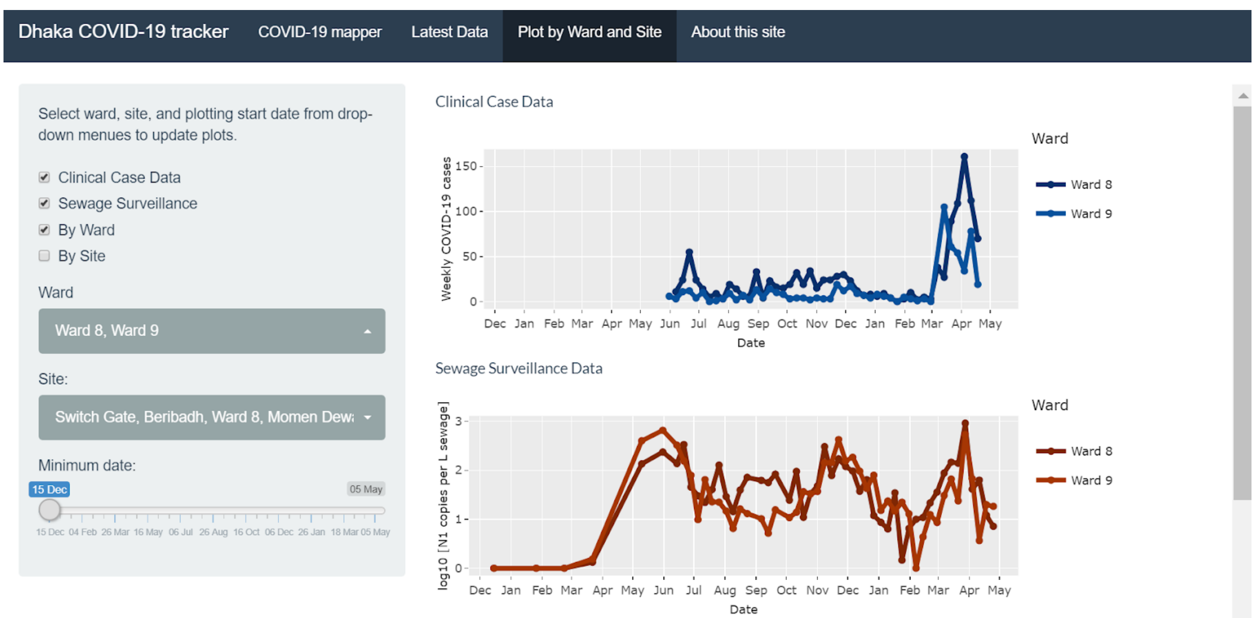
1. Case, J. B. *et al.* Replication-Competent Vesicular Stomatitis Virus Vaccine Vector Protects against SARS-CoV-2-Mediated Pathogenesis in Mice. *Cell Host Microbe* **28**, 465-474.e4 (2020).
2. Novel coronavirus structure reveals targets for vaccines and treatments. *National Institutes of Health (NIH)* https://www.nih.gov/news-events/nih-research-matters/novel-coronavirus-structure-reveals-targets-vaccines-treatments (2020).
3. CDC. Coronavirus Disease 2019 (COVID-19). *Centers for Disease Control and Prevention* https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html (2020).
4. Myelle Lansat. Dhaka, Bangladesh is the most crowded city in the world: Photos - Business Insider. https://www.businessinsider.com/dhaka-bangladesh-most-crowded-city-in-the-world-2018-11 (2018).
5. Bodrud-Doza, M., Shammi, M., Bahlman, L., Islam, A. R. M. T. & Rahman, M. M. Psychosocial and Socio-Economic Crisis in Bangladesh Due to COVID-19 Pandemic: A Perception-Based Assessment. *Front. Public Health* **8**, (2020).
6. Deutsche Welle. Why Bangladeshis no longer fear the coronavirus. *Taiwan News* https://www.taiwannews.com.tw/en/news/4019864 (2020).
7. Jubayer, M. F., Kayshar, M. S., Al Emran, M., Uddin, M. N. & Alam Soeb, M. J. Response to coronavirus disease 2019: Case study of one baking industry in Dhaka, Bangladesh. *J. Agric. Food Res.* **2**, 100077 (2020).
8. Ahmed, F. *et al.* First detection of SARS-CoV-2 genetic material in the vicinity of COVID-19 isolation Centre in Bangladesh: Variation along the sewer

network. *Sci. Total Environ.* **776**, 145724 (2021).
9. Kitajima, M. *et al.* SARS-CoV-2 in wastewater: State of the knowledge and research needs. *Sci. Total Environ.* **739**, 139076 (2020).
10. Sherchan, S. P. *et al.* First detection of SARS-CoV-2 RNA in wastewater in North America: A study in Louisiana, USA. *Sci. Total Environ.* **743**, 140621 (2020).
11. Kumar, M. *et al.* First proof of the capability of wastewater surveillance for COVID-19 in India through detection of genetic material of SARS-CoV-2. *Sci. Total Environ.* **746**, 141326 (2020).
12. Trottier, J. *et al.* Post-lockdown detection of SARS-CoV-2 RNA in the wastewater of Montpellier, France. *One Health* **10**, 100157 (2020).
13. Daughton, C. G. Wastewater surveillance for population-wide Covid-19: The present and future. *Sci. Total Environ.* **736**, 139631 (2020).
14. Ahmed, W. *et al.* 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. *Sci. Total Environ.* **728**, 138764 (2020).
15. Hellmér, M. *et al.* Detection of Pathogenic Viruses in Sewage Provided Early Warnings of Hepatitis A Virus and Norovirus Outbreaks. *Appl. Environ. Microbiol.* **80**, 6771–6781 (2014).
16. Ahmed, W. *et al.* Comparison of virus concentration methods for the RT-qPCR-based recovery of murine hepatitis virus, a surrogate for SARS-CoV-2 from untreated wastewater. *Sci. Total Environ.* **739**, 139960 (2020).
17. Kitajima, M. *et al.* SARS-CoV-2 in wastewater: State of the knowledge and research needs. *Sci. Total Environ.* **739**, 139076 (2020).
18. Gertjan Medema, Leo Heijnen, Goffe Elsinga, Ronald Italiaander, & Anke Brouwer. Presence of SARS-Coronavirus-2 RNA in Sewage and Correlation with Reported COVID-19 Prevalence in the Early Stage of the Epidemic in The Netherlands | Environmental Science & Technology Letters. https://pubs.acs.org/doi/10.1021/acs.estlett.0c00357 (2020).
19. Medema, G., Heijnen, L., Elsinga, G., Italiaander, R. & Brouwer, A. Presence of SARS-Coronavirus-2 RNA in Sewage and Correlation with Reported COVID-19 Prevalence in the Early Stage of the Epidemic in The Netherlands. *Environ. Sci. Technol. Lett.* **7**, 511–516 (2020).
20. Kent, J. Big Data Dashboards to Help US Colleges Track COVID-19 Spread. *HealthITAnalytics* https://healthitanalytics.com/news/big-data-dashboards-to-help-us-colleges-track-covid-19-spread (2020).
21. Hagedorn, B. L., Gauld, J., Feasey, N. & Hu, H. Cost-effectiveness of using environmental surveillance to target the roll-out typhoid conjugate vaccine. *Vaccine* **38**, 1661–1670 (2020).
22. Microsoft Bing COVID-19 Tracker. http://bing.com/covid/local/york_maine_unitedstates?dynamicSharing=true&shtp=Facebook&shwth=900&shh=800&shtk=Y29yb25hdmlydXMgdHJhY2tlciB1cGRhdGVz&shdk=dGVzdA%3D%3D&shth=OSH.Mmq%2BwuM5WWl/TcdViNGxBA&redirect_uri=http%3A//veeraux%3A81/covid/local/unitedstates%3FdynamicSharing%3Dtrue&ref=Coronavirus&alias=websec&fdtrace=1.
23. Bangladesh Coronavirus (COVID-19) Tracker Map | AccuWeather. https://www.accuweather.com/en/bd/national/covid-19.
24. Ahmed, W. *et al.* Comparison of virus concentration methods for the RT-qPCR-based recovery of murine hepatitis virus, a surrogate for SARS-CoV-2 from untreated wastewater. *Sci. Total Environ.* **739**, 139960 (2020).

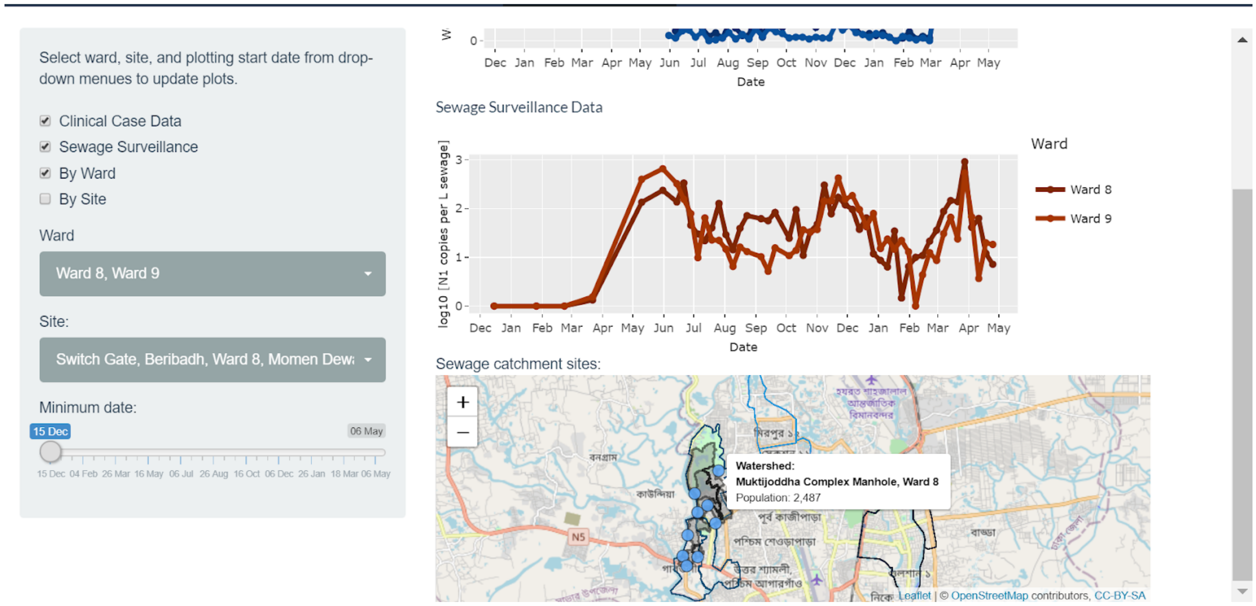
Supplemental Material



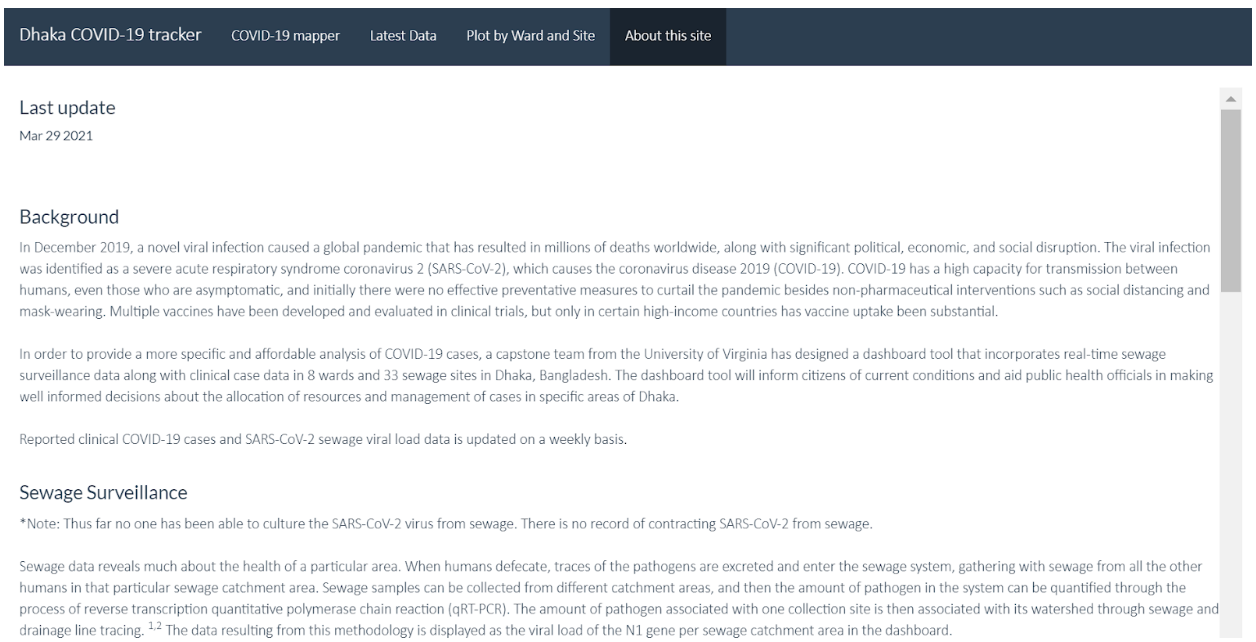
Supplemental Figure 1: The "Latest Data" tab of the dashboard, which provides a snapshot summary of the latest clinical case incidence and viral load data for health officials to use.



Supplemental Figure 2: The "Plot by Ward and Site" tab of the dashboard, which displays linear trends of the clinical case and sewage surveillance data.



Supplemental Figure 3: The mapping feature on the "Plot by Ward and Site" tab, which updates according to user input of the wards or sites that they wish to view.



Supplemental Figure 4: The "About this site" tab, which provides background information about the website and capstone project.

Walk-through Usability Testing for our COVID-19 Dashboard

Name: ___ ______

Thank you for helping us enhance the usability of our dashboard for our capstone project.
Our dashboard aims to aid public health officials in Bangladesh conduct widespread testing of its citizens and make important decisions to stop the spread of COVID-19. **Please complete the following questions in order.**

1. What is the Viral Load at the Gulshan Circle-1 Manhole for the week of March 9th, 2021?
   2.614635
2. What was the confirmed number of clinical cases in Ward 10 for the week of September 8th, 2020?
   13
3. What city is this COVID-19 prevalence data in?
   Dhaka, Bangladesh
4. How many total clinical cases were reported in the study areas for the week of January 19th, 2021?
   243
5. What is the population of the catchment area for the Nobaberbagh Bus Stand Manhole?
   12,747
6. On what date is the viral load of the catchment of Momen Dewan Bosti at its absolute maximum?
   2020-08-25
7. How many weekly COVID-19 cases were reported in Ward 18 on the date that it reached its absolute maximum?
   178
8. On what date did sewage surveillance start for the wards in Region 3? Dec. 22, 2020
9. \# of wards monitored? 8 \# of sewage catchment sites surveyed? 33
10. What is your understanding of the current COVID-19 situation in Dhaka?
   - seems like the worse they ever been
   the COVID situation is still faring bad ranging moderate-high viral load & the upper end of cases per week

On a scale of 1-10, how easy was the site to navigate and understand for you? 7

Comment: What would have made the site easier to navigate and understand?
- being able to type in site name on latest data tab
- little hard to see the side panel on covid-19 mapper / took a little bit to figure out what data was on it (small font)

Supplemental Figure 5: User sheet filled out by a random user for Usability Testing.

Walk-through Usability Testing Observer Sheet

https://usabilitygeek.com/an-introduction-to-website-usability-testing/ (site used to aid in usability test design)

User's Name completing testing: ______________________

Date of usability test: 4/17/21

* thought 3rd tab was helpful, but didn't understand log10 axis

total time: 19:03

Time it took the user to complete each question

1. 1:12 went back (2:15 skipped) 2. 42 sec 3. 18 sec 4. 2:10 5. 2:15

6. 1:39 7. 37 sec 8. 1:14 went back (1:34 skip; ↑ gave hint on where) 9. 23 sec (18 sec) 10. 4:22

Additional Notes on any Problems the User Had:

Question 1 - skipped finding viral load - on wrong tab

Q4. struggled to figure out COVID mapper date tab add instructions? couldn't read well

* might want instructions for the legends

Q5. confused if watershed population was same as catchment site population

• Didn't know which graph was viral load - label as viral load graph instead of VLN1

- skipped Q8. couldn't find regions. but figured out Q1. while looking

User's Name completing testing: ______________________

Date of usability test: 4/18/21

total time: 15:32

Time it took the user to complete each question

1. 3:30 wrong right (1:06 but) 2. 1:10 wrong 3. 2:01 4. 3:52 5. 1:29

6. 0:56 7. 0:33 8. 1:43 × 9. 1:33 (42 sec) 10. didn't really answer

Additional Notes on any Problems the User Had:

Q1. didn't change date - box to tell to move slider? need more instructions for first tab panel

Figured out Q1 when doing Q4 - figured out date

Note: did not try and click on map in Times Series tab

Q8. Found date using times series, but then found on about this site tab

Notes: Header to explain differences between time series "Clinical Case Data" vs "Sewage Surveillance"

Explain Interpretation table

Supplemental Figure 6: Example observer sheet with notes on two separate random users.