Public Meeting Regulatory New Drug Review: Solutions for Study Data Exchange Standards

Meeting Summary

Food and Drug Administration

White Oak, MD

November 5, 2012
10am – 4pm

1. Background

On November 5, 2012 The FDA held a public meeting entitled "Regulatory New Drug Review: Solutions for Study Data Exchange Standards" the purpose of which was to solicit input from industry, technology vendors, and other members of the public regarding the advantages and disadvantages of current and emerging open, consensus-based standards for the exchange of regulated study data. FDA also sought input from stakeholders and other members of the public on this topic and a set of pre-meeting questions discussed below. The federal register notice (FRN) and link to feedback received on the pre-meeting questions is available at www.regulations.gov by searching for Document ID FDA-2012-N-0780-0001.

2. Meeting Agenda

The meeting agenda was developed based on the input received in response to the FRN and the goal of the meeting was to obtain input from stakeholders and create an environment for open discussion. The speakers were scheduled on a "first-come, first-serve" basis and availability of meeting time and approximately 1.5 hours was dedicated for discussion in the afternoon. The meeting was organized in roughly four (4) groups:

1. Meeting Introduction and Overview
2. Part 1 – FDA Review of Current Environment and Challenges
3. Part 2 – Study Data Exchange Solutions
4. Open Discussion

A copy of the agenda is in Appendix A.

The following sections provide an executive level summary of each agenda topic grouping.

3. Meeting Introduction and Overview

FDA received over 28 comments from Industry, Academia, standards development organizations (SDOs), clinical research organizations (CROs), Consultants, and Technology Vendors. All the comments received are available from the Docket listed above.

Summary of the meeting drivers[1]:

[1] Presentations Part 1: Mary Ann Slack Welcome Public Meeting: Solutions for Study Data Exchange Standards

- FDA supports the American Standard Code for Information Interchange (ASCII)-based SAS Transport (XPORT) version 5 file format. It has served its purpose well, but…
    - XPORT v5 is not an extensible modern technology
    - Known limitations are causing technical issues; we anticipate these to become increasingly challenging
- Event and data relationships are not currently well captured.

In working towards a solution:

- Public input on potential replacement solutions and experiences
    - What solutions are out there?
    - Are they "fit for purpose" today? In 1 year? In 5 years?
    - What should we know before we evaluate?
- Large scale changes cannot be made immediately
- FDA will complete an evaluation to determine the cost and benefit to both FDA and regulated Industry of any migration to a new exchange format
    - Any solution will need to meet FDA requirements

To facilitate characterizing requirements, FDA is providing draft scenarios to illustrate data needs (see link in Section 7).

4. Executive Summary of Part 1 – FDA Overview of Current Environment

The limitations of the current exchange format SAS Transport (XPT) v5 were discussed. They generally fall into two categories:

1. **Technical** – Character limitations for variable names, labels, character fields, large file sizes due to "empty space"[2].
2. **Structural** – Two-dimensional "flat" data structure for hierarchical/multi-relational "round" data; lack of a robust information model with the standard. Important meaning is lost when exchanging 2-dimentional flat files, making some interpretations and analyses difficult or impossible.[3]

An overview of the high-level, exchange, functional needs to enable a modern review environment was provided.[4] These functional requirements include:

- The ability to have an audit trail to enable reviewers to understand what was done with the data. For example, data is collected from sources (e.g., case report forms) and placed into databases where clean up and reconciliation often occurs. It is challenging or impossible to understand these data management activities and their implications on the results.

[2] Presentations Part 1: 1. Doug Warfield Problems/Challenges Faced within Current Environment and General Requirements

[3] Presentations Part 1: 2. Armando Oliva Challenges of Current Study Information Exchange Format

[4] Presentations Part 1: 3. Chuck Cooper CDER CSC Functional Needs for a Modern Review

Analysis specific audit trails are important too. Understanding the approach taken with an analysis provides reviewers a greater understanding of the analysis output.

- Greater flexibility to adapt to new content requirements. Study content requirements and content standards continue to evolve rapidly, so to the extent an exchange standard can incorporate new content requirements and standards should be considered (e.g., minimal need to modify implementation guides, information systems, data management and analysis tools).
- Support for data integration. It is inevitable that reviewers will want to conduct additional analyses so the need to integrate data is critical.
- Robust metadata needed for reviews. Reviewers rely heavily on the define file to ensure they understand each data element. If the quality of that metadata is poor, it slows down the review. An exchange solution needs to support robust metadata exchange so that the data are understandable to reviewers and tools.

5. Executive Summary of Part 2 – Study Data Exchange Solutions

Overall, there were five options presented for replacing the current exchange format, SAS Transport v5:

1. SAS Transport v5 extensions
2. Clinical Data Interchange Standards Consortium (CDISC) Operational Data Model (ODM)
3. Health Level Seven (HL7) v3 including Clinical Document Architecture (CDA)
4. Semantic Web (Resource Description Framework (RDF); Web Ontology Language (OWL))
5. Analytical Information Markup Language (AnIML)

- Bill Gibson (SAS) presented new capabilities available to address field name size, field name characters, field label size, and character value size available in the SAS v5 Transport File w/Extensions. [5]
- Melissa Binz and Peter Mesenbrink (Novartis) discussed the importance of controlled terminology management, good communication and flexibility as clinical data standards evolve, effective standards adoption policies (e.g., ensuring sufficient advance public notice), having greater consistency across review divisions on standardized data needs (including any legacy data conversions) during the product lifecycle. [6]
- Mathias Brochhausen (University of Arkansas) [7] and Charlie Mead (W3C) [8] discussed the semantic web approach which focuses on the semantics; one can convert existing standards (CDISC, HL7, Systematized Nomenclature of Medicine (SNOMED)) to OWL and expose the semantics in a way that can help solve information exchange problems very

[5] Presentations Part 2, Bookmark: 1. Bill Gibson Extended SAS v5 Transport

[6] Presentations Part 2, Bookmark: 2. Peter Mesenbrink, Melissa Binz Solutions for Study Data Exchange Standards

[7] Presentations Park 2, Bookmark: 3. M Brochhausen W.R.Hogan Study Data Exchange and Open Source Ontologies

[8] Presentations Part 2, Bookmark: 5. Charlie Mead W3C Solutions for Study Data Exchange Standards

easily and more quickly. Initial solution can quickly focus can be placed on a small set of requirements and expand the solution over time.

- Gary Kramer (American Society for Testing and Materials (ASTM)) presented how XML-based AnIML can be used to exchange information and results for any study. Currently it's used for studies of analytic instruments (e.g., mass spectrometers) but can easily be extended.[9]
- Armando Oliva (FDA) gave an update of the HL7 v3/CDA research and development (R&D) efforts. CDER is conducting a proof of concept to test HL7 Study Data Standards which is in parallel with ongoing robust CDISC implementation efforts and Therapeutic Area (TA) standards development. There is incremental testing of: Phase 1: Patient Narrative / Clinical Investigator Info (1572), Phase 2: Structured Protocol, Phase 3: Subject Data.[10]
- There were two presentations focused on CDISC ODM: David Gemzik (Medidata) [11] and Wayne Kubick[12]. David Gemzik described the implementation of ODM throughout the service environment at Medidata. Wayne Kubic reviewed ODM as a vendor neutral XML Schema for exchange and archive of clinical trials metadata and data: snapshots, updates, archives. ODM is an XML wrapper that can be used to exchange any kind of study data and is already used by vendors. Can capture relationships that are protocol-defined. It is human and machine readable.
- Fred Wood (Octagon Research Solutions/Accenture) highlighted feedback to the questions posed in the federal register notice. He highlighted that ODM has been demonstrated to be an efficient mechanism for reusing metadata across the data lifecycle (i.e., from study design/set up to submission), there is much more experience and knowledge of CDISC standards than those of HL7 in clinical areas, implementation of an exchange standard needs to occur in phases and support of multiple versions of an exchange standard should be evaluated.[13]
- Diane Wold (GSK) highlighted the need to address short term solutions to SAS Transport v5 limitations and long term solutions to data relationships. Working in partnership and being clear on the problems to be solved is needed.[14]

6. Summary of Afternoon Discussion

The afternoon session was a facilitated discussion session about the different solutions presented and thoughts on next steps for FDA.

General Discussion

[9] Presentations Part 2, Bookmark: 4. Gary Kramer Analytical Information Markup Language

[10] Presentations Part 2, Bookmark: 6. Armando Oliva HL7 Study Data Standards Project

[11] Presentations Part 2, Bookmark: 7. David Gemzik Medidata and Open Industry Standards

[12] Presentations Part 2, Bookmark: 8. Wayne Kubick Using CDISC ODM for Study Data Exchange

[13] Presentations Part 2, Bookmark: 9. Fred Wood Solutions for Study Data Exchange Standards

[14] Presentations Part 2, Bookmark: 10. Diane Wold Solutions for Study Data Exchange Standards

- There was interest in exploring SAS v5 extensions as a short term solution for technical limitations to the current format. It's clear that it would not solve the structural limitations and a longer-term solution would also need to be identified. Attendees discussed that this would be a lower level of effort to assess as a short term solution but more information is needed.
- There was strong support from the audience to see FDA conduct an ODM pilot.
- Attendees discussed ODM as good option but it was clear that there are known challenges to be addressed (e.g., lack of information model, not ISO 21090 compliant, problem with relationships).
- Several attendees throughout the session indicated there is a need to understand FDA's requirements to ensure a clear understanding of what problem the FDA is looking to address. Also there was interest in knowing what pilots are ongoing at FDA.
- Clear business requirements are needed so that alternatives can be assessed objectively and a decision can be made. Metrics from a pilot or a comparative pilot are also important to assess the economic impact of a change.
- An attendee discussed the challenges in identifying available resources with the necessary expertise for HL7 implementations.
- One attendee expressed that we should harmonize with EHR standards (to leverage the investment in HL7 in healthcare and enrich the Study Data Tabulation Model (SDTM) content), and minimize data management and exchange burden with investigators and electronic health record (EHR) systems/vendors.
- A group of attendees discussed the continued need to accommodate legacy data and the challenges with standardizing these data.
- Numerous speakers advocated not using HL7, referencing its complexity and lack of experience with the standard, and the ongoing investment made with CDISC implementations.
- There was general interest in semantic web, but more information is needed to better understand its potential use.
- There was feedback that Biomedical Research Integrated Domain Group (BRIDG) was not discussed enough and its role in a solution.
- When discussing the burden of implementing a change, attendees expressed the need for clarity of FDA's requirements; that doing work partially across multiple standards is not desirable, and that time would be needed to identify vendors or to implement a solution based on new requirements.

7. Next Steps

FDA will post presentations and meeting summary on http://www.fda.gov/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/ucm249979.htm.

FDA will also publish study information exchange scenarios to illustrate some of the challenges associated with study information exchange. These can be used to develop more formal requirements to assess potential solutions.

Appendix A. Meeting Agenda

Center for Drug Evaluation and Research

Regulatory New Drug Review: Solutions for Study Data Exchange Standards

Public Meeting Agenda
White Oak Campus
November 5, 2012

Time		
10:00	**Welcome and Introductory Remarks**	Mary Ann Slack Deputy Director, CDER
10:15	**FDA Drivers for this Meeting Topic**	Mary Ann Slack
10:30	**Discussion – Problems/Challenges Faced Within Current Environment and General Requirements** **Speaker 1**. Doug Warfield (FDA) **Speaker 2**. Chuck Cooper (FDA) **Speaker 3**. Armando Oliva (FDA)	
11:30	**Discussion – Data Exchange Standards and their Advantages and Disadvantages** **Speaker 1**. Bill Gibson (SAS) **Speaker 2.** Peter Mesenbrink and Melissa Binz (Novartis) **Speaker 3**. Mathias Brochhausen and William Hogan (University of Arkansas)	
12:00	**Lunch**	
1:00	**Continue Discussion – Data Exchange Standards and their Advantages and Disadvantages** **Speaker 4**. Gary Kramer (ASTM Subcommittee) **Speaker 5**. Charlie Mead (W3C) **Speaker 6**. Armando Oliva (FDA) **Speaker 7**. Dave Gemzik (Medidata Solutions) **Speaker 8**. Wayne Kubick (CDISC) **Speaker 9**. Fred Wood (Octagon) **Speaker 10.** Diane Wold (GSK) **35 min Open Discussion**	
2:45	**Break**	
3:00	**Continue Open Discussion**	
3:50	**Summary & Next Steps**	Mary Ann Slack
4:00	**Meeting Adjourned**	