The Data Modernization Initiative

Jim Craver

Acting Deputy Director for Management and Operations

NCHS Board of Scientific Counselors
October 22, 2021

What is the Data Modernization Initiative?

- CDC is at the heart of a national effort to create modern, integrated, and real-time public health data and surveillance that can protect us from any health threat.

- The goal is to move from siloed and brittle public health data systems to connected, resilient, adaptable, and sustainable 'response-ready' systems that can help us solve problems before they happen and reduce the harm caused by the problems that do happen.

DMI Solves Longstanding Problems That Put Us at Risk

Siloed information
Disconnected and/or proprietary disease systems driven by disease-specific budget lines keep us from seeing the complete picture

Outdated skills
The public health workforce needs training to use today's technologies more effectively

Point-to-point data transmission
Providers in healthcare and at health departments are burdened with sending data to many places in many different ways

Outdated technologies
Older systems at health departments are not flexible, do not use cloud, and are not scalable

Public health is not a part of the healthcare data ecosystem
Public health got left behind as federal incentives and regulations helped healthcare systems to be able to easily share data automatically in the Electronic Health Record

When the next emergency happens, we will have:

- A foundation for data sharing across all levels of public health for coordinated, scalable and timely case investigation, management, and reporting

- Shared analysis capabilities for rapid identification of trends within and across jurisdictions, including:
  - Race/ethnicity-specific trends and risk factors
  - Improved forecasting and response capabilities

NCHS Data Modernization Projects

- **Data Sourcing and Production Improvements**
  - NVSS Modernization
  - Electronic Health Records
  - NHIS Pilot Projects
  - Model-Based Estimates
  - Linkage

- **Data Access Improvements**
  - **Data Query System**
  - Website Redesign
  - **Data Science Activities**
  - **Virtual Data Enclave**

Data Query System Logic Model

- Improve the ability of users to identify data sources

Activities	Outputs	Outcomes
Design a query system that improves the **discoverability, identification, and interoperability** of NCHS data assets	Data query system with **modernized data accessibility** design and development Website support services and overhauled website	**Reduced time and burden** to access identified data assets Improve the **ability of users to identify data sources**
Develop **enterprise-wide access tools** that guide users to data	Enterprise-wide access tools that effectively **guide users to available data**	Improve the ability of users to identify data sources Data assets incorporate **standard metadata, support APIs, and are search engine optimized**

Metadata, APIs, and Other Data Science Logic Model

- Simplify and facilitate access to and use of microdata

Activities	Outputs	Outcomes
Build a repository of federated data with **standardized and searchable catalogues, metadata** on sources and uses, and links to external sources	Repository of federated data with more standardized, searchable catalogues, metadata on sources and users, and linkages to external sources	Available NCHS microdata are available in a single data repository to approved users, complete with standard and searchable metadata
Develop analyses and visualizations in applications connected to unrestricted microdata for enterprise-wide use	API to federated data repository **Documentation of API with use cases explicitly for visualization**	**Usage of API for research** and public-health related goals **Improved data sharing** and interoperability
Extend enterprise analytics on complex sample survey data files, vital statistics, admin and clinical data sets, and other research outputs	Measures of APIs and other digital outputs **Increased understanding of research outputs** via bibliometrics	Simplify and facilitate access to and use of microdata by responding to user needs Identify gaps where research or data products can be delivered

Virtual Data Enclave Logic Model

- Provide approved researchers with access to restricted-use data

Activities	Outputs	Outcomes
Finalize **partnership with VDE provider**	Partnership established with VDE provider and **VDE becomes operational**	**Improved continuity of data access**
Establish **researcher accounts** in the VDE	Researcher accounts have been established on the VDE and **researchers access data and produce output**	**Improved ability for researchers to produce** approved output

- As CDC and NCHS' Data Modernization Initiatives and related efforts mature, we are actively working to identify new innovations that will directly address some of the long-standing public health challenges

- We are seeking the Board's input, ideas, and suggestions

For more information, contact CDC
1-800-CDC-INFO (232-4636)
TTY: 1-888-232-6348   www.cdc.gov

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.