FDA’s Enterprise Data Modernization Private Workshop
Summary of Key Takeaways

On July 20th and 21st 2021, the Duke-Margolis Center for Health Policy, in partnership with the U.S. Food and Drug Administration, hosted a private workshop entitled FDA’s Enterprise Data Modernization. The workshop convened stakeholders from across the FDA’s centers, data companies, sponsors, and academia to discuss ways in which the FDA might modernize its data infrastructure in accordance with the goals outlined in the Data Modernization Action Plan (DMAP), released in March of 2021.

Key Takeaways

The workshop featured sessions centered around both non-clinical trial data, such as real-world data, and supply chain data. This document details key themes that spanned across the various sessions.

The role of automation: Throughout the workshop, stakeholders highlighted the benefits of automation, including its role in converting data from “human-readable” data to “machine-readable” data. For example, the participants described how medical devices already operate by converting data in a similar capacity; participants theorized on how this might be built upon and operationalized. Additionally, the participants discussed how it may be possible to automate some aspects of data ingestion and evaluation, which could give FDA review staff more time to focus on other important administrative review tasks. The participants also discussed the importance of rendering automated processes interpretable by humans. Decisions made by humans interacting with automated processes should be documented to reveal patterns that might support iterative improvements to automated processes.

Data selection & prioritization: During this workshop, participants also raised the question about what data types the FDA considers most valuable. Participants discussed how the FDA might adapt and build on its current data infrastructure depending on the types of data the Agency would like to collect. For instance, if the FDA wanted to collect more real-time data from wearables, then the Agency could build its capabilities accordingly. The workshop participants also spoke about the purpose and intent behind the FDA’s data collection efforts and error thresholds. For instance, could datasets that are merely being used for additional context, such as demographic data, have a lower threshold for error than data used to inform regulatory decision-making? And what criteria could inform sound decision-making regarding such thresholds?

The need for continued collaboration: Participants highlighted the need for collaboration—mainly the desire for increased partnerships between the government, academia, patient advocates, industry, and other stakeholders. These collaborations could potentially take a format similar to that of the FDA’s Center for Devices and Radiological Health’s (CDRH) Collaborative Communities or to public-private partnerships. Through these collaborations, the FDA could help drive momentum to improve data standards across the broader landscape.

The utility of data standards: Another key theme that arose throughout the workshop was how to ensure that data standards are or can remain consistent across the data landscape. Participants described how the use of common data elements, open data dictionaries, and
taxonomies could help to reduce discrepancies among the data. However, one participant noted that even with consistent data, two statisticians could analyze the same dataset and arrive at starkly different conclusions. Thus, it is also important to consider natural variance in human analysis and interpretation. To further elaborate on the need for data dictionaries, participants provided examples from both a non-clinical trial data perspective as well as a supply chain data perspective on how there can be numerous definitions of the same data point. Two of these examples included 1) how substantially the classification of an FDA-regulated company could vary in a single supply chain database and 2) how COVID-19 definitions differed or had multiple classifications across specific tests or claims data within a single database.

Overall, the workshop served as an introductory step to the enormous undertaking of modernizing the FDA’s data infrastructure. While the FDA’s data modernization efforts are still in early phases, there is a desire among all stakeholders to see the FDA achieve success. As data modernization plans progress, the FDA will be able to take inspiration from both outside organizations and within the FDA to inspire broader progress.

Funding for this workshop was made possible in part by a cooperative agreement (U01FD006807) from the U.S. Food and Drug Administration Center for Drug Evaluation and Research. The views expressed in written workshop materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services nor does mention of trade names, commercial practices, or organizations imply endorsements by the U.S. Government.