

## Advancing the Development and Implementation of Analysis Data Standards: Key Challenges and Opportunities

Tommy Douglas Conference Center • Silver Spring, MD  
Wednesday, June 12, 2019

### Discussion Guide

#### Background

The receipt of standardized, high-quality data is critical to the U.S. Food and Drug Administration's (FDA) ability to efficiently and effectively review medical products. An important component of this review process is ensuring medical research is adequately and appropriately documented. This includes how data are collected, analyzed, and ultimately used to inform regulatory decisions to bring new medical products to the market.

Each year, the Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER) receive more than 150,000 submissions, amounting to millions of data points. The data can arrive in a wide variety of formats and even on paper, complicating the review process and potentially hindering timely approval of new therapies that could benefit patients. To help simplify the process and improve the efficiency and quality of reviews, FDA began requiring that certain submissions be delivered in electronic format and meet specific data standards.<sup>1</sup>

Data standards can help FDA review these submissions, answer critical research questions, and gain key insights. Study data standards describe a standard way to exchange clinical and nonclinical research data between computer systems. They provide a consistent framework for organizing study data, including templates for datasets, standard names for variables, and standard ways of doing calculations with common variables.<sup>2</sup> These standards do not define how scientists should conduct their research, but rather enable the reproducibility of research. This added value not only expedites regulatory review, but also benefits biomedical research in general.

The FDA's Data Standards Catalog<sup>2</sup> lists the data standards and terminologies that FDA supports for use in regulatory submissions. These include the following Clinical Data Interchange Standards Consortium (CDISC) standards relevant to the conduct of clinical studies bulleted below:

- Standard for Exchange of Nonclinical Data (SEND) for nonclinical data
- Study Data Tabulation Model (SDTM) for clinical data
- Analysis Data Model (ADaM) for analysis of clinical data
- Case Report Tabulation Data Definition Specification (Define-XML) for the metadata that accompany SEND, SDTM, and ADaM datasets

Collectively, these standards help FDA receive, process, review, analyze, and archive submissions more efficiently by reducing the effort needed to process less-structured data. When properly implemented they can help ensure the integrity of data submission, traceability to source data, and repeatability of

---

<sup>1</sup> <https://www.fda.gov/media/82716/download>

<sup>2</sup> <https://www.fda.gov/industry/fda-resources-data-standards/study-data-standards-resources>

data analyses. This allows reviewers more time to focus on scientific review and opportunities to combine data from multiple studies to explore new research questions and gain new insights.

FDA has taken a number of steps to promote the use of data standards for regulatory submissions in electronic format, including the publication of binding guidance documents. These documents specify the submission types that must be submitted electronically, supported data standards, exemptions from the guidance, criteria for waivers of the electronic submission requirements, and the implementation process and timetable.<sup>3</sup> The focus of this conference is on the data standards used primarily for submission analysis: SDTM and ADaM. The following guidance documents relate to Analysis Data Standards (ADS):

- Providing Regulatory Submissions in Electronic Format – Submissions Under Section 745(a) of the Federal Food, Drug, and Cosmetic Act<sup>4</sup>;
- Providing Regulatory Submissions in Electronic Format — Certain Human Pharmaceutical Product Applications and Related Submissions Using the eCTD Specifications Guidance for Industry<sup>5</sup>; and
- Providing Regulatory Submissions in Electronic Format — Standardized Study Data.<sup>1</sup>

Important accomplishments with the implementation of data standards for electronic submissions are documented in a report prepared for FDA by Booz Allen Hamilton and published in 2017.<sup>6</sup> This report found that in Fiscal Year 2016, 99% of combined submissions to CDER and CBER were in electronic format. Further, over half of CDER submissions and one third of CBER submissions contained at least one study with SDTM data. This has led to positive review experiences and fewer resources needed to conduct the review and supports the CDER's Office of New Drugs' efforts to implement safety review initiatives that utilize standardized analytics and applications to detect potential signals of interest. This requires data that have been appropriately mapped to key tables, which enables new approaches to organize data for more efficient analysis.

Despite these successes, there remain key implementation challenges that prevent the consistent use of ADS, which impacts FDA's review of evidence submissions. While data standard structures are defined in SDTM and ADaM, the FDA has observed challenges with stakeholders implementing these standards consistently across therapeutic areas. Key issues include the ability to trace data as it is transformed and described using CDISC standard terminology, and then mapped to SDTM and ADaM data structures. Standardized terminology provides a critical foundation for standards implementation, but this terminology is not always consistent across standards. Reducing the variation in how these standards are implemented will not only improve the consistency of submitted study data, but also enable FDA scientists to potentially explore new research questions by combining data from multiple studies as a result of more uniform study data.

---

<sup>3</sup> As of May 5, 2017, new drug applications (NDAs), biologics license applications (BLAs), and abbreviated new drug applications (ANDAs) must be in eCTD format, whereas commercial investigational new drug applications and master files must be in eCTD format as of May 5, 2018.

<sup>4</sup> <https://www.fda.gov/media/88120/download>

<sup>5</sup> <https://www.fda.gov/media/120094/download>

<sup>6</sup> <https://www.fda.gov/media/106231/download>

Additionally, the Agency recognizes the increasing opportunity to leverage real-world data (RWD) -- defined as data relating to patient health status and/or the delivery of health care that is routinely collected from a variety of sources -- for regulatory decision making.<sup>7</sup> However, there is still uncertainty whether and how existing data standards supported by the FDA could accommodate RWD, or if new standards may be needed.

On June 12, 2019, under cooperative agreement with the FDA, the Duke-Robert J. Margolis, MD, Center for Health Policy at Duke University and Critical Path Institute will convene a public workshop to solicit feedback from stakeholders on how to advance implementation of ADS. Discussion will identify and explore implementation and submission challenges, as well as opportunities to enhance the development and use of ADS to improve the predictability and quality of data submissions. This feedback will inform FDA's strategic planning around improving the efficiency of regulatory review, and continued development of the Agency's efforts to support and enable standardized study data for electronic submissions.

## Summary of Workshop Sessions

### Session I: FDA Efforts to Support Analysis Data Standards for Product Development and Review

This session will feature two level-setting presentations for the day's discussion. The first presentation will orient the audience to FDA's review process and the lifecycle of data submissions and highlight Agency guidance and policy development efforts to support greater use of ADS. The second presentation will discuss a clinical reviewer's approach to the data submitted in a marketing application, development of a level 2 guidance pathway for the technical specifications, and how these standards will help new programs to standardize safety review processes in the Office of New Drugs to improve the efficiency and quality of submission reviews.

### Session II: Industry Experience with Data Standards during Product Development and Review

This session will highlight industry experience implementing FDA's finalized guidance for meeting submission requirements using an electronic format. This session will specifically focus on key industry challenges using ADS, and solutions that have been developed. There will be two presentations detailing specific industry experiences implementing the guidance, and panel reactants will further comment on implementation issues. Key discussion topics for the panel include potential difficulties sponsors face using standard structures such as SDTM and ADaM to develop analysis data files and improving the traceability of data as it is transformed and mapped to SDTM and ADaM standard structures for electronically submitting results. Discussion will also explore potential opportunities to reduce the variation of how SDTM and ADaM standard structures are implemented in order to improve consistency and quality of submissions for review as well as to support better data integration of submitted results within a therapeutic area or class of products.

#### *Discussion Questions:*

- An important aim of ADS is to support a more consistent compilation of study results for product submissions. What are the key implementation challenges using ADS such as ADaM and SDTM to ensure and maintain consistency across submissions?

---

<sup>7</sup> <https://www.fda.gov/media/120060/download>

- Does the structure of ADaM or SDTM potentially increase the variability of how study results are presented? What solutions are being implemented to improve the consistency of reported study data?
- How are updates to these standards impacting the ability to standardize study results? What mechanisms are in place for standards development organizations to receive feedback and requests for updates?
- What mapping tools is industry using to trace data from ADS back to source data?
- In the lifecycle of data submissions, are there metrics for sponsor compliance with standard terminology and structure?
- How is the development of CDISC Therapeutic Area User Guides or FDA Technical Specification Documents supporting the implementation of ADS? In what ways have they improved implementation, and are there any implementation barriers that have arisen?
- What training resources and methods are being used to help data managers and data analysts learn best practices for data standardization and submission?

### Session III: Additional Applications and Impact of Data Standards on Clinical Research and Development Outside of Industry

This session will obtain feedback from academic and nonprofit groups utilize ADS as part of basic research, medical product development, or regulatory submission processes. Speakers and panelists will discuss implementation of data standards in original and translation research, challenges that stakeholders face leveraging these standards during the data lifecycle, and solutions being developed to support more systematic implementation.

#### *Discussion Questions:*

- What approaches are currently being taken within research institutions to adopt standard terminologies and data structures across the study and trial landscape?
- What are key considerations for enabling adoption of data standards in research across institutions and stakeholders?
  - Are there opportunities to improve how academia and key stakeholder groups such as sponsors could strengthen collaboration around use of ADS?
  - How might the adoption of standard templates aid such efforts?
- How can professional and patient stakeholder groups contribute to defining relevant gaps in standard terminology for therapeutic areas (markers, measures, endpoints, etc.)?
- Common data elements (CDE) could be useful in developing analysis data sets. How successful have CDE initiatives been in academia and have they contributed to standards development and adoption?
- What incentives would be helpful to increase adoption of FDA-supported data standards in research settings outside of industry?
- For academic institutions and non-profit organizations, what are the barriers to learning and utilizing ADaM?
  - How does the necessary support documentation enhance or reduce these challenges?
  - Are there any barriers associated with developing expertise in both SDTM and ADaM? Does the necessary expertise differ for each standard and how might this impact submissions?
- Does the point at which ADS are implemented impact the standardization of reporting analysis results (e.g., implementing ADS in Phase 1 vs Phase 3) given cost and resource considerations?

#### Session IV: Key Opportunities to Improve the Implementation of ADS

This session will solicit a range of stakeholder perspectives on key opportunities to improve implementation of ADS. This will include FDA's observations of sponsors' use of ADaM standard structures in New Drug Applications, how stakeholders could leverage publicly available resources to assist with implementing data standards and supporting services available to sponsors provided by FDA. The panel will also consider opportunities for the Agency to facilitate innovation and keep pace with changes in technology and data collection. This includes utilizing open source tools for more efficient and cost-effective submissions, and potential lessons learned from key stakeholders on how the Agency could engage and obtain feedback from drug sponsors on data standards implementation.

##### *Discussion Questions:*

- What efforts are underway to synchronize content updates to data standards supported by the FDA?
- What is the utility of existing resources including CDISC Therapeutic Area User Guides or FDA technical specifications for implementing ADS?
- What guidance is needed from FDA to support more consistency and higher quality sponsor submissions which incorporate ADS?
- What are the opportunities to improve implementation of beginning-to-end life sciences and translational research during drug development?
- How can ADS be incorporated into study designs to reduce the effort needed to submit analyses in supported formats?
- How might FDA-sponsor engagement earlier in the drug development program enhance the incorporation of ADS in study design?
- How can the further development of other CDISC data standards such as the Clinical Data Acquisition Standards Harmonization (CDASH), and SEND, improve implementation of ADS?
- How could collaborations be fostered to support uptake of commercially available open source tools and CDISC standardized tools?

#### Session V: Emerging Trends and Innovations for the Development and Use of ADS

Given growing interest at the Agency to better leverage RWD and resultant real-world evidence (RWE) for regulatory decision-making, this session will begin to consider how standards could support submissions that include RWD and RWE. RWD may have varying levels of curation and standardization relative to the structured data collected in traditional clinical trials. The panel will refer to current experience collecting and submitting RWD/RWE to the Agency and will consider whether existing ADS could be utilized or modified for these types of submissions. In addition, the panel will discuss potential opportunities to incorporate or modify data standards that are not currently part of FDA's standards catalog, including standards already emerging for RWD, to support such submissions.

##### *Discussion Questions:*

- What data and submission standards should a sponsor be expected to follow when submitting RWD/RWE to FDA?
  - What is the breadth of healthcare data available that could support evidence submissions utilizing RWD and RWE?
  - A key advantage of using RWD/RWE is the ability to potentially link disparate sources to identify insights not possible when relying on a single data source for the analysis. What types of RWD might be appropriate for linkage and how could ADS support linkages?

- How can health information exchange data standards such as those developed by HL7 and CDISC meet the RWD needs of FDA for submission of evidence packages?
- Presently, the data standards catalogue only identifies the SDTM and ADaM standards supported by FDA. Can these standards accommodate RWD or would data model extensions be needed to incorporate a more diverse range of data?
- Are there any standards outside of the FDA data standards catalogue that should be considered? If so, how could the transition from a fully mature and adopted standard for FDA best be managed?
- Is there an opportunity to develop electronic data capture standards to automatically extract RWD from electronic health records into analysis data models?
- What initiatives are underway to develop more standardized terminology across RWD sources such as patient registries and electronic health records?
- What examples exist of RWD being successfully incorporated into FDA data submissions?

Funding for this workshop was made possible in part by a cooperative agreement 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.

## Appendix: Suggested Reading

1. [Shankar RG, Battacharya S, et al. RImmPort: An R/Bioconductor package that enables ready for analysis immunology research data Bioinformatics. 33\(7\), 2017, 1101-1103.](#)
2. [Sim, I; Carini S; Tu, S et al. The Human Studies Database Project: Federating Humans Studies Design Data using the Ontology of clinical Research. Summit on Translat. Bioinforma. 2010:51-55.](#)
3. [Van Valkenhoef G, Tervonene T, De Brock B, Hillege H. Deficiencies in the transfer and availability of clinical trials evidence: a review of existing systems and standards. BMC informatics and Decision Making 2012, 12:95.](#)
4. [Tu, SW; Peleg M, Carini S, Bobak S, Ross, J; Rubin D; Sim I. A practical method for transforming free text eligibility criteria into computable criteria. J of biomedical Informatics 44 \(2\) :239-250. 2011.](#)
5. [Hudson, L; Kush, R; Navarro E et al. Global Standards to Expedite Learning from Medical Research Data. Clin Transl Sci \(2018\) 11, 342–344; doi:10.1111/cts.12556](#)
6. [Booz Allen Hamilton \(2017\).](#)
7. [U.S. Food and Drug Administration. \(2018\). Framework for FDA’s Real-World Evidence Program.](#)
8. [Mulberg AE, Bucci-Rechtweg C, Giuliano J, Jacoby D, Johnson FK, Liu Q, Marsden D, McGoohan S, Nelson R, Patel N, Romero K, Sinha V, Sitaraman S, Spaltro J, Kessler V. Orphanet J. Regulatory strategies for rare diseases under current global regulatory statutes: a discussion with stakeholders. Rare Dis. 2019 Feb 8;14\(1\):36.](#)
9. [Stegall MD, Troy Somerville K, Everly MJ, Mannon RB, Gaber AO, First MR, Agashivala N, Perez V, Newell KA, Morris RE, Sudan D, Romero K, Eremenco S, Mattera M, Spear N, Porter AC, O’Doherty I. The importance of drug safety and tolerability in the development of new immunosuppressive therapy for transplant recipients: The Transplant Therapeutics Consortium’s position statement. Am J Transplant. 2019 Mar;19\(3\):625-632.](#)
10. [Perrone RD, Mouksassi MS, Romero K, Czerwiec FS, Chapman AB, Gitomer BY, Torres VE, Miskulin DC, Broadbent S, Marier JF. Total Kidney Volume Is a Prognostic Biomarker of Renal Function Decline and Progression to End-Stage Renal Disease in Patients With Autosomal Dominant Polycystic Kidney Disease. Kidney Int Rep. 2017 Jan 16;2\(3\):442-450.](#)
11. [Conrado DJ, Karlsson MO, Romero K, Sarr C, Wilkins JJ. Open innovation: Towards sharing of data, models and workflows. Eur J Pharm Sci. 2017 Nov 15;109S:S65-S71.](#)
12. [Larkindale J, Abresch R, Aviles E, Bronson A, Chin J, Furlong P, Gordish-Dressman H, Habeeb-Louks E, Henricson E, Kroger H, Lynn C, Lynn S, Martin D, Nuckolls G, Rooney W, Romero K, Sweeney L, Vandenborne K, Walter G, Wolff J, Wong B, McDonald CM, Duchenne Regulatory Science Consortium Imaging-Dmd Consortium And The Cnrg Investigators MO. Duchenne Regulatory Science Consortium Meeting on Disease Progression Modeling for Duchenne Muscular Dystrophy. PLoS Curr. 2017 Jan 12;9.](#)
13. [Haas M, Stephenson D, Romero K, Gordon MF, Zach N, Geerts H; Brain Health Modeling Initiative \(BHMI\). Big data to smart data in Alzheimer’s disease: Real-world examples of advanced modeling and simulation. Alzheimers Dement. 2016 Sep;12\(9\):1022-1030.](#)
14. [Jauregui B, Hudson L, Becnel L, et al. The Turning Point for Clinical Research: Global Data Standardization. Applied Clinical Trials 2019.](#)