

Discussion Guide: FDA's Data Enterprise Modernization

2-Day Private Virtual Meeting

July 20-21, 2021

Overview:

On July 6th 2021, the U.S. Food and Drug Administration reinforced its commitment to modernizing the FDA's data and technology infrastructure [FDA's Budget: Data Modernization and Enhanced Technologies | FDA](#). As mentioned in the announcement:

*"The U.S. Food and Drug Administration's mission to protect and promote the public health impacts the life of every American, every day. Our work encompasses an ever-widening array of foods, additives, cosmetics, medicines, and medical devices as the FDA regulates approximately 20 cents of every dollar spent annually by American consumers. The industries we regulate are constantly innovating, attracting skilled professionals, and upgrading their technological capabilities. **The FDA must keep pace.**"*

"Scientific breakthroughs have enabled the development of new, more personalized therapeutic treatments, advanced manufacturing, and modern data solutions such as blockchain, genomic information, and real-time analytics. As a byproduct of these exciting advancements, the amount and variety of data that the FDA generates, needs, and uses is rapidly increasing. But we are often hampered by antiquated methods...."

The FDA's Data Modernization Action Plan (DMAP) [Data Modernization Action Plan \(fda.gov\)](#), which launched in March 2021, is a key foundational framework to guide the FDA's data modernization efforts and to help the FDA "keep pace". Fundamental to modernizing data at the FDA is understanding how the Agency can develop new approaches to ingesting, analyzing, maintaining, and interfacing with data that drive regulatory, policy, and operational decision-making across the Agency. As mentioned in the DMAP:

"Although FDA's legacy technology and data systems allow the Agency to meet its regulatory responsibilities, FDA urgently needs new, robust, and flexible capabilities to avoid losing future opportunities."

The DMAP recognizes that change does not happen overnight rather through iterative improvement. The DMAP will be implemented through 3 pillars – Driver Projects, Data Practices and Talent Development.

The FDA's Chief Data Officer team is focused on driving improvement by creating enterprise level solutions to help resolve Center level challenges. The team actively partners with Centers to identify common data issues and to co-create solutions.

Meeting Objective:

The goal of this two-day workshop is to address the data infrastructure needs for two critical FDA driver initiatives: Non-clinical Trial Data, such as real-world data (RWD), and Supply Chain. The focus of the meeting is to identify challenges and opportunities, translate those insights into needs, and recommend potential data solutions. While policy is important and critical to the FDA's success, this is not a policy focused discussion. This is a discussion on the data infrastructure to support these key initiatives. The

objective of this meeting is to leverage specific use cases from FDA Centers to identify key challenges ingesting data, addressing supply chain issues affecting public health, and to propose solutions (short and long-term) to improve the FDA’s data infrastructure, processes and systems. This includes the full lifecycle of data from identifying, acquiring, ingesting to governance, security and appropriate use for generating actionable insights.

Meeting Scope

In scope	Out of scope
<ul style="list-style-type: none"> • What are the design and resource constraints due to current policy? How does the FDA work within those constraints to build a modern infrastructure? • Infrastructure as <ul style="list-style-type: none"> • technical needs (ex. tools, platforms), • process needs (ex. ability to work across FDA centers or with external partners) • people needs (ex. skillsets, training, hiring) • How does the FDA develop the ability to ingest and analyze data the way it wants to? • How do we build an infrastructure capable of handling both RWD and clinical trials data? • What pilot projects or initiatives can or should the FDA implement to determine infrastructure needs and solutions? 	<ul style="list-style-type: none"> • Policy challenges (ex. Is data fit-for-purpose? How do we determine data reliability?) • Legal challenges (ex. HIPAA or contractual constraints to data sharing)

Overview of Workshop Sessions

Day 1: Introduction

Overview and Discussion of the Data Modernization Action Plan and FDA Data Strategy

In this introduction portion of the meeting, FDA Chief Data Officer Ram Iyer will present an overview of the FDA’s DMAP related plans before being joined by Robert Califf to discuss some of the key questions about FDA’s data strategy that the DMAP seeks to address.

Day 1: Real-World Datasets

Session 1: How the FDA Utilizes Non-Clinical Trial Data

The goal of this session is to leverage use cases to highlight the challenges Centers experience when analyzing and ingesting non-clinical trial data during the regulatory process. The presentations will focus on data issues within the medical space including digital health and

oncology. The session will explore specific data challenges when ingesting, processing and interpreting data with a focus on RWD. While the examples may highlight common RWD issues, this is not a standard RWD discussion. The goal is to leverage the use cases to go beyond the traditional RWD conversation to include action-oriented data infrastructure or process solutions.

Case 1: Opportunities for FDA Infrastructure to Meet Current and Future Regulations

Digital health has an opportunity to collect the evidence of safety and effectiveness to understand the benefits and risks of the products. This is a shift from the information and knowledge we are used to collecting and underpins the concept of pre-certification which builds in real world performance and organizational excellence. An open question is when and how much data needs to be shared with FDA. More information on the pre-cert program can be found at: <https://www.fda.gov/medical-devices/digital-health-center-excellence/digital-health-software-precertification-pre-cert-program>.

- **Example: Current Use of RWD for Market Authorization**
 - Natural Cycles is an over-the-counter web and mobile-based standalone software application that monitors a woman’s menstrual cycle using information entered by the user and informs the user about her past, current and future fertility status. To support the marketing submission, Natural Cycles leveraged OUS real-world data from 15,570 registered users of the app to validate the accuracy of the algorithm in the identification of ovulation.
 - Background Info:
 - https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN170052.pdf
 - <https://www.fda.gov/news-events/press-announcements/fda-allows-marketing-first-direct-consumer-app-contraceptive-use-prevent-pregnancy>

Case 2: Real-World Data for Oncology Drug Approvals

Randomized clinical trials (RCTs) are the evidentiary gold standard for medical product approvals; however, there is an increasing interest in additional sources for evidence generation in oncology, including the use of Real-World Data especially in the context of trial modernization and evidence development within 21st Century Cures Legislation. RCTs can have challenges in eligibility and accessibility, where generalizability can be challenging because of restrictive inclusion criteria, may not include representative population subgroups, and may be limited to patients able to access traditional trials. When RCTs are unethical, infeasible, or there is a lack of equipoise and well as in special populations (rare cancers, pediatrics), the complementary use of real-world evidence may be informative when appropriate and constructed with valid methodological approaches. The increasing availability and submission of RWD requires thoughtful evaluation, characterization, and standardization to facilitate regulatory review. The introduction of the FDA Technology Modernization Action Plan is evaluating ways to build on advances such as RWD to appropriately inform regulatory decision making, build data tools for regulatory science, and communicate with external stakeholders.

Discussion Questions:

- For each use case, what are the operational challenges facing FDA staff that need to use this data? Are there commonalities between the use cases?
- For each use case, what tools or knowledge does the FDA need to further modernization efforts?
- In developing an Enterprise scale service(s), we need to consider center specific regulations (both current and evolving) in designing the data infrastructure. What strategies should the agency adopt in building and maintaining RWD data infrastructure to meet these objectives?

Session 2: Implementing Solutions to Modernize the Use of Non-clinical Trial Data

Building on Session 1, this session will discuss opportunities to address the challenges of utilizing non-clinical trial data. Panelists will highlight potential solutions and approaches the Agency can take to modernize the FDA's data infrastructure and data processes. The discussion will focus on opportunities to implement iterative as well as transformational solutions at the enterprise level.

Focus Area 1: Using Non-clinical Trial Data

Discussion Questions:

- Where can implementation of data standards improve FDA processes?
- Given that the FDA deals with a broad range of data types, what cross-center interoperability standards are needed to improve the agency's capacity for surveillance and other public health data needs?
- In what areas, if any, does the FDA need direct access to large quantities of RWD or other non-clinical trial data?
- What database tools does the agency need to efficiently handle an influx of data?
- What skillsets does the FDA need to develop to fully implement a modern infrastructure?

Focus Area 2: Data Quality

Discussion Questions:

- What kind of information does FDA need to understand data lineage? Does FDA have the capability to ingest this type of metadata efficiently? What internal mechanisms does the FDA need to assess the reliability of submitted data?
- What kind of systems can FDA set up to ensure that reliability assessments are consistent across centers?
- What kind of processes can FDA setup to acknowledge and adapt for data collection algorithms that may change over time?
- What are the alternatives in areas where high quality real-world data is not easily available or not created at the pace required (ex. rare diseases)? What is the role of synthetic data?

Focus Area 3: Projects

- What type of projects or initiatives would be well positioned to address the considerations discussed?

Day 2: Supply Chain and Next Steps

Session 3: Using Data to Monitor Supply Chains

In this session, presenters will highlight use cases to discuss how the Agency's ability to track potential supply chain disruptions, product shortages, and recalls may be improved through data modernization efforts. Three use cases will be presented that explore challenges faced by the FDA related to 1) processing medical product recall data; and 2) tracking medical projects including devices and drugs used during the COVID-19 pandemic. Panelists will discuss data and technical challenges facing current supply chain modernization efforts.

Case 1: Respirator Supply Chain Challenges During COVID-19

In this session, presenters will highlight use cases to discuss how the Agency's ability to track potential supply chain disruptions, product shortages, and recalls may be improved through data modernization efforts and the use of real-time data. Panelists will discuss regulatory and technical challenges facing such a modernization effort. Two use cases will be presented that explore challenges faced by the FDA related 1) processing product recall data and 2) tracking medical devices used during the COVID-19 pandemic.

Case 2: Leveraging Data for Product Recalls

Recalls are one of the most effective ways to protect the public from FDA regulated products that may cause harm. As an agency, we are working to continually improve. How can we process external data more efficiently?

Case 3: Modernizing FDA Drug Supply Chain Data Infrastructure

This presentation will provide a high-level overview of CDER's vision for modernizing the drug supply chain data we analyze and the infrastructure we use for such analysis. Data modernization efforts that will enhance our visibility into each step of the drug supply chain, improve our collection of high quality/structured data, and increase utilization of advanced analytics including the development of predictive models are needed to support a more proactive approach to drug supply chain disruptions.

Discussion Questions:

- What is the breadth of use cases where FDA uses or could use supply chain information?
- How important is real-time access to high-quality data for the FDA?
- To what extent does the FDA face issues of access to sufficient data on supply chains or recalled products?
- What data or data tools does the FDA need to manage risks to medical product supply chains?
- Does the FDA need access to large quantities of raw data for supply chain surveillance?

- Are there ensemble methods that the agency can use instead of running models from raw data?

Session 4: Implementing Solutions to Address Challenges in Utilizing Supply Chain Data

Building on Session 3, panelists will highlight potential iterative and transformational solutions the FDA can implement to advance supply chain data efforts at the Center and enterprise level.

Focus Area 1: Using Supply Chain Data

Discussion Questions:

- Where can implementation of data standards improve FDA supply chain processes?
- Given that the FDA deals with a broad range of data types, what cross-center interoperability standards are needed to improve the agency's capacity for supply chain surveillance and related data needs?
- What tools does the FDA need to develop to fully implement a modernized infrastructure for the discussed use cases (ex. AI or cloud-based solutions)?
- What skillsets does the FDA need to develop to fully implement a modern supply chain data infrastructure?
- What database tools does the agency need to efficiently handle an influx of real-time supply chain data?

Focus Area 2: Data Quality

Discussion Questions:

- How best can the FDA access high-quality, real-time supply chain data?
- What kind of information does FDA need to understand data lineage? Does FDA have the capability to ingest this type of metadata efficiently? What internal mechanisms does the FDA need to assess the reliability of submitted data?

Focus Area 3: Projects

Discussion Questions:

- What type of projects or initiatives would be well positioned to address the considerations discussed?

Session 5: Key Themes and Future Steps

Panelists will react to the conversations from both days of the workshop and explore short- and long-term next steps for the FDA as a whole. Participants will discuss challenges faced across the FDA and consider potential projects for implementing the DMAP in the context of the discussed use cases.

Discussion Questions:

- Are there other use cases that could be helpful to think about as the FDA works on updating the Agency's data infrastructure?
- How can the challenges raised in this workshop be addressed during the implementation process?
- How would fully implementing these use cases at FDA impact current regulatory processes? Is there an incremental approach for phasing in solutions?
- Based on these use cases, what broad capabilities, tools, and skillsets does the FDA need to implement a modern data infrastructure?

- How can the FDA best align cross-Agency to ensure the strategy outlined in the DMAP benefits all centers? What steps can be taken to enable the FDA to efficiently answer questions that rely on data from multiple centers?
- How can we deliver incremental value while building a complex data infrastructure to address these and other driver projects and use cases?
- What type of driver projects would be well positioned to address the considerations discussed?