Data Modernization Initiative

Case-based surveillance capabilities and technology recommendations

August 2021

Public Health Informatics Institute

Table of Contents

Acknowledgements

This project was supported by cooperative agreement number 6-NU38OT000316, funded by the Centers for Disease Control and Prevention. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention or the Department of Health and Human Services.

Acronyms

APIs	Application programming interfaces
AI	Artificial intelligence
APHL	Association of Public Health Laboratories
ASTHO	Association of State and Territorial Health Officials
CDC	Centers for Disease Control and Prevention
CSELS	Center for Surveillance, Epidemiology, and Laboratory Services
CoAg	Cooperative agreement
CSTE	Council of State and Territorial Epidemiologists
DMI	Data Modernization Initiative
DHIS	Division of Health Informatics and Surveillance
EHR	Electronic health record
FHIR	Fast Healthcare Interoperability Resources
HIE	Health information exchange
IIS	Immunization information system
MPI	Master patient indexes
PHII	Public Health Informatics Institute
RLS	Record locator services
RCKMS	Reportable Conditions Knowledge Management System
STLT	State, tribal, local and territorial
TEFCA	Trusted Exchange Framework and Common Agreement
USCDI	U.S. Core Data for Interoperability

Executive Summary

The Centers for Disease Control and Prevention (CDC) has invested in a data modernization initiative (DMI) to improve public health surveillance. Through the DMI, surveillance systems will contain more timely, accurate and comprehensive data through enhanced interoperability across the public health ecosystem. The CDC engaged the Public Health Informatics Institute (PHII) to identify the capabilities and technical requirements of future surveillance systems that will meet those specifications. PHII identified five key partner groups—public health jurisdictions, public health associations, public health informatics, technology and the CDC—that routinely interact with current surveillance systems and could inform what is needed to enhance their flexibility, scalability and interoperability with healthcare and public health.

Methods

PHII developed recommendations for future surveillance systems based on three distinct information gathering methods. First, we reviewed previous assessments of case-based surveillance systems and similar publications to understand the evolution of current functions and capabilities. Following the review of literature, each of the five key partner groups participated in a listening session to discuss critical components of future surveillance systems. Finally, using the information gathered from these sources, PHII developed recommendations for future surveillance systems. An expert panel reviewed the recommendations and offered feedback that shaped the final recommendations for case-based surveillance presented in this paper.

Recommendations

The recommendations reflect a forward path for system technological functions and capabilities as well as opportunities to evolve case-based surveillance strategies and approaches. Note that these recommendations address both social and technical aspects of surveillance systems; the current challenge is as much about the nature of people and organizations as it is about technology.

1. Establish automation through integration of surveillance systems and case management systems.
2. Balance the push and pull of data from electronic health records (EHR) to public health departments.
3. Use artificial intelligence (AI) to automate processes and predict trends.
4. Establish uniform system standards and functionality to promote cross-jurisdictional interoperability by creating functional requirements to collect the same information across programs.
5. Harmonize data across public health programs by using common data models and standardized data sets.
6. Ensure flexibility and modularity of systems development and maintenance.
7. Support scalable systems so they are readily available as the demand of the moment changes, including the use of cloud technologies.
8. Improve entity and patient matching to improve overall data quality.
9. Improve partnerships and cross jurisdictional coordination.
10. Support intensive data access required for dedicated query systems (e.g., data lakes, data warehouses).

11. Streamline or simplify data transport.

In addition to the surveillance system recommendations, socio-technical recommendations emerged during this process. There was also a need to strike the right balance between creating systems that can accommodate the unique requirements of individual jurisdictions as well as promote interoperability across agencies and industries.

- **Socio-technical recommendations for future surveillance systems:** There was resounding agreement that identifying the correct technology and defining the correct capabilities will only address some of the limitations of current surveillance systems. People, policies, procedures and funding all impact the system and the ability to apply technology to the full extent necessary to modernize surveillance systems.
- **Finding the right balance:** Critical decisions must be made that will impact the speed of implementation and the way in which future efforts to further modernize systems unfold. These decisions will often present trade-offs and create tension in selecting one option over another.

Background

The COVID-19 pandemic highlighted the need for timely and accurate surveillance data for federal, state, local, tribal and territorial public health agencies. The Centers for Disease Control and Prevention's (CDC) Data Modernization Initiative (DMI) is part of a national effort to develop integrated public health data and surveillance systems to support the COVID-19 response and prepare for future public health emergencies. Case-based surveillance systems play a significant role in conducting both routine and emergency public health functions. With increasing demands to access and report data in near real-time, surveillance programs need to modernize their systems to adapt to changing needs and leverage new technologies.

The Public Health Informatics Institute (PHII) received funding through CDC's data modernization cooperative agreement (CoAg) to develop recommendations for state, tribal, local and territorial (STLT) health departments. This report identifies critical case-based surveillance system capabilities and supporting technology standards needed to modernize public health surveillance during the next five to ten years.

Methods

To develop this report and recommendations, PHII conducted a series of listening sessions with experts on public health surveillance technology. PHII convened an expert panel in June 2021 to provide feedback on the recommendations.

The listening sessions involved key partners from public health jurisdictions, public health associations, CDC, and experts in public health informatics and technology. They shared input on priorities, key capabilities and future technologies for case-based surveillance systems. Prior to the listening sessions, PHII reviewed past assessments of case-based surveillance systems and prior recommendations for capabilities. This information was used to develop key questions and discussion points for the listening sessions.

Next, PHII and staff from the Center for Surveillance, Epidemiology, and Laboratory Services (CSELS) at CDC identified thought leaders from the public and private sectors and invited them to participate in a 60-minute listening session. Six listening sessions were conducted in April and May 2021. Thirty-seven participants representing five partner groups participated (Appendix 1). Participants were assigned to listening sessions based on their partner group to encourage discussion and information sharing. Each session began with an overview of the project, the goals for the listening session and description of the format, followed by a participant-led discussion. To guide the discussion, participants were provided with key questions that the recommendations will address and talking points that were customized for each partner group (Appendices 2 and 3). All sessions were recorded with verbal permission from the participants. Audio files were transcribed using Speechpad and qualitative data analysis techniques—including coding and grouping of keywords— were used to identify themes.

Following the listening sessions, an expert panel convened for one 90-minute session. The panel was composed of leaders in the fields of public health, informatics, information technology and surveillance systems. A summary of the themes and recommendations that emerged during the listening sessions was provided to the panel for review prior to their session. During the meeting, the panel discussed the findings from the listening sessions and suggested refinements to the recommendations.

PHII enlisted staff with expertise in evaluation and public health information technology to listen to all sessions conducted as a part of this project. These experts participated in the synthesis of the listening sessions to draft the recommendations.

Findings

PHII reviewed the transcripts from the participant discussions and highlighted key focus areas for each of the five partner groups. Within each focus area, themes were identified and grouped based on similarity and categorized in the final recommendations. Table 1 below outlines the key focus areas and themes that emerged during listening sessions.

Table 1. Listening session focus areas and themes by partner group

Partner group	Discussion Focus Areas and Themes
Public health jurisdictions	• Automate processes to reduce program burden and streamline ingesting large amounts of data • Use artificial intelligence (AI) to streamline data processes and conduct predictive analytics • Harmonize data across programs to improve data consistency and completeness • Configure systems to support core surveillance system functionality while allowing jurisdictions to adjust to their needs and supporting flexibility as new use cases arise, such as data exchange with immunization information systems (IIS) • Leverage Fast Healthcare Interoperability Resources (FHIR) and application programming interfaces (APIs) to facilitate data exchange • Integrate reporting platforms including data lakes and data warehouses • Facilitate patient record matching across systems and data sources through the use of master patient index

	• Enhance workforce capacity to support more and better training for public health informaticians with epidemiology and information technology skills; increase ability for jurisdictions to compete with the private sector to recruit and retain staff
Public health associations	• Standardize system functionality expectations similar to IIS functional standards • Integrate with case management systems • Use data lakes to allow jurisdictions to control the flow of surveillance data into the system • Reduce or eliminate funding siloes to facilitate collaboration across systems and programs to achieve common goals • Harmonize data to increase consistency in data collection across programs

CDC	• Case-based surveillance system should seek to adopt FHIR and RESTful APIs
	• Case-based surveillance systems should be flexible, agile and sustainable:
	  ○ Flexibility in data formats and transport methods
	  ○ Processing data without a lot of software development/coding
	  ○ Reduce processing burden on system and programs
	  ○ Store data for streamlined access and analysis without adding load to system
	• Case-based surveillance systems should adopt common data models, core data sets and common data sets such as United States Core Data for Interoperability (USCDI)
	• Case-based surveillance systems should be more standardized across jurisdictions
	• AI should be used to “sniff” messages and make decisions based on what is detected
	• Case-based surveillance systems should be more modular while being conscious of challenges with swapping components and the impact of changes on interoperability

Surveillance system vendors	• Harmonize data across systems and jurisdictions, including a standardized format and structure to improve consistency and completeness of data • Standardize data transports to facilitate interfaces with other data systems • Address policies that hinder data flow within and across jurisdictions • Leverage AI for predictive analytics • Reassess the purpose of surveillance systems to address modern needs for real-time flow of data and more timely reporting
Electronic health record vendors	• Develop a standardized case-based surveillance platform and potentially integrate case-based surveillance systems across programs

- Take advantage of FHIR and leverage health information exchanges (HIEs) for data exchange
- Develop a centralized approach for data exchange
- Leverage AI for predictive analytics
- Encourage public health to participate in standards definition forums
- Build public health informatics expertise, especially at the policy level

Surveillance System Functions and Priorities

During the listening sessions, the participants highlighted interoperability, data quality and management of case records as key case-based surveillance system functionalities that should be prioritized for enhancement and support in the near future.

Interoperability

Case-based surveillance systems must be able to establish, re-establish and modify interfaces with other information systems for the electronic exchange of case and demographic data. The interface should enable the surveillance system to facilitate electronic data exchange with other information systems, including electronic health records systems (EHRs), electronic laboratory systems (ELRs) and vital records, including electronic birth and death records systems. The surveillance system should be able to interface with internal health department systems such as syndromic surveillance systems as well as external systems. It should also be able to exchange data electronically, as well as monitor and troubleshoot data exchange.

Data quality

The case-based surveillance system should be capable of ensuring data quality through patient, event and entity matching and deduplication functions. The system should be able to identify and manage duplicate and potential duplicate patient records and event entries such as laboratory results and electronic case reports.

Managing case records

Case-based surveillance systems should be able to integrate data from multiple health information databases or systems into a single repository. Specifically, the surveillance system should be able to link cases across multiple systems and data sources whether the data is being submitted directly to the system or accessed from an external data repository.

Recommendations

These recommendations are a synthesis of the information gathered from the listening sessions and expert panel participants. They reflect a forward path for system technological functions and capabilities as well as opportunities to evolve case-based surveillance strategies and approaches. Note that these recommendations are socio-technical in nature: the current challenge is as much about the nature of people and organizations as it is about technology.

Establish automation through integration	As the public health mission evolves, so should the systems that support it. Public health surveillance systems should bridge the difference between "case management" and "surveillance" systems; few systems can completely automate the surveillance process end to end. However, an integrated reporting platform can be created by adding two systems together (e.g., contact tracing and immunization data to case management systems). As public health surveillance needs have become more immediate, there is a need to enhance systems that were not designed to be real-time. In addition, these systems should be built to require less human interaction. Finally, programs can help each other by sharing investments and deploying shared services rather than replicating all aspects of their technical implementations.
Balance "push" and "pull"	Systems should query ("pull") EHRs for necessary data rather than always expecting EHRs to push data to them. While not replacing traditional public health reporting ("push"), this shift will allow surveillance systems to better control the flow of certain data into the systems and reduce processing and storage burden.
Use artificial intelligence (AI)	As the ability to access and ingest large amounts of data across multiple sources has increased, there is a need to be more predictive and less reactive. Artificial intelligence is a method that can be used to conduct automated processes, reducing the need for manual intervention by program staff. Predictive analytics conducted by AI could also be used to anticipate disease outbreaks and detect trends more quickly. While AI may be a helpful solution for public health surveillance, foundational improvements to systems are more urgent.
Establish uniform system standards and functionality	Uniform surveillance system standards would facilitate easier data exchange with trading partners. System integration problems are based on inconsistent standards across programs. This refers to everything from standard functional requirements, to standard ways of collecting the same information across programs, to technical standards for implementation. Public health participation in standards development and selection will only improve the "fit" of standards for public health needs. More centralized services, for example, the Reportable Conditions Knowledge Management System (RCKMS) would allow a balance of jurisdictional flexibility with uniformity. Continued adoption of FHIR in public and private settings could potentially address uniformity as well as facilitate interoperability.
Harmonize data across programs	Public health program silos (e.g., disease-specific programs) still exist, even after years of attempting to reduce and eliminate them. Informatics needs to span programs and make them more coherent through more common data models. Data today is coming rapidly from different and often non-traditional sources, requiring cleaning and preparation before it can be absorbed. The use of common standardized data sets such as USCDI may help improve data consistency and completeness. Much of what we do with data rises and falls with common understanding of the meaning of the data,

	which is often inconsistent. Improved semantic interoperability by using consistent terminology will ultimately enhance data sharing and use.
Ensure flexibility and modularity of systems	An agile approach to system development and maintenance is needed to quickly meet the demands of a rapidly changing, dynamic world. Thinking in terms of *functions* instead of *systems* promotes modularity and flexibility. A mitigating strategy may include a core system supplemented by a "lite," more flexible version to help deal with immediate, urgent needs. Modular systems should also be considered. Use of microservices (an architecture that assembles systems out of small building blocks) allows additional functionality and choice in addressing emergent needs. Modularity would have to be genuine and substitutable with components that are truly interoperable through application program interfaces (APIs).
Support scalable systems	Systems should be readily scalable as the demands of the moment change. Cloud technologies—now essentially a commodity service—enable this well.
Improve entity and patient matching	The answer to patient matching may be the use of full agency master patient indexes (MPI), supported by record locator services (RLS). Improved matching will improve overall data quality, especially across programs and potentially across jurisdictions. In addition, there is the need for better directories for healthcare (and other) organizations relevant to public health data.
Improve partnerships and cross jurisdictional coordination	Data must move between jurisdictions effectively, and agencies must consider data sharing outside their borders as a fundamental requirement for their systems. National health information exchange efforts (such as eHealth Exchange, CommonWell Health Alliance, Carequality, and soon TEFCA) help make this possible through common standards and broad clinical participation. Jurisdictions could establish vendor user groups to share resources and experiences of surveillance systems.
Support intensive data access required for dedicated query systems	Transactional surveillance and case management systems cannot bear the load of increased demand for data for various audiences. Data warehouses and data lakes are needed to manage the flow of data into the system and allow collection of data for specific purposes rather than ingesting large amounts of information regardless of case report type.
Streamline and simplify data transport	Public health should attempt to get as much data through one communications "pipe" as possible. Clinical care has requested this, and public health should be able to accommodate. HIEs can be a great help in this area by providing a single connection to multiple public health systems, including case-based surveillance systems. Transport should also be standardized around fewer choices to streamline data exchange while providing some accommodation to data partners.

Conclusion

These recommendations lend themselves to more standardized, real-time and interoperable systems. However, the right balance between systems that are modular and interoperable versus those that are uniform and tightly integrated needs to be considered. System improvements that are speedier and result in short-term gains by using existing systems versus long-term (and even delayed) gains from implementing more standards-based solutions is another area for consideration. A balance between speed, customization and cost needs to be identified during the implementation process to successfully modernize systems timely and in a way that maximizes cross-entity interoperability.

There was resounding agreement that identifying the correct technology and defining the correct capabilities will only address some of the limitations of current surveillance systems. The people, policies, procedures and funding all impact the system and the ability to apply technology to the full extent necessary to modernize surveillance systems.

Appendix 1: Listening session participants

Partner groups	Number of participants	Jurisdictions/organizations represented
Public health jurisdictions	14	● Alabama ● Connecticut ● City of Houston ● Massachusetts ● North Carolina ● Oregon ● Utah ● Virginia
Public health member associations	3	● Association of Public Health Laboratories (APHL) ● Association of State and Territorial Health Officials (ASTHO) ● Council of State and Territorial Epidemiologists (CSTE)
Surveillance system vendors	7	● Conduent (Maven) ● Inductive Health ● STC ● Sunquest
Electronic health records vendors	5	● Epic ● Cerner ● Allscripts
Centers for Disease Control and Prevention	8	● Division of Health Informatics and Surveillance (DHIS) ● U.S. Digital Services

Appendix 2: Listening session key questions

- What key functionality is needed in case-based surveillance systems to support surveillance activities more effectively?
- What will make systems flexible and responsive for emergency events?
- Which functionality should be integrated into the surveillance system and what should be addressed outside the surveillance system through interoperable systems or services?
- What functionality is essential for data collection and data exchange?
- What will allow the systems to take advantage of technology to improve efficiency or effectiveness?
- What technology should surveillance systems be taking advantage of?
- What functionality is a priority to improve data integration, data cleaning and data linkage?

Appendix 3: Discussion prompts by participant group

Partner type	Discussion prompts
Public health jurisdictions	1. Shared infrastructure, enterprise-wide approaches and cloud solutions in use or considered for future use 2. Functional enhancements to improve system flexibility and responsiveness in emergency events 3. Integrating communicable and non-communicable condition surveillance systems in your jurisdictions 4. Gaps and priorities for future state of case-based surveillance systems
Public health member associations	1. Innovative technologies being used or implemented in state and local public health jurisdictions. 2. Integrating communicable and non-communicable condition surveillance systems in state and local jurisdictions 3. Gaps and priorities for future state of case-based surveillance systems
Electronic disease surveillance system vendors	1. Next generation technology that could be adapted or implemented for public health case-based surveillance systems 2. Case-based surveillance system functions and shared services 3. Standards that should be updated or developed to facilitate the collection and sharing of surveillance data and case information across jurisdictions and with CDC
Electronic health records vendors	1. Strengthening partnerships between public health jurisdictions and EHR vendors to support EHR integration with public health surveillance systems 2. Interoperability standards that should be updated or developed to facilitate the collection and sharing of EHR data with public health jurisdictions and CDC 3. Next generation technology that should be considered for implementation by public health jurisdictions to facilitate data sharing with EHRs 4. Opportunities for public health to enhance the value of electronic data exchange with EHRs
CDC	1. Innovative technologies being used or implemented in state and local public health jurisdictions 2. Next generation technology that should be considered for implementation by public health jurisdictions to enhance case-based surveillance systems 3. Gaps and priorities for future state of case-based surveillance systems 4. Shared services platforms or technologies that CDC is prioritizing to support public health case-based surveillance data infrastructure at the jurisdiction level