Journal of the American Medical Informatics Association, 29(5), 2022, 958–963
https://doi.org/10.1093/jamia/ocac025
Advance Access Publication Date: 23 February 2022
Case report

Case report

A real-time COVID-19 surveillance dashboard to support epidemic response in Connecticut: lessons from an academic-health department partnership

Abhinav Suri [1], Melanie Askari [1], Jennifer Calder [1,2], Charles Branas [1], and Andrew Rundle [1]

[1]Department of Epidemiology, Columbia Mailman School of Public Health, New York City, New York, USA, and [2]Formerly The Stamford, Connecticut Department of Health, Stamford, Connecticut, USA

Corresponding Author: Abhinav Suri, BA, MPH, 885 Tiverton Dr, Los Angeles, CA, USA; as6182@cumc.columbia.edu

Received 10 November 2021; Revised 27 December 2021; Editorial Decision 7 February 2022; Accepted 9 February 2022

ABSTRACT

In response to the coronavirus disease-19 (COVID-19) pandemic, numerous institutions published COVID-19 dashboards for reporting epidemiological statistics at the county, state, or national level. However, statistics for smaller cities were often not reported, requiring these areas to develop their own data processing pipelines. For under-resourced departments of health, the development of these pipelines was challenging, leading them to rely on nonspecific and often delayed infection statistics during the pandemic. To avoid this issue, the Stamford, Connecticut Department of Health (SDH) contracted with the Columbia Mailman School of Public Health to develop an online dashboard that displays real-time case, death, test, vaccination, hospitalization, and forecast data for their city, allowing SDH to monitor trends for specific demographic and geographic groups. Insights from the dashboard allowed SDH to initiate timely and targeted testing/vaccination campaigns. The dashboard is widely used and highlights the benefit of public-academic partnerships in public health, especially during the COVID-19 pandemic.

Key words: epidemiological dashboards, epidemiological surveillance, public health informatics, public academic partnerships, COVID-19

INTRODUCTION

From the beginning of the coronavirus disease-19 (COVID-19) pandemic, it was evident that real-time data would be critical to informing an optimal public health response. State governments created COVID-19 dashboards that showed near real-time state and county level data.[1] However, these data products often did not meet the needs of municipal health departments who were left to rely on periodic city- or town-level reports provided by their state departments of health (often on an ad-hoc basis). Here we describe the aims, process, outcomes, and lessons learned from an academic-local health department partnership between the Columbia University Mailman School of Public Health (MSPH) and the Stamford Connecticut Department of Health (SDH) to develop and launch a municipal COVID-19 dashboard that addressed the aforementioned shortcomings of city-level dashboards and supported local response to the COVID-19 pandemic.

Stamford is the third largest city in Connecticut (CT) with over 129 000 residents with significant racial and ethnic diversity.[2] The CT State Department of Public Health (CTDPH) released weekly reports of COVID-19 case rates by town, but these reports lacked information regarding case, testing, and death rates across demographic subgroups and did not contain neighborhood-level data showing where COVID-19 outbreaks happened. Having this level of granular data analysis would be useful in Stamford's ability to target localized COVID-19 awareness campaigns and alleviate disparities in the virus's impact across various demographics/areas in the city. Thus, SDH decided to partner with MSPH to generate real-time, specific data and analysis of COVID-19 cases, hospitalizations, deaths, hospital capacity and later, vaccination rates, by sociodemographic strata, census tract, and by neighborhood for the city of Stamford. To achieve these goals, the production of an online web-based dashboard was set as the final deliverable product (ref Figure 1 for final dashboard image). The dashboard format was selected due to its prior success in the pandemic response visualization and to allow users to interact with the data and display a variety of graphs in real time from any computer/mobile device without need for additional data processing or need to run a script.[3]

MATERIALS AND METHODS

Over the course of 3 months, an MSPH programmer, faculty, and staff developed an interactive dashboard that allowed SDH employees and leadership to visualize key trends in Stamford's experience of the COVID-19 pandemic. Due to the customization requirements, the dashboard was programmed from scratch. Other alternatives considered were R-Shiny and ArcGIS; however, the ability to customize the dashboard was limited in the former and the cost associated with the latter was unsustainable long term. Ultimately, the MSPH team decided to develop a customized Python web application (with server hosting provided by Amazon Web Services under Columbia's Business Affiliate agreement). Given that the dashboard would be recording sensitive information, the Columbia University Irving Medical Center Information Technology team reviewed the

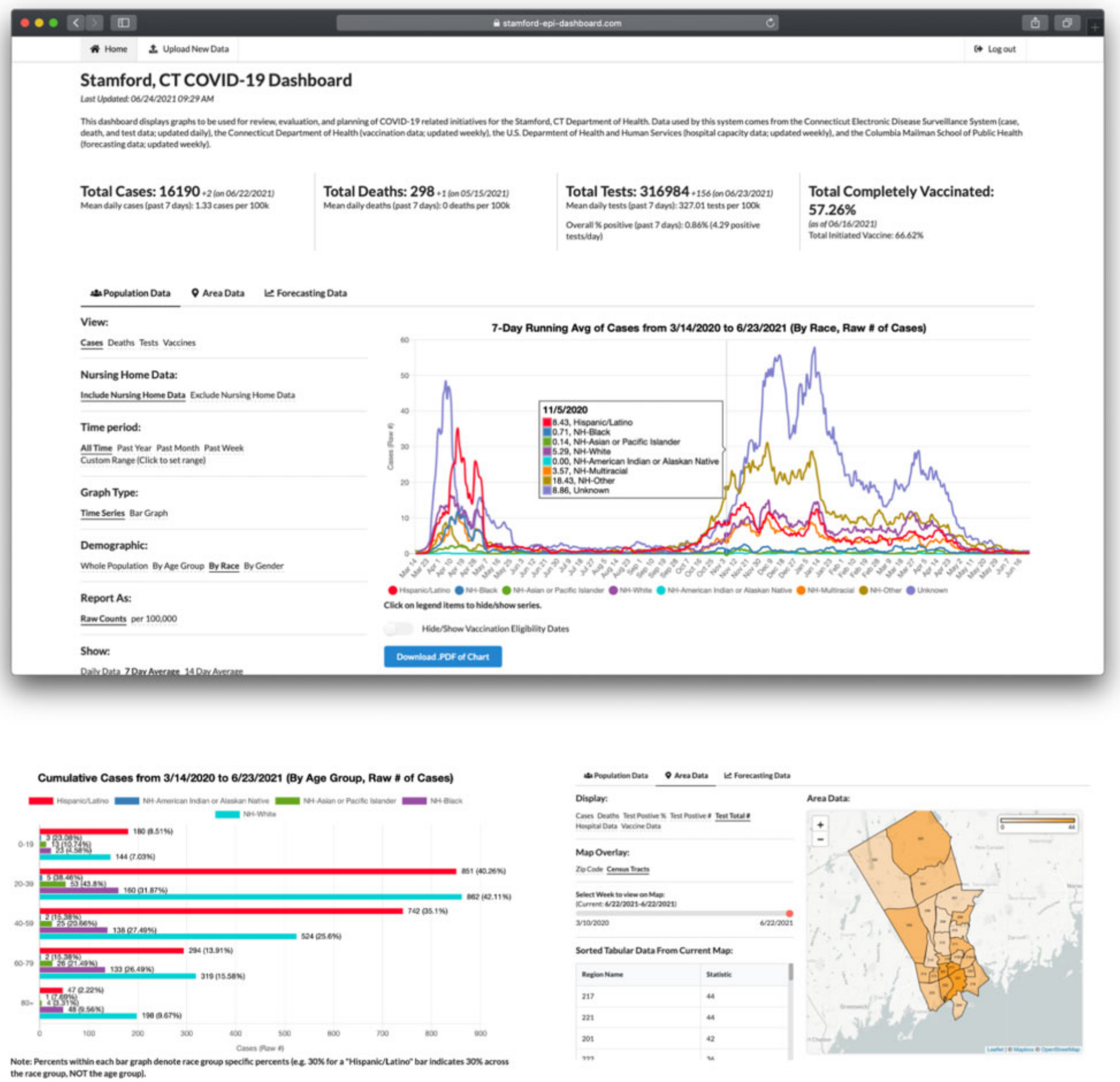

Figure 1. Snapshots from the Stamford, Connecticut Health Department COVID-19 data dashboard, 2020–2021. Top panel shows the front page of the dashboard. The bottom left panel shows an example graph generated from the dashboard (cumulative age × race case counts). The bottom right panel shows the geographic panel of the dashboard that shows (in this configuration) total # of tests by census tracts.

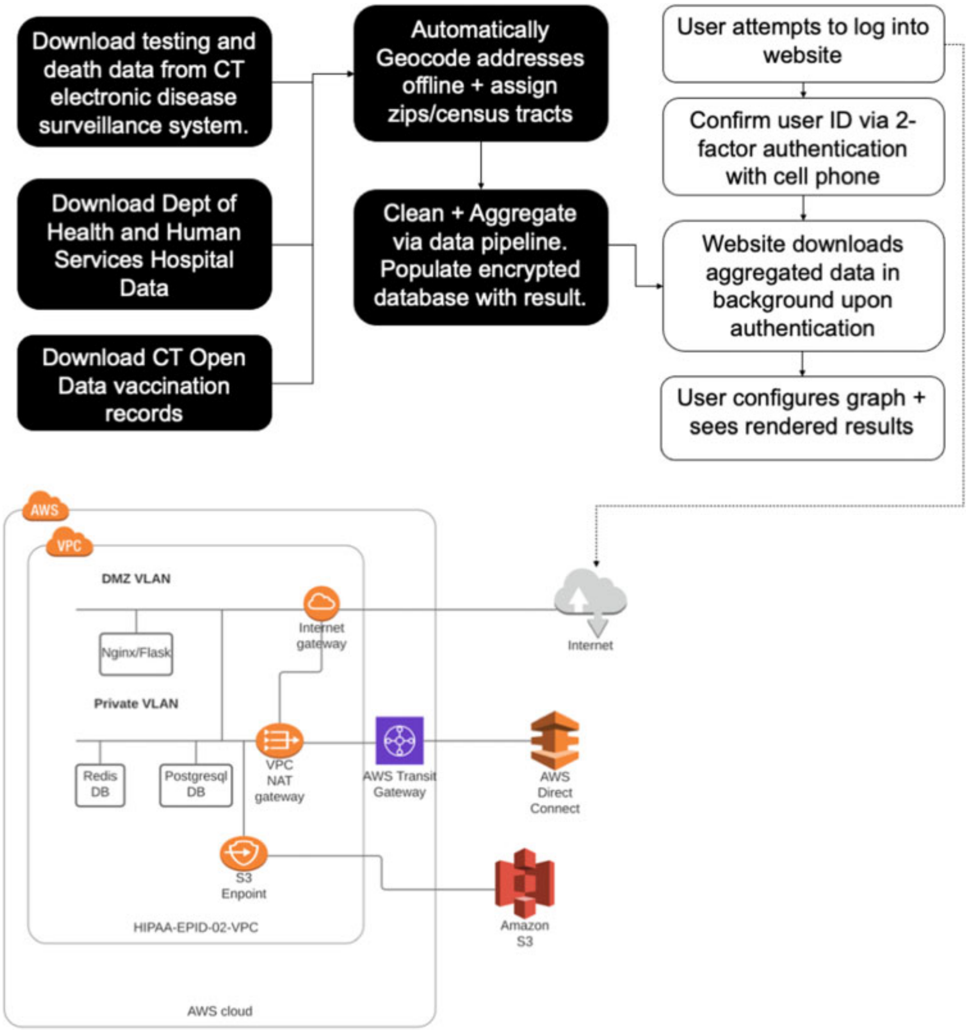

Figure 2. Outline of the dashboard data flow and server configuration. The top part of the panel shows the general flow of information (from Table 1) in cleaning data for the dashboard. Additionally, the bottom part of the dashboard shows the server configuration of the HIPAA compliant virtual environment on Amazon Web Services used to host the dashboard.

security of the dashboard and the configurations of the server and made sure that the latest encryption standards were implemented, along with 2-factor authentication to ensure that only authorized individuals had access to the dashboard (Figure 2).

Technical architecture

The backend of the application was developed using Flask, a lightweight Python server application framework. This Flask app coordinated responses to user requests to the websites and appropriately fetched data from a database (Postgres) to provide users with data. To process data without disruption to the site, a separate asynchronous job queue, handled by Redis, was created to handle processing new data in the background. The website served by the Flask application was created using HTML, CSS, and JavaScript. The JavaScript program (written in TypeScript to aid in minimizing potential errors) comprised of 20 different files that handled user interaction, graph, and map display logic and employed the usage of multiple lightweight libraries to aid in fast rendering times (so the users could rapidly switch between graph views). From a programmatic standpoint, this architecture was chosen due to its ability to be highly adaptable to changes in data formats and addition of new data sources. For example, when vaccination data became available, only 5 files needed to be altered to account for the new data source. From a user-interaction standpoint, the architecture enabled fast responses and easier customization of graphs without the user needing programming knowledge (thus enabling users with any level of epidemiological experience to interact with the dashboard).

Sustainability

To ensure sustainability of the programming process, the team adopted a scrum methodology to continuously implement features as differing needs for SDH arised.[4] Over 2–3-week periods (occurring consecutively in the first 3 months postdeployment), several "sprints" (ie, development periods for high-priority features) were undertaken.[4] Features were released at the end of sprints and feedback from SDH leadership was gathered (usually within hours of deployment) to fix errors and fine-tune the dashboard. Most features' requests from SDH involved fine-tuning the display/user interaction aspect of the graphs in the dashboard. High-priority features requests focused on the integration of new data sources (such as vaccination data once that was released to the SDH staff) and were integrated into the current or following sprint period. Any errors that were prioritized for immediate fixing were handled within 2–3 h of notice (average time to fix = 2.34 ± 1.45 h, $n = 11$). These errors primarily stemmed from errors in test/case/death data sources (eg, 1 error came from a lab inputting a result as occurring on January 1, 1920 instead of January 1, 2020). As errors came to the attention of the developer, programmatic testing of features was employed to minimize the presence of errors caused by erroneous input. Critical to this entire process was the input from users. To solicit constant input from users of this dashboard, biweekly meetings were set up with SDH leadership and all other users to gather input on the design and functionality of the dashboard in its planning and postdeployment stage.

Data sources

To populate the dashboard, COVID-19 testing, current case, and death data were downloaded daily from the CT Electronic Disease Surveillance System (CTEDSS) for reportable diseases and were uploaded to the dashboard. An individual was designated as a case in CTEDSS on the date they first test positive (individual only considered reinfection if a positive test occurred at least 90 days after the first positive test). CTEDSS was originally designed to inventory disease cases, facilitate contact tracing, and record case interactions with the healthcare system. CTEDSS was not designed to calculate key epidemiological statistics required for tracking the course of the epidemic in the population. Accordingly, a new data pipeline was created on the server to clean and process the uploaded CTEDSS data into configuration usable for epidemiological statistics (critically this pipeline reconciled any duplicate cases across multiple days). Data were also integrated from CT Open Data (vaccination statistics), the Department of Health and Human Services (ICU/hospital capacity data), and from the MSPH COVID-19 Projections team (Table 1).[5]

If any errors were encountered during the automated processing of data (eg, due to new output data formats from any of the sources), the data cleaning processing job would abort, the database would revert to its state before the addition of the data, and an error message would be displayed to the user outlining the reason for the failure (along with the full error stack trace for developer usage). Error messages would also be sent to the developer of the application. This approach ensured that unexpected data format changes would not affect the state of the dashboard and could be handled in a timely fashion. Additionally, encrypted database backups were automatically created every 24 h via AWS Relational Database System (RDS) to allow developers to revert to an earlier dashboard state if necessary. Figure 2 shows the overall organization of the data processing pipeline.

User interface

The users of this dashboard included members from the SDH pandemic response team as well as other city administrators. Among the people who had access to this dashboard were as follows: the head/commissioner of the SDH, the medical advisor to SDH, public safety

Table 1. Overview of Stamford, Connecticut Department of Health COVID-19 dashboard data sources, 2020–2021

Data source	Data type	Description
Connecticut electronic disease surveillance system (CTEDSS)	Tests, cases, and deaths	CTEDSS is an internal tool that provided information on the tests, case, and death data via daily reports from the Connecticut Department of Public Health. Test reports provided information regarding individuals first date of a positive test (also considered the date they were marked as a case) along with the race, address (along with whether said address was a long-term care facility), birth date, and gender of the individual test. Death data contained information about the race, zip code, birth date, and date of death of the individual affected.
CT Open Data Portal[7–9]	Vaccinations	The CT Open Data Portal provided weekly aggregate reports with town-specific vaccination rates divided into % initial dose administered and % completely vaccinated starting in April 2021. Information about vaccines stratified by age group and census tract designation were also provided.
Department of Health and Human Services[10]	Hospital Data	The Department of Health and Human Services (DHHS) provided information regarding ICU capacity for hospitals across the United States including the Fairfield County area (where Stamford is located). From the descriptors DHHS reported, we output % COVID cases in hospital, % ICU occupied, and % COVID patients in ICU to give SDH officials a better idea as to the number of acute active cases in the area.
Columbia University Mailman School of Public Health[11]	Forecasting Data	MSPH provided weekly forecasts (projecting 3 weeks) of cases in the Fairfield County area. Daily predictions were accompanied by a measure of confidence with upper and lower bounds on the predictions for number of cases.

Notes: This table provides an overview of all the data sources used to plot graphs and make maps for usage by SDH employees and leadership. Additionally, information about the type of each dataset ("Data Type" column) and characteristics of each data source ("Description" column) are also described.

data analysts for Stamford, the mayor & mayoral office, and several interns/project managers in those offices. As a result, the dashboard was designed to accommodate the differing levels of epidemiological experience present in the user base. We designed a highly flexible user interface that to allowed users to easily graph and map the data according to a number of customizable characteristics. The dashboard allowed users to generate 150 possible graph and map representations of the pandemic (Figure 1).

Graphs could be configured to view cases, deaths, tests (including or excluding nursing homes), vaccination rates (first dose or completed all doses), hospitalization data, and forecasting data. Additionally, case/death/test data could be stratified by age, sex, race/ethnicity, or census tract/zip. All data could be viewed as counts, rates, and age standardized rates across user specified date ranges. The number of different stratification variables proved useful to the SDH users since they could use those variables to visualize disparities in pandemic response across the city by race and area. Additionally, graphing parameter defaults for cases, death, test, and hospitalization graphs were set to allow less epidemiologically experienced individuals (ie, mayoral office members) to quickly get an accurate picture of pandemic spread. These simpler graphs displayed 7-day running averages for the aforementioned characteristics over time to allow less experienced users to get a sense rough trend in the pandemic progression. The full list of variables that could be graphed is as follows:

- View cases, deaths, tests including or excluding nursing home data. Also view vaccination data (first dose, completed all doses), hospitalization data (% COVID cases in hospital [# COVID positive cases/total individuals in hospital that week], % ICU occupied [# of ICU beds occupied/# of ICU beds available], % COVID patients in ICU [# of ICU beds occupied by COVID positive cases/# of ICU beds occupied]), and forecasting data.
- View all of the above over time or aggregated within preset or user customizable date ranges.
- View cases, death, and test data by age group, race, sex, or age by race.
- Display raw counts, per 100 000 counts, or percent (for tests [# positive tests/# of total tests] and vaccination data [# vaccinated/total population]).
- Display age adjusted rates where applicable.
- Display daily data, 7-day rolling average, or 14-day rolling average for time trend data.
- Display cases, deaths, tests, and vaccinations by zip code or census tract and view trends over time by zip code or census tract.

RESULTS

Initial reports were generated from the dashboard starting in December 2020 and the full online dashboard was launched in January 2021. SDH employees and leadership, Stamford's Mayor and mayoral staff, public safety leadership, and interns/members of the aforementioned departments were provided secure access to the dashboard. Upon being provided access, users went through a 15-min training session with the developer of the dashboard (this session showed users how to set up 2-factor authentication to log into the dashboard?, how to navigate the dashboard?, and how to customize visualizations of graphs/geographic data?). Feedback from users regarding ease of use of the dashboard was gathered, and we found that users thought the dashboard was intuitive to use (even without prior epidemiological experience). After the dashboard was deployed, new data sources (eg, vaccination data) were integrated as that data became available. These sources were added to the dashboard in a prompt manner (within 1–2 weeks of release) without disruption to daily updates of the dashboard.

For security and evaluation purposes, MSPH recorded all interactions between users and the dashboard. As of December 26, 2021, the dashboard has been viewed 3485 times (1 view equals a user interaction lasting greater than 30 s) by SDH members ($n = 12$) and has also been visited daily by MSPH programmers, faculty, and staff. Across SDH employees/leadership specifically, an average of 1.78 h/week was spent viewing and interacting with the main dashboard. Additionally, approximately 2–3 graphs were exported from the system per week for distribution to city officials. Prior to the dashboard development, graphs stratified by race or age could only be generated on a monthly basis due to the time consuming nature of processing data manually in Excel and one-off SAS scripts. Additionally, these manual processing pipelines did not visualize geographic data (due to lack of programming experience in the department), preventing SDH from getting a sense of disparities in COVID-19 impact across census tracts.

Given the ability to view information on a more frequent basis, numerous insights could be gathered from the dashboard. For example, per the SDH leadership and epidemiologists, a key insight gained from the dashboard within the first month of launching was that there were disparities in COVID-19 infection rates by race/ethnicity that leadership were not aware of from CT state reports (which only gave information on nonstratified cases/tests/death rates). This insight allowed them to target messaging to groups identified as more susceptible to COVID-19 based on the dashboard. Additionally, the dashboard's mapping tools provided critical information on day-by-day disease spread throughout the lifetime of the dashboard. This information was not accessible by SDH prior to the development of this dashboard. The dashboard enabled the city to identify specific regions that had rapid increases in cases, allowing SDH leadership to then carry out targeted testing and vaccination campaigns focusing primarily on these census tracts.

Early in the pandemic, SDH secured CT State funding for COVID-19 surveillance and epidemiology support which was used to fund the contract with the MSPH. In addition, a MSPH Masters student worked on the project to fulfill the school's practicum hours requirement and a Doctoral student and faculty volunteered time. The dashboard will remain active through the duration of the pandemic.

DISCUSSION

The development and usage of the dashboard proved to be critical in Stamford's ability to monitor and respond to the pandemic. Given that SDH did not have the bandwidth to develop a dashboard of their own, the partnership with MSPH enabled them to get a dashboard developed that helped them in their pandemic response. MSPH was able to complete the project using 275 h of developer time and 1 h/week of postdeployment maintenance/dashboard data updates which is accomplishable within 1–1.5 months using a small number (1–3 people) of experienced programmers and 1–2 staff. Given the relatively small amount of human resources needed for the development of these dashboards, it is likely within reach for small-to-medium sized cities to undertake the task of creating an epidemiological dashboard, especially given the potential insights it could offer into an evolving pandemic.

Similarly, other public health departments can look to outsourcing epidemiological analyses and monitoring software to academic institutions, allowing the departments of health to focus on carrying out interventions based on the data generated from academic partners. Such a partnership would likely allow both local and state departments to get specific data on the progression of the pandemic across different demographics and areas. Our experience shows the utility of bilateral partnerships between local health departments and schools of public health in responding to public health emergencies. Furthermore, the collaborative and transparent nature of the academic-public partnership allowed both stakeholders to fill in knowledge gaps and select the most appropriate epidemiological parameters to track in the dashboard. The collaboration between SDH and MSPH was also facilitated by the Commissioner of Health's pre-existing appointment as an Adjunct Professor at MSPH. For instance, the Commissioner had an existing University ID account, log-on access to the computer systems, a professional and social network within the school and administrative standing, allowing for a rapid integration between the SDH and MSPH teams. This suggests that formal relationships between schools of public health and local departments of health, such as faculty positions for key Department of Health personnel, should be considered as part of disaster and emergency preparedness planning.

Additionally, the increasing enrollment of Master and Doctoral students with prior computer programming and information technology skills into schools of public health proved critical to the success of this project and to the MSPH's other COVID-19 response projects. Of particular benefit was that the student who performed the computer programming understood epidemiologic concepts, such as rates and age standardization. This experience suggests that deeper ties should be fostered within universities between schools of public health and undergraduate and graduate computer science and/or data science departments. Furthermore, creating a pipeline between computer science departments and city public health departments could provide departments of public health with experience necessary for developing technological solutions to epidemiological problems. Increasingly, public health is becoming a career path for computer science majors and the health of the public benefits from their technical skills.

Local data and publicly accessible interfaces to make these data not just useful, but also used, are still in need to best respond to pandemics and other large-scale public health crises.[6] The development of a municipal COVID-19 surveillance dashboard proved to be critical to the City of Stamford's ability to monitor and respond to the pandemic and may serve as a model for other municipalities.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

AUTHOR CONTRIBUTIONS

Design of research, data acquisition, and analysis were done by AS, MA, and AR. Study design and conception by AS, AR, JC, and CB. Revision and approval of final manuscript by AS, MA, JC, CB, and AR.

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY

Data (user-specific dashboard usage data) cannot be shared for ethical/privacy reasons.

REFERENCES

1. New Jersey COVID-19 Dashboard. 2020. https://njhealth.maps.arcgis.com/apps/MapSeries/index.html?appid=50c2c6af93364b4da9c0bf6327c04b45&folderid=e5d6362c0f1f4f9684dc650f00741b24. Accessed August 26, 2021.
2. U.S. Census Bureau QuickFacts: Stamford city, Connecticut. 2021. https://www.census.gov/quickfacts/stamfordcityconnecticut. Accessed August 24, 2021.
3. Wissel BD, Van Camp PJ, Kouril M, *et al.* An interactive online dashboard for tracking COVID-19 in U.S. counties, cities, and states in real time. *J Am Med Inform Assoc* 2020; 27(7): 1121–5.
4. Sharma S, Hasteer N. A comprehensive study on state of Scrum development. In: 2016 International Conference on Computing, Communication and Automation (ICCCA). IEEE; 2016: 867–72; Greater Noida, India.
5. Zebrowski A, Rundle A, Pei S, *et al.* A spatiotemporal tool to project hospital critical care capacity and mortality from COVID-19 in US counties. *Am J Public Health* 2021; 111(6): 1113–22.
6. Committee on Applied Research Topics for Hazard Mitigation and Resilience, Policy and Global Affairs, National Academies of Sciences, Engineering, and Medicine. *Motivating Local Climate Adaptation and Strengthening Resilience: Making Local Data Trusted, Useful, and Used*. Washington, DC: National Academies Press; 2021: 26261.
7. COVID-19 Vaccinations by Town | Connecticut Data. 2021. https://data.ct.gov/Health-and-Human-Services/COVID-19-Vaccinations-by-Town/x7by-h8k4. Accessed August 24, 2021.
8. COVID-19 Vaccinations by Town and Age Group | Connecticut Data. 2021. https://data.ct.gov/Health-and-Human-Services/COVID-19-Vaccinations-by-Town-and-Age-Group/gngw-ukpw. Accessed August 24, 2021.
9. COVID-19 Vaccinations by Census Tract | Connecticut Data. 2021. https://data.ct.gov/Health-and-Human-Services/COVID-19-Vaccinations-by-Census-Tract/fj4n-gyni/. Accessed August 24, 2021.
10. COVID-19 Reported Patient Impact and Hospital Capacity by Facility | HealthData.gov. 2020. https://healthdata.gov/Hospital/COVID-19-Reported-Patient-Impact-and-Hospital-Capa/anag-cw7u. Accessed August 24, 2021.
11. Time Machine—Interactive Graphs. 2020. https://msph.shinyapps.io/timemachine/. Accessed August 24, 2021.