Dallas County Public Health Disease Surveillance and Investigation System

Dallas County Health and Human Services

Primary Contact Information:
• Role/Title: Philip Huang, MD, MPH – Director, Dallas County Health & Human Services
• Email: Philip.Huang@dallascounty.org

Clinical Project Lead:
• Role/Title: Philip Huang, MD, MPH – Director, Dallas County Health & Human Services
• Email: Philip.Huang@dallascounty.org

IT Project Lead:
• Role/Title: Rachel Corbin – Disease Surveillance Offering Lead & Salesforce Manager, Accenture
• Email: Rachel.Corbin@accenture.com

Executive Summary

The COVID-19 pandemic exposed existing gaps in public health disease surveillance and investigation infrastructure and posed new challenges for public health systems. In Dallas, three critical needs were identified: management of increase in data volume without sacrificing data quality, ability to conduct large-scale contact tracing to understand and control spread and improve disease investigation workflows to reduce the burden placed on public health professionals. To address these needs, Dallas County Health and Human Services (DCHHS) employed trusted enterprise-level software, including Salesforce, Informatica MDM, MuleSoft, Rhapsody, and Power BI. All these technologies were brought together to craft a seamless solution with streamlined user experience. These tools are adaptable and scalable, improving Dallas County's long-term preparedness to address future threats more readily. As industry-leading technologies, they also provide Dallas County security and peace of mind when it comes to data security.

Data from the Disease Surveillance and Investigation System during the COVID-19 pandemic were critical for monitoring disease trends and were provided as inputs to inform policy decision-making about interventions to reduce transmission. Data was also used for improving neighborhood-level visibility to identify at-risk areas – including those with higher burden of COVID-19 incidence and areas with lower rates of vaccination. Consequently, this system was pivotal to effectively reducing COVID-19 transmission in Dallas County, resulting in decreased rate of hospitalization and mortality related to COVID-19. This system has also proven applicability beyond COVID-19, as it was employed

to address the recent Mpox outbreak. Similarly, the system allowed Dallas County to rapidly begin tracking cases and contacts in Dallas County, providing the ability to identify potential at-risk populations and deploy appropriate public health campaigns. The ability to quickly spin up this surveillance capability allowed officials to rapidly contain the situation.

Define the Public Health Problem

Challenges

DCHHS serves a diverse county of 2.6 million people. The volume and variety of data generated and processed during the COVID-19 pandemic put immense strains on existing public health systems, also imposing an increased manual burden on disease investigators as they conducted large-scale contact tracing. The specific challenges that the Dallas County Public Health Disease Surveillance are discussed in further detail below. Addressing these challenges in light of COVID-19 has allowed DCHHS to build a robust infrastructure which is now being leveraged for routine disease surveillance activities and has improved preparedness

- **Data modernization**: High-quality data drives high-quality public health activities. The main challenges addressed were data quality (e.g., duplicate records), lag times in data availability, lack of census block level data, and burden of manual data quality management on public health professionals.
- **Integration:** The ability to readily ingest, parse, and make data available from electronic cases, vital records, the immunization registry, and electronic reportable laboratory results. Siloed data sources also increased the complexity of generating a single source of truth.
- **Contact tracing**: During the COVID-19 pandemic, DCHHS found a lack of infrastructure and capabilities to foster impactful large-scale contact tracing. Challenges included appropriate data infrastructure, manual routine processes, and lag times in outreach to cases and contacts. The ability to train, onboard, and sustain hundreds of profiles without slowing the system was also crucial.
- **Notifiable Disease Reporting:** Providing timely reports is critical for informed, impactful public health decision making. Key challenges addressed with regards to reporting included timeliness, standardization, and streamlined workflow.

Unique components of this public health problem

These technologies were deployed amid the COVID-19 pandemic to address the immediate demands, while conducting response activities in parallel. As such, system improvements began with imminent needs – high volumes of public health data, the need to quickly and accurately identify disease trends and which populations and communities were being disproportionately affected, the need to link multiple relevant data sources, and the need to have a system that could manage hundreds of contact tracers.

A high priority for DCHHS

The COVID-19 pandemic revealed how public health data infrastructure had previously been neglected, causing DCHHS to conduct a necessary review of outdated systems and technology. The impact of these systems on data quality and timeliness was quickly evident and needed to be addressed rapidly to allow DCHHS to make important and costly public health decisions. In addition, as the pandemic continued, the impact on the workforce was evident – public health staff were overwhelmed and attrition increased, furthering the burden placed on others. These challenges exposed the impact of the system on routine disease surveillance and showed the need for improvements for future preparedness as well.

Performance Data

Prior to implementation of this system, DCHHS was using a Red Cap database system that was quickly overwhelmed by the volume of reported COVID-19 cases. At its peak, Dallas County was reporting thousands of new cases per day. Before implementation, there were also significant delays in getting the data, as lab reports were still coming in by paper fax and there were piles of paper lab reports that required manual data entry. Also, prior to implementation of the system, there was no ingestion of hospital electronic case reports, no ingestion of COVID-19 vaccination data, and no ingestion of mortality, which have subsequently all been successfully matched and merged with automated electronic data ingestion, to provide a single complete, comprehensive person-based data source for all of this relevant public health information.

Performance Goals

Broadly, the performance goals for the new system were to reduce the amount of manual effort, improve data quality and timeliness, and enhance interoperability with needed systems for automated data ingestion. In addition, to bolster epidemiological investigations and reporting, the system was developed with the latest visualization technology to optimize insight generation. Long-term utility was also a critical consideration, so selected infrastructure needed to not only meet current needs but be nimble enough to address currently unanticipated future needs.

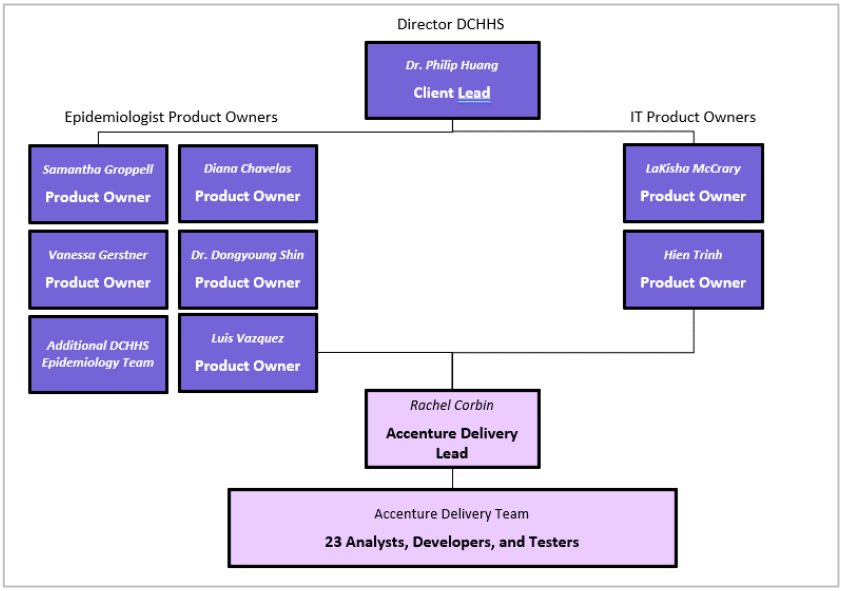
Currently, the system can process millions of data points, with the ability to expand the capacity even further. Data feeds are currently refreshed nightly with the ability to increase frequency based on the need, providing near real-time data for analysis, reporting, and decision making. To address the long lag time due to manual data entry, the new system was developed with the capability ingest, validate, clean, match existing records, and merge data automatically, based on business rules developed by DCHHS. Not only does this provide more rapid availability of data for investigation, but it also provides electronic queues for workload management and tracking. From all this information, the system creates a single complete, comprehensive person-based data source for all of this relevant public health information.

Design and Implementation Model Practices and Governance

Funding

Dallas County Health and Human Services (DCHHS) leveraged several grant resources for the project. Primary among those were the grant funded by CDC entitled National Initiative to Address COVID-19 Health Disparities Among Populations at High-Risk and Underserved, Including Racial and Ethnic Minority Populations and Rural Communities, Coronavirus Aid, Relief, and Economic Security (CARES) Act funded by the US Department of Treasury as well as the IDCU/COVID grant funded by the Texas Department of State Health Services. Additionally, DCHHS continues to identify additional funding sources to support the Data Modernization efforts.

Project Governance Structure



The DCHHS Department Director, Dr. Philip Huang, a physician, has been responsible for the overall oversight of the design, implementation, and managing of the information and technology interventions with the system. Additionally, over the course of the projects lifetime, over ten DCHHS program epidemiologists have provided critical subject matter expertise as product owners to provide detailed content knowledge and end-user perspective for the system, design, and implementation. The Dallas County, IT Director, and her leadership team have provided specific IT support for the project.

In addition, there has been a Steering Committee of external partners to provide additional subject matter expertise and perspective related largely to external factors required with implementation of the system. Steering Committee participants have included: representatives from other local health departments (C-Suite and HIT experts from Tarrant County Health Department and the Houston Health Department), area health systems (The Parkland Hospital system, the Parkland Center for Clinical Innovation, Baylor Scott & White, as well as the Texas Epic Collaborative), and the Texas Department of State Health Services. This group has served as a sounding board and working group to discuss

common issues and problems and identify strategies to work with key external partners, whose cooperation is needed for the success of this project.

The DCHHS project stakeholders work closely with a team from Accenture to bring the platform to life and continually evolve it to align with the business need. Rachel Corbin, delivery lead and manager at Accenture, works closely with Dallas County daily to design and align business requirements with technical requirements. After rounds of requirement grooming these technical requirements are then passed to an Accenture team of 23 analysts, developers and testers for development and implementation.

Project Governance Future

The governance structure may evolve as the system matures and requires less basic development.

Collaborations

Contract/Business Associate Agreement (BAA) with Accenture for system development and implementation.
Partnership and Memorandum of Understanding (MoU)/BAA with Parkland Center for Clinical Innovation (PCCI)for automated reverse geocoding/address validation and data analysis and modeling.

Workflow Design

For the development of this system, we follow an agile methodology. Throughout the development lifecycle the product owners provide guidance and testing to ensure the data and system workflows are as expected. This includes identifying when data wasn't processed, duplicates/logic updates needed, LOINC/SNOMED filtering updates, and more.

Ultimately, the new workflow was designed to streamline the process of investigations and data analysis and automate a portion of manual tasks. When the data is collected from different sources it is automatically mastered, cleaned and deduplicated in Informatica MDM based upon requirements set by the epidemiologists who used to do the task manually. The data is then archived in a Datawarehouse and sent to Salesforce for end users to leverage. Within Salesforce, the data is automatically related together and displayed giving the users a complete view without any human interaction. End users can then complete their investigations, automatically send data to the State of Texas, create reports, and visualize the data within the system.

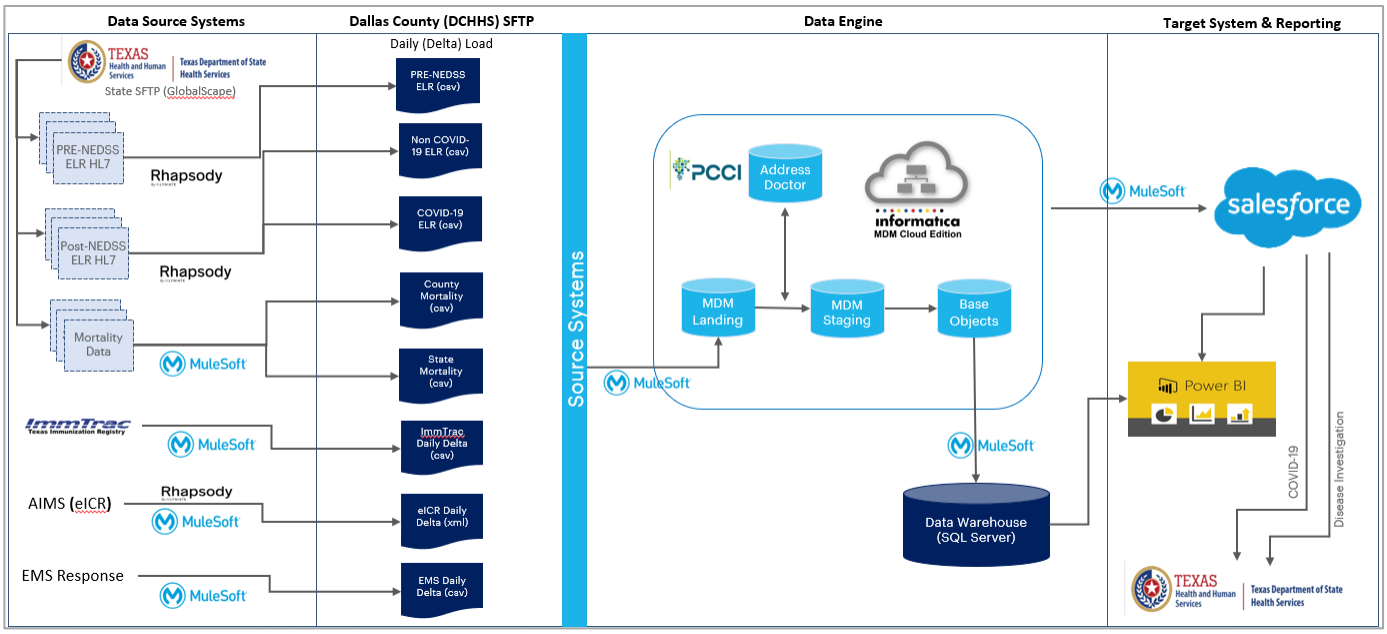
Training and Onboarding

End users are provided virtual group or one-on-one trainings that demonstrate the disease investigation process from start to finish within Salesforce. A number of these trainings have been recorded and are saved locally for current and future users to access. A user guide and other reference materials containing visuals and detailed instructions have been created to assist users with case investigations and data entry. Training materials are

updated regularly, new Salesforce features and functionality are demonstrated to staff in monthly Teams calls, and product owners are available for questions via email. All training documents are stored in a central location for ease of access. Feedback is used to improve system interfaces, workflows and design, and new users have greatly appreciated the improved capabilities and intuitive interfaces.

Information and Technology Solutions and Workflow

Longitudinal Public Health Workflow



Technologies Used

- **Data management**: Robust disease surveillance relies on trustworthy data. Using Informatica MDM, Mulesoft, and Rhapsody, Dallas County designed a Data Engine that leverages automation to reduce manual intervention and limits human error. It collects and aggregates data from previously siloed data sets – including electronic lab reports, hospital electronic case reports, vaccination data, vital statistics, and more. Using rules-based data cleaning and deduplication, it standardizes, enriches, and supports data quality.

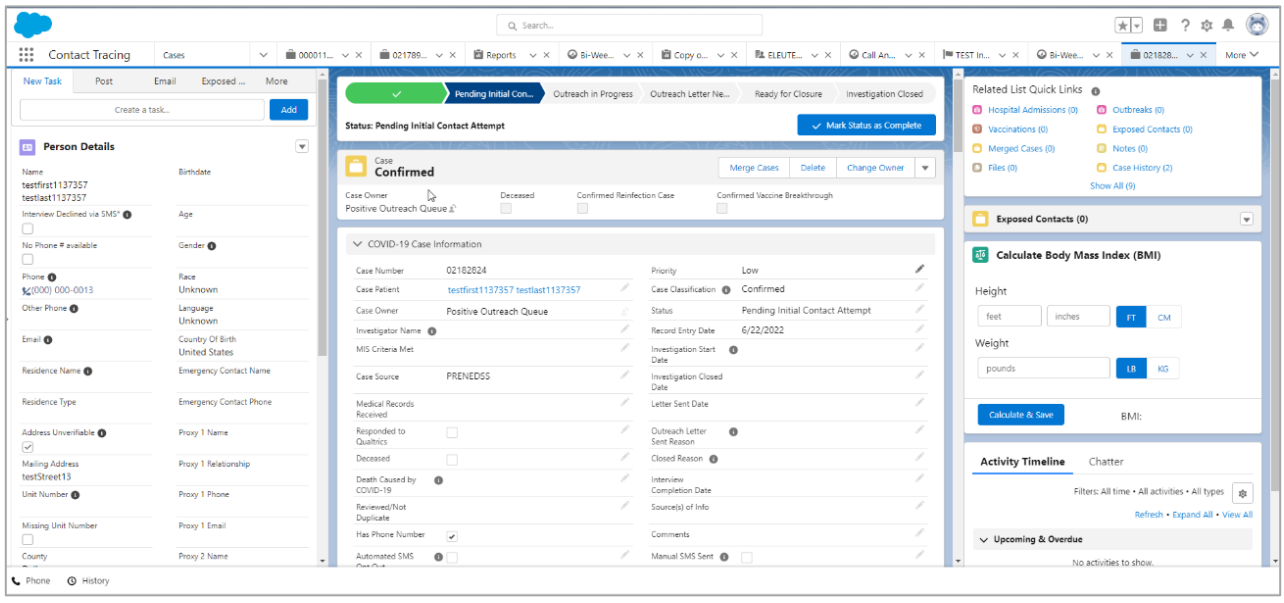
- **Disease surveillance, investigation, contact tracing, and monitoring**: Using Salesforce, Dallas County connected the Data Engine with case investigations, providing streamlined workflows and the ability to view longitudinal data for an individual across all notifiable health conditions. The integrated virtual agent allowed for automated outreach to patients via text and email. Direct integration with the Data Engine provided timely, high-quality data. In 2022, Dallas County expanded these capabilities to address immediate response needs regarding Mpox and the Ebola outbreak in Uganda (monitoring returning travelers). The system has

been expanded to cover over 100 reportable conditions and is scheduled to include all reportable conditions and sexually transmitted infections by July 2024.

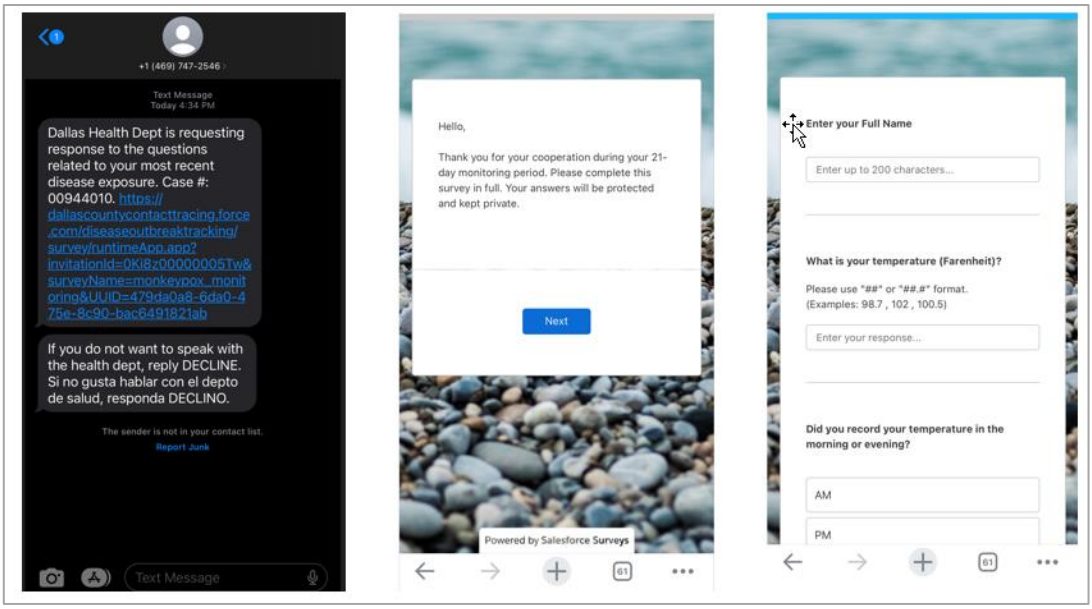
- **Analysis and Reporting**: The platform provides enhanced and automated reporting to Texas – allowing Dallas County officials to send clean, standardized data from within the investigation platform. The integrated, near real-time dashboards support data-driven and informed health policy decisions, such as managing disease outbreaks, community outreach, training disease investigators, modernizing disease reporting, and more.

Interfaces

The cleaned data is automatically ingested, creating test results, person accounts, condition histories and investigations, which SMEs can then pick up from a queue. If needed, there is also the ability to manually create any of these if the case was received through non-electronic mechanisms (e.g., phone call). With the data in the system, SMEs can continue and complete investigations while seeing the entirety of the person's history. If contact tracing is recommended for a given disease or situation, SMEs can use the disease monitoring functionality to identify and be in touch with potential contacts. If new diseases emerge or outbreaks identified, the systems also provide functionality to understand trends, identify hotspots, visualize trends, and even report directly to the State of Texas from the system, which ensures that all requirements are met before a report is sent, improving the quality and standardization of shared data.



Patient Monitoring



Call Monitoring

Data Points and Sources

Electronic Lab Reports (ELR) – Raw daily data feed of Dallas County electronic lab reports for all conditions, provided by the Texas Department of State Health Services.

Electronic Initial Case Report (eICR) – Daily data feed of hospital electronic case reports from APHL/AIMS. Currently ingesting COVID-19, MPox and Pertussis, but working on expansion to all other reportable conditions.

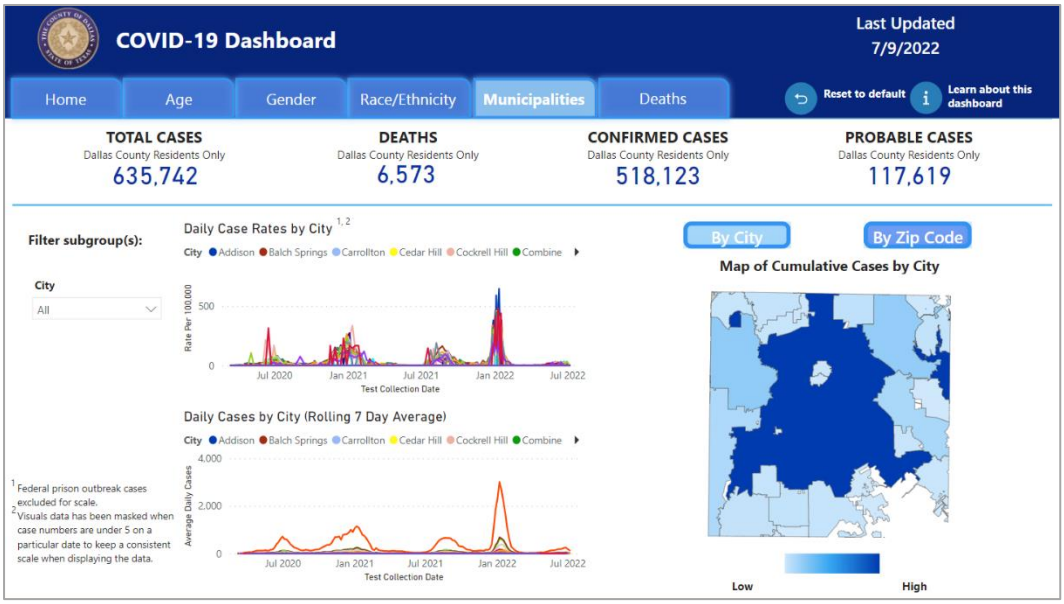
Vaccination/ImmTrac Data – Daily feed of COVID-19 vaccination data from the Texas Department of State Health Services.

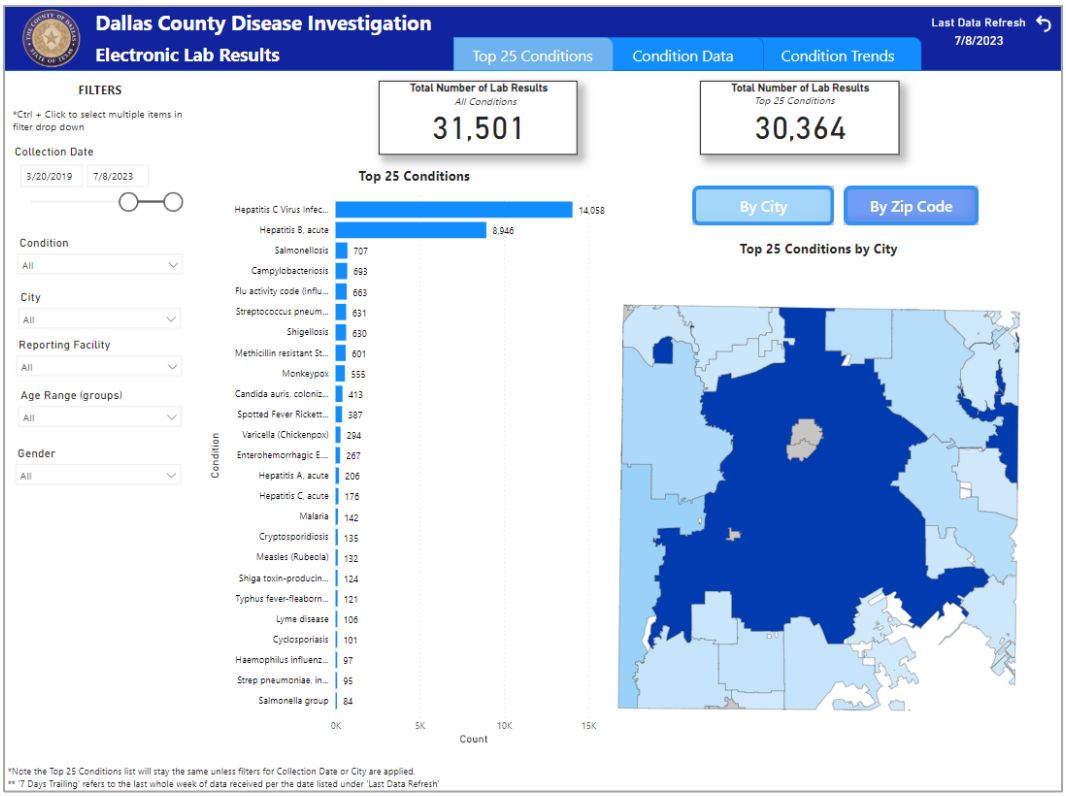
Mortality Data – Daily feed of state mortality reports for COVID-19 from the Texas Department of State Health Services, and a weekly feed of all-cause mortality data from the City of Dallas Office of Vital Records and Dallas County Clerk Office.

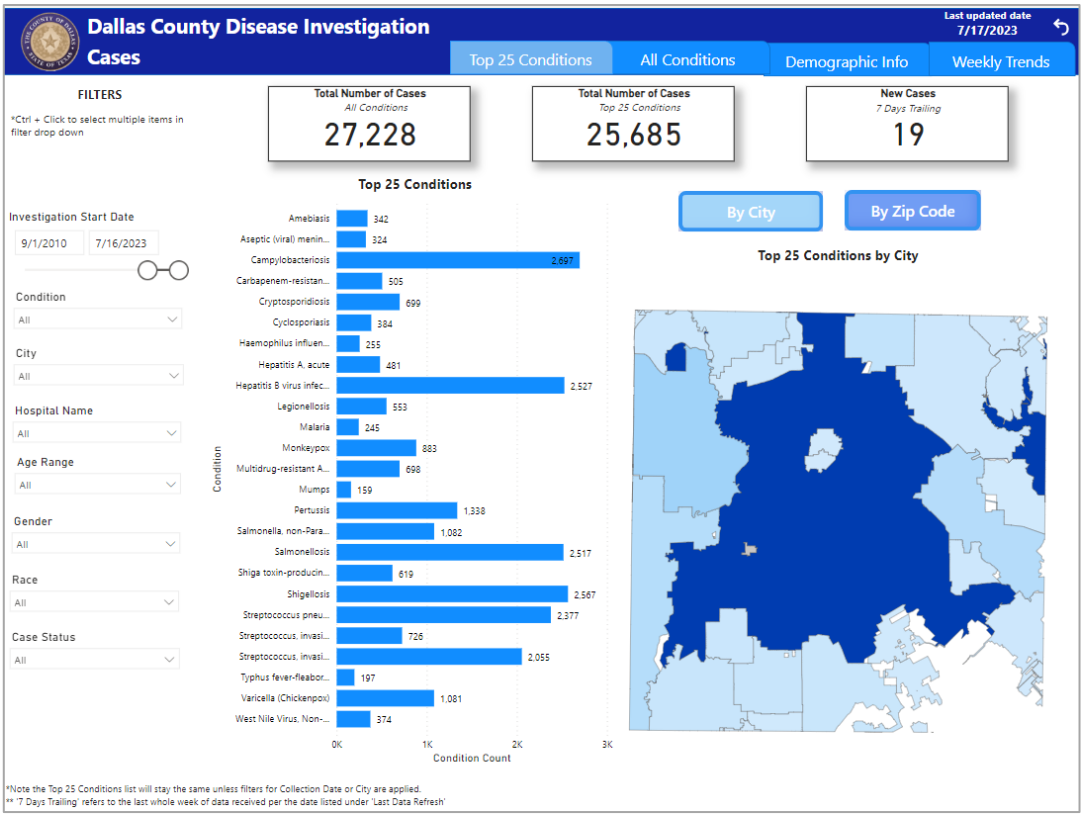
Legal authority for utilizing the data elements is based on local public health authority for epidemiological purposes. MoU's and data use agreements with the Texas Department of State Health Services and others were modified as needed. Negotiations are still underway trying to work through privacy concerns to add data on the homeless population, and the legal teams are working on appropriate MoU language. During the COVID-19 pandemic privacy concerns with school data were largely addressed through state-wide policies from the Department of Education.

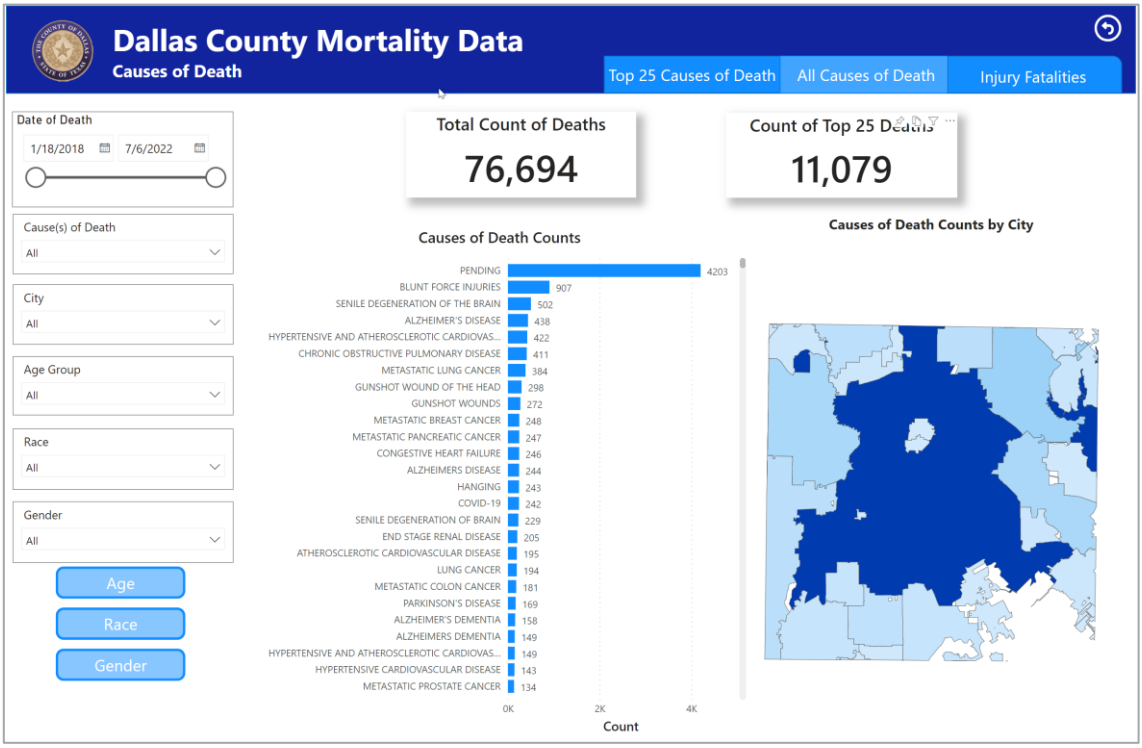
Data Visualization and Insight Generation

Data dashboards for COVID-19 have been on the public website for DCHHS throughout the pandemic. Data visualizations and other data collected by the system were presented to Dallas County Commissioners at every Commissioners Court session throughout the pandemic. Data was also presented to Dallas City Council Committee meetings, numerous presentations to community leaders, the media, and other decision-makers.


COVID-19 Dashboard
Last Updated
7/9/2022
Home
Age
Gender
Race/Ethnicity
Municipalities
Deaths
Reset to default
Learn about this dashboard
TOTAL CASES
Dallas County Residents Only
635,742
DEATHS
Dallas County Residents Only
6,573
CONFIRMED CASES
Dallas County Residents Only
518,123
PROBABLE CASES
Dallas County Residents Only
117,619
Filter subgroup(s):
City
All
Daily Case Rates by City 1, 2
City Addison Balch Springs Carrollton Cedar Hill Cockrell Hill Combine
Rate Per 100,000
500
0
Jul 2020
Jan 2021
Jul 2021
Jan 2022
Jul 2022
Test Collection Date
Daily Cases by City (Rolling 7 Day Average)
City Addison Balch Springs Carrollton Cedar Hill Cockrell Hill Combine
Average Daily Cases
4,000
2,000
0
Jul 2020
Jan 2021
Jul 2021
Jan 2022
Jul 2022
Test Collection Date
1 Federal prison outbreak cases excluded for scale.
2 Visuals data has been masked when case numbers are under 5 on a particular date to keep a consistent scale when displaying the data.
By City
By Zip Code
Map of Cumulative Cases by City
Low
High


Dallas County Disease Investigation
Electronic Lab Results
Last Data Refresh
7/8/2023
Top 25 Conditions
Condition Data
Condition Trends
FILTERS
*Ctrl + Click to select multiple items in filter drop down
Collection Date
3/20/2019
7/8/2023
Condition
All
City
All
Reporting Facility
All
Age Range (groups)
All
Gender
All
Total Number of Lab Results
All Conditions
31,501
Total Number of Lab Results
Top 25 Conditions
30,364
Top 25 Conditions
Hepatitis C Virus Infec... 14,058
Hepatitis B, acute 8,946
Salmonellosis 707
Campylobacteriosis 693
Flu activity code (Influ... 663
Streptococcus pneum... 631
Shigellosis 630
Methicillin resistant St... 601
Monkeypox 555
Candida auris, coloniz... 413
Spotted Fever Rickett... 387
Varicella (Chickenpox) 294
Enterohemorrhagic E... 267
Hepatitis A, acute 206
Hepatitis C, acute 176
Malaria 142
Cryptosporidiosis 135
Measles (Rubeola) 132
Shiga toxin-producin... 124
Typhus fever-fleaborn... 121
Lyme disease 106
Cyclosporiasis 101
Haemophilus influenz... 97
Strep pneumoniae, in... 95
Salmonella group 84
Condition
0K
5K
10K
15K
Count
By City
By Zip Code
Top 25 Conditions by City
*Note the Top 25 Conditions list will stay the same unless filters for Collection Date or City are applied.
** '7 Days Trailing' refers to the last whole week of data received per the date listed under 'Last Data Refresh'





Rapid Deployment

In 2022, Dallas County expanded the system capabilities to address immediate response needs regarding Mpox and the Ebola outbreak in Uganda (monitoring returning travelers). The system has been further expanded to cover over 100 reportable conditions and is scheduled to include all reportable conditions and sexually transmitted infections by July 2024. Rapid deployment proof-of-concept was demonstrated with development of a template for all requirements needed for adding a new condition. This was tested for Marburg Virus and the condition was added within one two-week sprint.

Impact on Public Health

Implementation of the Dallas County Disease Surveillance and Investigation System has enabled Dallas County Health and Human Services to successfully manage the high volume demands during the COVID-19 Pandemic related to data ingestion, disease investigation, contact tracing (including standing up system to accommodate hundreds of contact tracers), data reporting and visualization, and state and CDC reporting requirements. The system was expanded to address local needs related to the Mpox epidemic, including automated patient symptom monitoring, showcasing its ability to enhance DCHHS preparedness and response. It is also currently being expanded to accommodate all reportable conditions, and the future state will accommodate data related to other public health issues of importance such as chronic disease, opioid overdoses, and other conditions, providing a comprehensive view of public health. A proof-of-concept rapid deployment for any new emergent conditions was demonstrated for Ebola, with a recent epidemic seen in West Africa.

Timely and accurate data for COVID-19 were critical for local policy decision-making related to workplace closures, masking recommendations, etc. Obtaining accurate granular data were also critical for addressing health equity issues in the community to facilitate identification areas of higher need, and targeting of community outreach activities related to testing, vaccine administration, education, and other efforts to reduce health disparities.

Improved Efficiency and Accuracy

Prior to implementation of the Dallas Disease Surveillance and Investigation System, the following was the status:

1. Data Integrations: DCHHS staff were unable to import flat files (e.g. CSV files) of test results concurrently while manual data entries were occurring by large teams of staff members and all data entry or wait outside of operating hours

2. Manual Workflow: A less efficient manual entry process was required to manually enter fax information that was burdensome and meant less information was able to be converted into electronic records each day.

3. De-duplication: Prior to the Dallas Disease Surveillance and Investigation System, the process of de-duplication and address verification of all test reports required a cumulative of 15 hours per day by data management staff members on average

4. Address Validation: Prior to the integration of the geocoding service in the Salesforce platform, staff manually searched and validated the address for every report received by the health department to ensure the patient was a Dallas County resident and that the address was valid and accurate. Staff even had to manually search the Dallas County Tax Assessors website to confirm case jurisdiction.

5. Text Message Functionality: Previously this was not a feature and meant a significant amount of calls out to Dallas County residents were not successful as no prior notification had been sent to the patient.

6. Daily Reporting and Biweekly Analysis: Limited reporting was possible previously, which meant a lot of manual effort was required to process, combine, and clean the data prior to reporting. This effort repeatedly required staff to work well into the night to complete and produce reports.

After implementation of the Dallas Disease Surveillance and Investigation System

1. Data Integrations: Following electronic lab report ingestion and mapping of 7 different file formats and implementation of deduplication logic, much less time is now needed to prepare files for integration into the new database. eCR integration allowed COVID case report data to be directly imported into DCHHS' Salesforce platform from some of the largest hospitals and health systems in Dallas County. Ingestion of COVID-19 vaccination information and mortality data for all Dallas County residents were also integrated into the system and laid the foundation for similar ingestions for all other public health conditions of interest.

2. Manual Workflow: An enhanced workflow was created to allow positive cases to be manually entered into Salesforce from scanned electronic faxes to save printing costs, reduce the need for maintenance of paper files. This enabled existing limited staff resources to be used more efficiently.

3. De-duplication: Complex logic has been built into the data ingestion process into the Salesforce platform, to screen each positive case to reduce the introduction of duplicate patient records and therefore increase data accuracy within the platform.

4. Address Validation: The system has enabled integration with a custom geocoding service (e.g. API logic) that allows automatic address validation to more efficiently identify patient residences that are located within the borders of Dallas County and each respective city, in order to increase data accuracy in the platform.

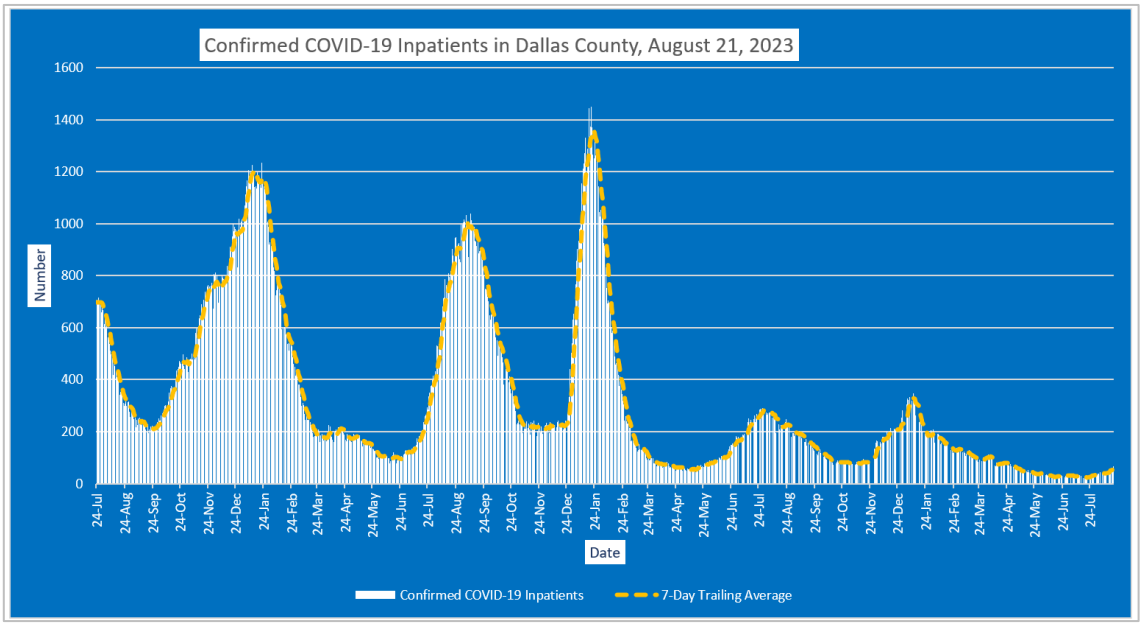
5. Text Message Functionality: During the peak of COVID-19, the system implemented an automated 2-way text messaging campaign and inbound call queue process that notifies patients in advance of any call attempts and provides them the opportunity to call case investigators directly. This significantly increased the amount of successful contacts. This capability is available to meet future needs.

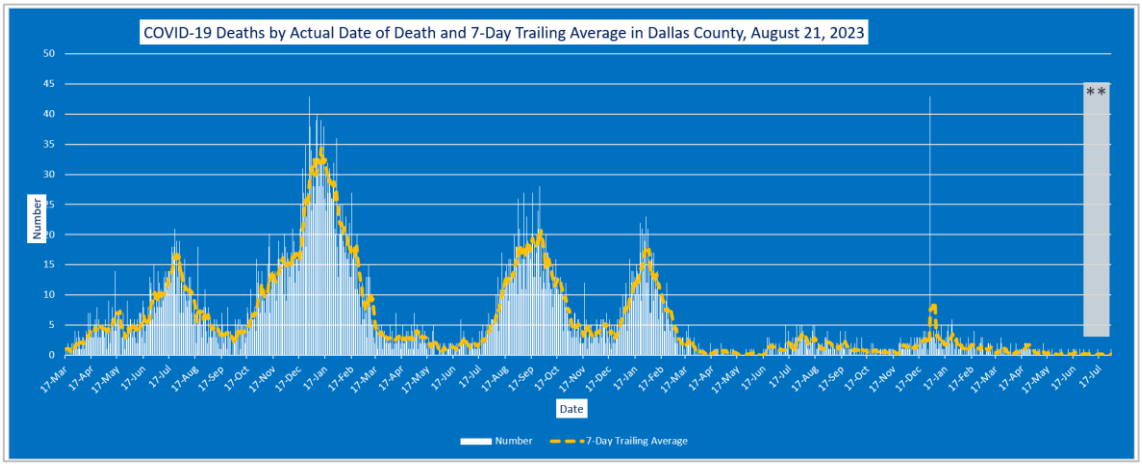
6. Daily Reporting and Biweekly Analysis: Automatic reports and dashboards that show productivity and highlight areas when efficiencies can be made have significantly

reducing the amount of overall manual effort previously required to process, combine, and clean the data prior to reporting.

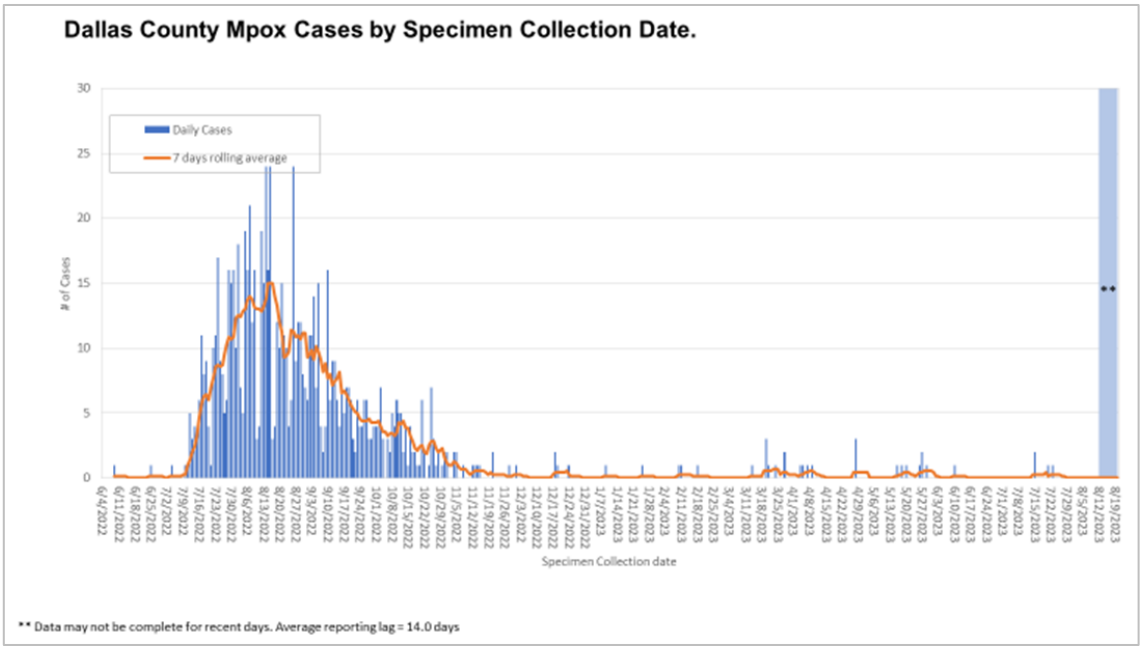
Increases in Life and Healthy Life Expectancy

Implementation of the Dallas Disease Surveillance and Investigation System for COVID-19 has been critical to facilitating the public health response for community education, targeted priority community vaccine efforts, and ultimately community-wide impact on controlling spread of disease, reducing hospitalizations, and reducing deaths:



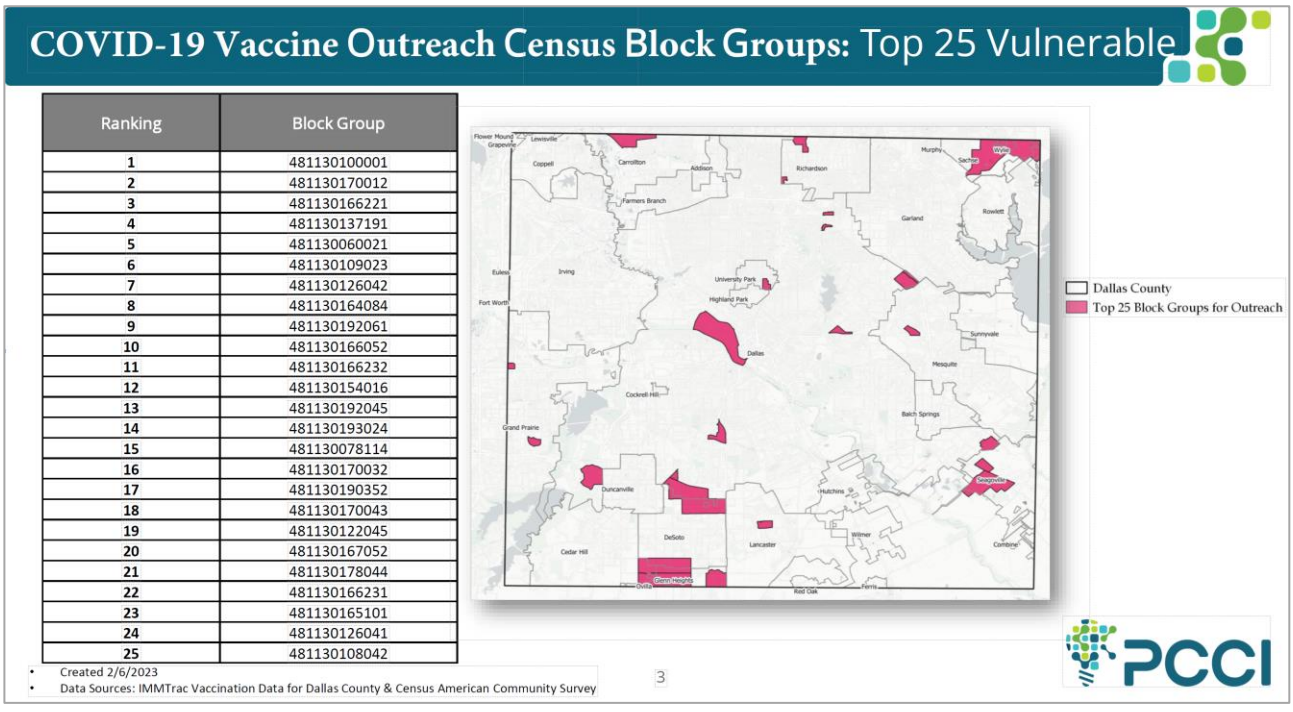


Implementation of the Dallas Disease Surveillance and Investigation System for MPox has also been instrumental in facilitating the public health response for community education, vaccine administration, and ultimately control of this epidemic.



Improves Health Equity and Access to Health and Community Services

Implementation of the Dallas Disease Surveillance and Investigation System has been critical to address health equity and access issues through enabling the capability of providing granular COVID-19 data to identify census block-level analysis to identify highest priority areas with greatest burden and highest need for targeting of outreach and community vaccination efforts:

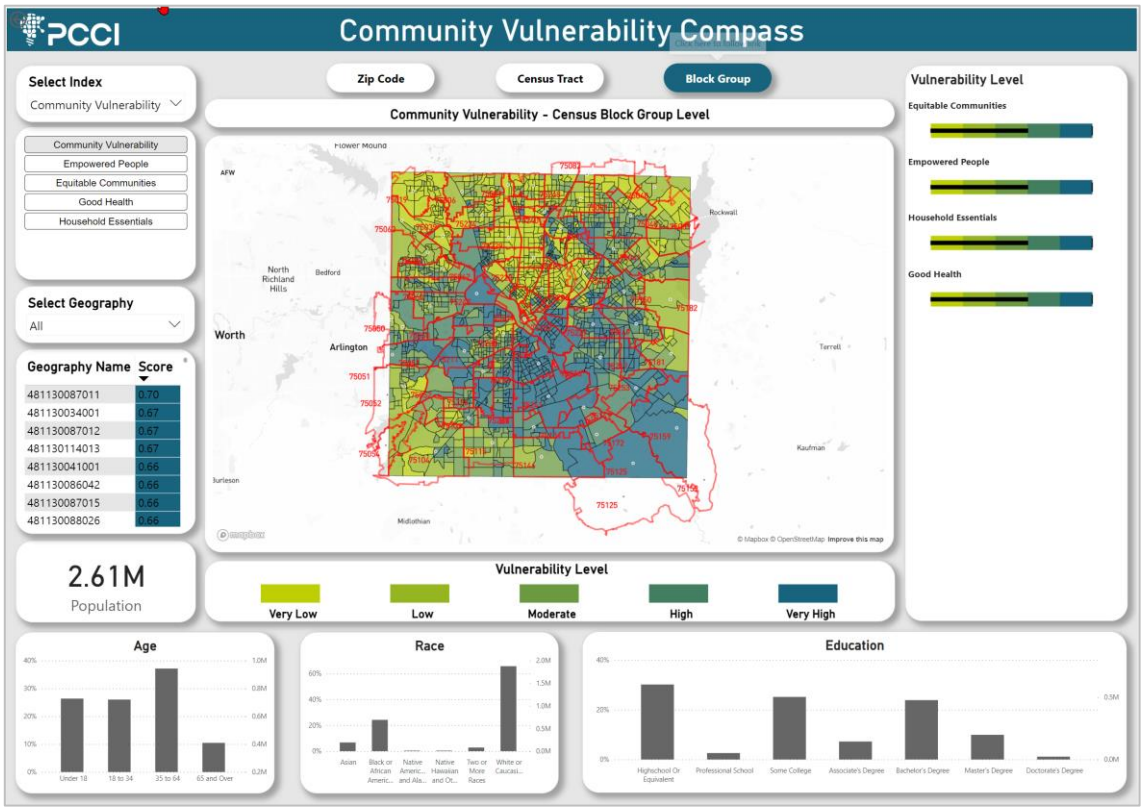


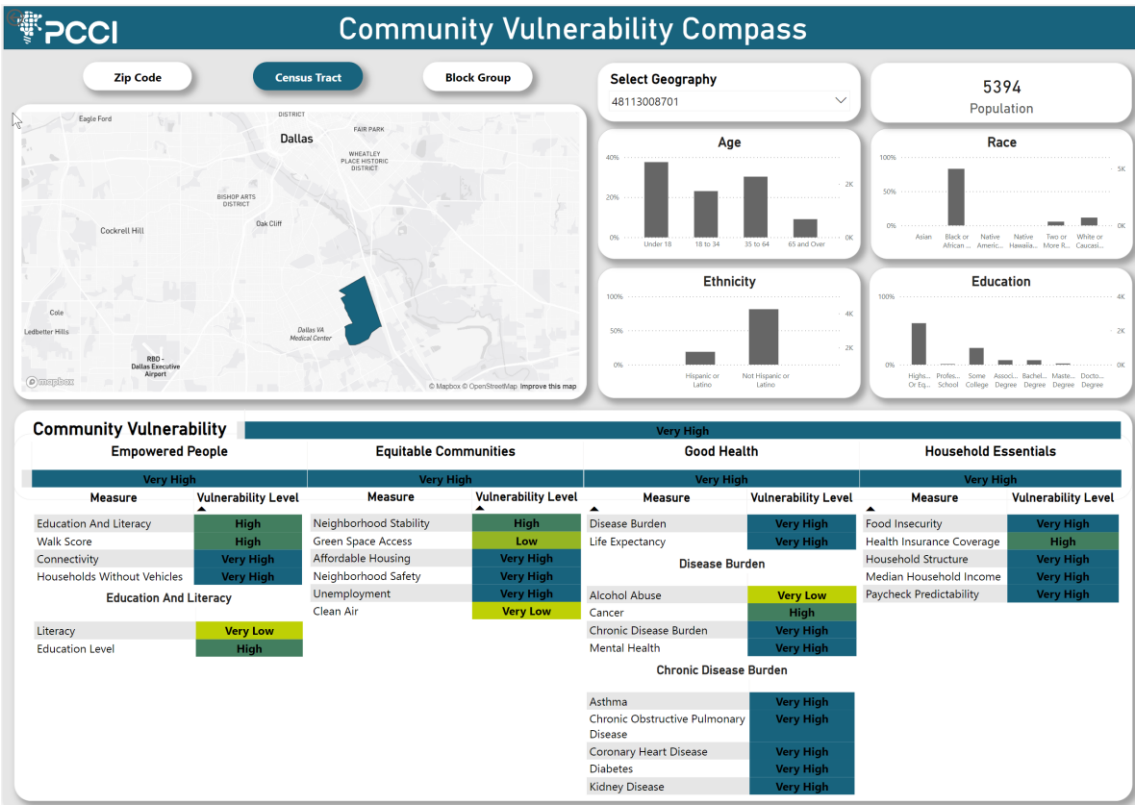
Ranking	Block Group
1	481130100001
2	481130170012
3	481130166221
4	481130137191
5	481130060021
6	481130109023
7	481130126042
8	481130164084
9	481130192061
10	481130166052
11	481130166232
12	481130154016
13	481130192045
14	481130193024
15	481130078114
16	481130170032
17	481130190352
18	481130170043
19	481130122045
20	481130167052
21	481130178044
22	481130166231
23	481130165101
24	481130126041
25	481130108042

Upcoming Enhancements

Community Vulnerability Compass Integration

DCHHS's partner, the Parkland Center for Clinical Innovation (PCCI), has developed a community vulnerability compass that facilitates simplified access to a comprehensive summary of social determinate information for any specific address, census block, census tract or zip code. This summary is being incorporated into the Disease Investigation platform to provide this associated information for any individual locations address.





Master Data Management (MDM) SaaS Migration

The Informatica MDM platform is currently moving from a hosted environment to a SaaS environment. Some of the new features that will result include:

- Data Profiling – Ability to quickly profile any dataset and immediately get a holistic understanding of the data quality and design quality rules for transformation and standardization.
- Use of artificial intelligence (AI) to automate direct file ingestion mapping and entity recognition.
- Use of AI and natural language interface to enable teams to ask questions about the data without requiring any knowledge of the SQL. This will reduce time for data classification, enhance data discovery, uncover key data insights, and improve productivity.
- Use of machine learning to derive data driven match rules to better increase the match accuracy with ability to train and retrain the match model

HIMSS Global Conference Audience Guidance

Topic Guidance: Check three which apply to this case study

Clinical Informatics and Clinician Engagement Clinically Integrated Supply Chain

Consumer/Patient Engagement and Digital/Connected Health

Consumerization of Health

Culture of Care and Care Coordination

Data Science/Analytics/Clinical and Business Intelligence

Disruptive Care Models

Grand Societal Challenges

Health Informatics Education

Health Information Exchange

Interoperability

Data Integration, and Standards

Healthcare Applications and Technologies Enabling Care Delivery

Healthy Aging and Technology

Improving Quality Outcomes

Innovation, Entrepreneurship, and Venture Investment

Leadership, Governance, and Strategic Planning

Population Health Management and Public Health

Precision Medicine and Genomics

Process Improvement, Workflow, and Change Management

Social, and Behavioral Determinants of Health

Telehealth

User Experience (UX)

Usability

User-Centered Design