Abstract

- This tutorial describes the CDISC standards (SDTM, ODM, ADaM, LAB, define XML and protocol), demonstrating how the models can be leveraged to achieve the true eClinical trial.
- The tutorial details, at a practical level, the flow of information using the standards from protocol setup through data capture, analysis and onwards to submission.
Aknowledgements

- CDISC Standards are developed by groups of volunteers and it would be impossible to name them all here, but we would like to thank them here for the great job they have done.
- This tutorial uses a number of slides developed by important CDISC contributors: Dave Iberson-Hurst, Diane Wold, Philippe Verplancke, Julie Evans and Frank Newby. Many thanks to them!

Learning Objectives

At the conclusion of this tutorial, participants should be able to:

- Discuss the basics of the SDTM, ODM, define .xml, LAB and ADaM standards.
- Explain the CDISC standards and their value to eClinical trials.
- Describe the data flow, using the CDISC standards, from clinician to submission.
- Explain how to leverage the standards to improve regulatory compliance.

Agenda

- Pierre-Yves Lastic
  – CDISC : End-to-End Overview
  – Protocol, TDM, CDASH, ODM and LAB
  – eCRF setup, data capture and mapping
- Steve Wilson
  – Regulatory Background
  – Submission and Review
  – SDTM, ADaM and define.xml
CDISC Snapshot

- **Global, open, multi-disciplinary non-profit organization**
  - Founded in 1997; incorporated in 2000
  - Liaison A Status with ISO TC 215
  - Charter agreement with HL7 since 2001
  - ~ 250 member organizations
  - Active Coordinating Committees
    - Europe, Japan, China, Korea
  - Additional activities
    - Australia, India, S. America and Africa

- **Established industry standards to support the electronic acquisition, exchange, submission and archiving of data to support regulated clinical research**
  - Freely available on the CDISC website (www.cdisc.org)
  - Developed through open, consensus-based approach

Clinical Information Flow
The CDISC Way

![Clinical Information Flow Diagram]

- Protocol
- Form Setup & Config
- Data Capture
- Mapping
- Analysis
- Submission
- Review

- CDASH
- ODM
- SDTM(SEND) & ADaM
- LAB
- XML
- Controlled Terminology

Review of the Standards
Overview

Protocol & BRIDG

Sponsor

Investigator

CRO

Subject

LABs

ODM

Archive

Operational Data Model

CRO

Sponsor

ODM

Archive

Subject

LABs

ODM

SDTM

ADaM

ODM

Define.XML

• Exchange & Archive of clinical data
• Production Version 1.3
• XML Schema

Original Use Cases

• Data Interchange – Transfer of information between two or more parties than maintains the integrity of the contents of the data.
• Data Archive – Long term storage of files that are no longer in active use
Other Use Cases

- Set up of systems
- Acquisition
  - eCRF
  - ePRO
  - EHR
- eSource
- Trial Registry
- Metadata Submission
  - Define.xml

Laboratory Data Model

- Exchange of LAB data
- Production Version 1.0.1
- Implementations through SAS, ASCII, XML/ODM and HL7 V3 RIM message

Use Case

- Support the bulk transfer of laboratory data
Study Data Tabulation Model

- Submission data (Case Report Tabulations; analysis data)
- SDTM Production Version 1.2, with Implementation Guide V. 3.1.2 (November 12, 2008);
- Referenced as a specification in FDA Guidance - 21 July 2004; updated – 30 October 2009

Analysis Dataset Models

- Analysis Data Model Version 2.1 and Implementation Guide Version 1.0, December 17, 2009

Terminology

- Data Tabulations = SDTM data
- Analysis Datasets = ADaM data

- Two sets of data, both are representations of the clinical trial data
- Each with a specific purpose
SDTM & ADaM Datasets

- **SDTM**:  
  - observations from a clinical trial  
  - are particularly useful in medical officer evaluation of safety (with appropriate tools)
- **ADaM**:  
  - restructured and contain additional information (derived variables, flags, comments, etc.)  
  - analysis-ready

Source: Susan Kenny, Inspire Pharmaceuticals Inc

CDASH

- FDA Critical Path Opportunity #45  
- Continues ACRO’s CRF Standardization Initiative  
- Goal: To develop a set of ‘content standards’ (element name, definition, metadata) for a basic set of global data collection fields that will support clinical research studies. The initial scope will be the ‘safety data domains’ to support clinical trials.

Protocol Representation

- HL7–CDISC–NCI Collaboration  
- Objective to develop a standard, structured, machine-readable clinical protocol representation
Main Protocol Use Cases

- To support CDISC Study Data Tabulation Model (SDTM)
  - Trial Design
  - Planned Assessments
  - Planned Interventions
  - Inclusion/Exclusion criteria
  - Statistical Analysis Plan
- To support study tracking databases, e.g. EudraCT, clinicaltrials.gov, or other trial registry or results databases, or databases that support project management tools
- To support the development of the clinical trial protocol document

Source: Protocol Team, CDISC

The BRIDG Model

- **Vision**: Create a domain analysis model for clinical research domain
  - **Key Goals**:
    - to harmonize clinical research standards among each other – i.e CDISC Standards
    - to harmonize standards between clinical/medical research and healthcare

Source: BRIDG Team, CDISC
Biomedical Research Integrated Domain

Protocol --> CDAS (H LAB) --> SDTM (SEND) --> ADaM

5TH ANNUAL CLINICAL FORUM BASEL 2011

Terminology

- Covers the work of all teams
Terminology Collaboration

EVS = NCI Enterprise Vocabulary Services

High Level Outline

Protocol
  Set Up
  Execute
  Analysis
  Submission

Conventions

SDTM Tabulations = CDISC Standard (content)

ODM XML = CDISC Standard (physical form) used to transport the content using World Wide Web Consortium (W3C) XML Standard
Protocol

Capture

Interchange
The CDISC Blueprint

The Same Picture

Protocol
PR Standard Hierarchy

- Top level sections from ICH E6 are shown as grey lines
- Next hierarchical level shown as light blue/aqua lines.
- Elements in each sub-section are in clear/white lines.
- The elements re captured in a spreadsheet & linked with definitions, sources, cardinality & other information

Sections

- Document Type
- General Information
- Background Information
- Trial Objectives and Purpose
- Trial Design
- Subject Selection and Withdrawal
- Subject Participation Study Design
- Treatment of Subjects
- Efficacy Assessments
- Assessment of Safety
- Statistics
- Direct Access to Source Documents
- Quality Control and Quality Assurance
- Ethics
- Data Handling and Record Keeping
- Financing and Insurance
- Publication Policy
- Supplements
Clinical Trial Register Elements

<table>
<thead>
<tr>
<th>Protocol Title</th>
<th>Clinical Trials Activities: Study.longTitle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protocol Short Title</td>
<td>Clinical Trials Activities: Study.shortTitle</td>
</tr>
<tr>
<td>Protocol Identifier</td>
<td>Clinical Trials Activities: Study.id</td>
</tr>
<tr>
<td>Clinical Trials Phase</td>
<td>Clinical Trials Activities: Study.phaseCode</td>
</tr>
<tr>
<td>Study Synopsis</td>
<td>Clinical Trials Activities: Study.description</td>
</tr>
<tr>
<td>Participation Type</td>
<td>Clinical Trials Activities: Study.multiInstitutionInd</td>
</tr>
<tr>
<td>Trial Status</td>
<td>Clinical Trials Activities: Study.status</td>
</tr>
<tr>
<td>Target study population description</td>
<td>Clinical Trials Activities: Study.populationDescription</td>
</tr>
<tr>
<td>Target Disease Condition</td>
<td>Clinical Trials Activities: Study.targetConditionCode</td>
</tr>
<tr>
<td>Date of First Enrollment</td>
<td>Clinical Trials Activities: PlannedStudy.TBD</td>
</tr>
<tr>
<td>Duration of Subject Participation</td>
<td>Clinical Trials Activities: PlannedStudy.plannedSubjectParticipationDuration, PlannedStudy.plannedSubjectInterventionDuration</td>
</tr>
<tr>
<td>Targeted Accrual</td>
<td>Clinical Trials Activities: PlannedStudy.targetAccrualNumber</td>
</tr>
<tr>
<td>Study Purpose</td>
<td>Clinical Trials Activities: Study.intentCode</td>
</tr>
<tr>
<td>Study Investigation Type</td>
<td>Clinical Trials Activities: Study.TBD</td>
</tr>
</tbody>
</table>

Trial Design Model

- Led by Diane Wold, GSK - from SDTM Team
- Allows description of key aspects of the planned conduct of a clinical trial in a standardized way
  - The planned arms of the trial
  - What happens to a subject in each arm
  - The planned schedule of visits
  - The inclusion and exclusion criteria for the trial
Example: Crossover Trial

Screen

Randomization

- Placebo
- Rest
- 5 mg
- Rest
- 10 mg
- Follow up

- 5 mg
- Rest
- Placebo
- Rest
- 10 mg
- Follow up

- 5 mg
- Rest
- 10 mg
- Rest
- Placebo
- Follow up

Example: Open Trial with Dissimilar Arms

Screen

Randomization to Chemo + Radiation or Chemo + Radiation + Surgery

- Induction Chemo + RT
- Additional Chemo
- Off Treatment Follow-up

- Induction Chemo + RT
- Additional Chemo + RT Boost
- Off Treatment Follow-up

- Evaluation for disease progression
- 3-5w Rest
- Surgery
- Additional Chemo
- Off Treatment Follow-up

- 4-6w Rest
- Additional Chemo
- Off Treatment Follow-up

- Follow-up

Trial Design Model in XML

- As ODM-extension incl. extension XML-Schema

- Easy to map to SDTM
  - i.e. automated conversion to SDTM tables possible

- Easy to understand for toolmakers and technology vendors
  - It must be easy for vendors to create software tools that export a trial design as TDM-XML

- Easy to visualise
TDM in XML – an example

TDM-XML Rendering

SDTM

Table of Activities

TDM-XML

HTML

SVG Images
Forms Setup and Configuration

1. ACRO Standard Form
2. CDISC SDTM Standard
3. ACRO Form + CDISC SDTM Standard = Annotated Form
4. Annotated Form + ODM Standard = Standard electronic metadata (XML)
5. Standard electronic metadata configures collection system
The Business Benefit of an ODM Based Study Specification Process

- Study Specification time reduced
  - Study CRF requirements fell from 300 hrs to 110 hrs
- Reduction in review and approval times
  - Requirements review per customer fell from 48 hrs to 12 hrs per person
- Total 492 hrs down to 158 hrs (3x faster)
  - Doesn’t include time saved on database build
  - Doesn’t include time saved on database test
  - Doesn’t include downstream time benefits
    - Machine readable spec will promote automation and other opportunities
Project Snapshot

• Addresses Critical Path Opportunity #45 – Streamline data collection at investigative sites.
• Continuation of ACRO’s Initiative
• Started October 2006
• Supported by a Collaborative Group of 17 organizations
• Core team of 16 members manages...
  - 11 working groups
  - Comprised of between 8-40 volunteers
• ~190 working group volunteers

• 16 Safety data domains developed.
• Consolidated document posted for Public review in May 2008.
• Received over 1800 comments from 46 organizations.
• All 3 ICH Regions were represented in the public comment process.
  - US
  - Europe
  - Japan

History

• March 16, 2004 FDA white paper:
  – “Innovation/Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products”
  – Describes urgent need to modernize the medical product development process – the Critical Path – to make product development more predictable and

History

Association of Clinical Research Organizations (ACRO)

- Suggested standardizing CRFs as one small way to increase efficiency in drug development
- ACRO Ethics and Clinical Practice Committee suggests AE and CM as core safety issues and reasonable first “standardized” CRF products (11/19/04)
- 10 member companies provide sample AE and CM forms to ACRO (12/6/04)
- AE and CM forms and ‘guidance’ shared with PhRMA, BIO, ACRP, others (3/29/05)

History - Handoff to CDISC

- Jan 2006 - CDISC requested to take leadership role
- June 2006 – Initial Collaborative Group (10) announced by Dr. Woodcock at Annual DIA Meeting in Philadelphia
- October 2006 – CDASH Collaborative Project Kickoff - ~ 80 volunteers

CDASH Purpose & Scope

- To develop a set of ‘content standards’ (element name, definition, metadata) for a basic set of global industry-wide data collection fields that support clinical research.
- The initial scope - ‘safety data/domains’
- These safety domains cut across all therapeutic areas (TA independent)
- . . . and make certain that all SDTM “required” variables are addressed AND that all CDASH collection fields map into the
Collaborative Group Members

- American Medical Informatics Association (AMIA)
- Association of Clinical Research Organizations (ACRO)
- Association of Clinical Research Professionals (ACRP)
- Baylor College of Medicine
- Biotechnology Industry Organization (BIO)
- Clinical Data Interchange Standards Consortium (CDISC)
- Clinical Research Forum
- Critical Path Institute
- Duke Clinical Research Institute (DCRI)
- Society for Clinical Data Management (SCDM)
- Food and Drug Administration (FDA)
- NIH - NCI - caBIG
- NIH - Clinical Research Policy Analysis & Coordination Program
- National Clinical Research Resources (NCRR)
- NIH - National Institute of Child Health & Human Development (NICHD)
- National Library of Medicine (NLM)
- Pharmaceutical Research and Manufacturers Association (PhRMA)

Who Participated?

Team Membership:

- Statisticians
- Medical Monitors / Clinical Scientists
- Regulatory Affairs
- Drug Safety
- Data Managers
- Clinical Study Coordinators
- Clinical Research Associates
- Investigators
- Clinical Program Managers
- Statistical Programmers
- Database programmers

International Collaboration

- The Association of Clinical Data Management (ACDM)
- The International Network of Clinical Data Management Associations (INCDMA)
- The French Association for Statistics and Data Management (DMA)
- Dutch Association for Statistics and Data Management (PSDM)
- All 3 ICH Regions were represented in the public comment process.
  - US
  - Europe
  - Japan
How was CDASH Developed?

• Started with Study Data Tabulated Model (SDTM)
• Focused on CRF Content, not CRF Layout
• Referred to ACRO CRF Samples
• Collected CRF samples
• Evaluated commonalities/differences of CRF samples
• Documented data points included/excluded with justifications

How was CDASH Developed?

• Agree on basic data collection fields
• Map to SDTM
• Terminology - proposals shared with the Terminology Team
• Write definitions and completion instructions for clinical site and Sponsors
• Proceed to the next step in the Consensus Process

CDASH Delivers Domain Tables

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Collection Field</td>
<td>Variable Name (CDASH variable name shaded)</td>
<td>Definition</td>
<td>Case Report From Completion Instructions</td>
<td>Additional Information for Sponsors</td>
<td>CDASH Core Designations</td>
</tr>
</tbody>
</table>

Can be either a topic description or the Question text.
## CDASH V1.0 - Table of Contents

### Sections

1. Orientation
   - Purpose
   - Organization of Document
2. CDASH Alignment with Other Standards
   - SDTM
   - CDISC Controlled Terminology
   - Other Standards (Beyond CDISC)
3. Best Practice
   - Introduction to Best Practices
   - Recommended Methodologies for Creating Data Collection Instruments
   - Recommended CRF Development Workflow
   - FAQs on Best Practices for Creating CRF Content and Structure
4. Overview of CDASH Domain Tables
   - Introduction
   - Data Collection Fields Considered not Necessary to Collect
5. CDASH Domain Tables
   - Common Identifier Variables
   - Common Timing Variables

### Appendices

7.1 Commonly Used CDISC Controlled Terminology
7.2 Regulatory References
7.3 CDASH Project Development Process
7.4 CDASH Core Team Members and Participating Companies
7.5 List of Abbreviations and Glossary
7.6 Acknowledgements
7.7 Revision History
7.8 Intellectual Property

### Domains Overview

- Common Identifier Variables
- Common Timing Variables
- Adverse Events (AE)
- Concomitant Medications (CM)
- Comments (CO)
- Drug Accountability (DA)
- Demographics (DM)
- Disposition (DS)
- Protocol Deviations (DV)

- ECG (EG)
- Exposure (EX)
- Inclusion Exclusion (IE)
- LAB Test Results (LB)
- Medical History (MH)
- Physical Exam (PE)
- Vital Signs (VS)
- Subject Characteristics (SC)
- Substance Use (SU)

### Core Designations

- **Highly Recommended**
  - A data collection field that should be on the CRF (e.g., a regulatory requirement (if applicable)). (e.g. Adverse Event Term)
- **Recommended/Conditional**
  - A data collection field that should be collected on the CRF for specific cases (may be recorded elsewhere in the CRF or from other data collection sources). (e.g. AE Start Time)
- **Optional**
  - A data collection field that is available for use if needed. (e.g. Were there any AE Experienced?)
Expectations

- Highly Recommended data collection variables should always be present on the CRF
- Sponsors will need to add data collection fields as needed to meet protocol specific and other data collection requirements
  - e.g. therapeutic area specific data variables and others as required per protocol, business practice and operating procedures.

Adverse Event (AE)

- Recommend for use in the collection of adverse events that are:
  - Non-solicited, or
  - Pre-specified

Non-solicited: AEs should be elicited with minimal connotations. For example: “Have you had any health problems since your last visit?”

Vs. solicited more in terms of determining efficacy/effectiveness of the study drug, like a pain scale...from a scale of 1–10 how would you rate your pain etc...which is usually collected on a diff. CRF

Pre-specified: Maybe AE’s that have been identified in your investigator brochure, not necessarily pre-printed but can

Adverse Events (AE)

<table>
<thead>
<tr>
<th>Data Collection Field</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Were any Adverse Events experienced? (Optional)</td>
<td>Intention/purpose of collecting this field is to help with data cleaning and monitoring. It provides verification that all other fields on the CRF were deliberately left blank. Note: This will not be part of the AE data for submission</td>
</tr>
<tr>
<td>Line # (Optional)</td>
<td>A sponsor-defined reference number or line identifier on the CRF (e.g., pre-printed or handwritten number on the CRF). Note: Can be beneficial to use in data queries issued to site to communicate specific AE record in question.</td>
</tr>
<tr>
<td>Adverse Event (Highly Recommended)</td>
<td>The verbatim description of the AE. In most cases, will be coded to a standard dictionary such as MedDRA.</td>
</tr>
<tr>
<td>Start Date (Highly Recommended)</td>
<td>The Date the AE started using the CDASH-recommended date format (e.g., 08-AUG-2008). For SDTM: CDASH Start Date &amp; Time (if collected) are concatenated into SDTM Variable AESTDTC using the ISO 8601 date format</td>
</tr>
<tr>
<td>Start Time (Recommended/Conditional)</td>
<td>The time (as complete as possible) that the AE began. Note: May be collected in Phase 1 studies.</td>
</tr>
</tbody>
</table>

Serious Event Type:
- Cancer
- Congenital anomaly
- Prolonged hosp.
- Resulted in death
- Life threatening
- Overdose
- Serious Event Type – Additional categories for seriousness
Adverse Events (AE)

<table>
<thead>
<tr>
<th>Data Collection Field</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>End Date</td>
<td>The Date the AE resolved using the CDASH-recommended date format (e.g., 08-AUG-2008). For SDTM: CDASH End Date &amp; Time (if collected) are concatenated into SDTM Variable AEENDTDC using the ISO 8601 date format.</td>
</tr>
<tr>
<td>End Time</td>
<td>The time (as complete as possible) that the AE resolved. Note: May be collected in Phase 1 studies.</td>
</tr>
<tr>
<td>Ongoing</td>
<td>Indicates that an AE is ongoing when no End Date is provided. Note: This is not a direct mapping to the SDTM Variable AEENRF. The date of data collection in conjunction with End Date and the Ongoing fields would determine how the SDTM Variable AEENRF will be populated.</td>
</tr>
<tr>
<td>Severity</td>
<td>Description of the severity of the AE. Either Severity (AESEV) or Severity CTCAE Grade (AETOXGR) must appear on the CRF. Some studies may mandate the collection of both. Use the appropriate CDISC controlled terminology. Note: For cancer studies Severity CTCAE Grade is the Highly Recommended field and the collection of the Severity field is Optional.</td>
</tr>
</tbody>
</table>

**Relationship to study drug:** mild, mod., severe or Toxicity scale like the CTC Grade that is used in onco.studies

**Relationship type:** Where you may need to capture or determine relationship to not just study drug, but study procedures, or study drug and device used to deliver the study drug.

---

Adverse Events (AE)

<table>
<thead>
<tr>
<th>Data Collection Field</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serious Event</td>
<td>Indicates whether the event is deemed &quot;serious&quot; according to the seriousness criteria defined in the protocol.</td>
</tr>
<tr>
<td>a. Serious Event Type – Congenital Anomaly or Birth Defect</td>
<td>If the details regarding a Serious AE need to be collected in the clinical database, then it is recommended that the individual Serious Event Type variables listed as items a - f be included. In many cases sponsors are already collecting this data in a separate pharmacovigilence database and choose not to collect it again in the clinical database. In that case only the Serious Event variable will be collected.</td>
</tr>
<tr>
<td>b. Serious Event Type – Persistent or Significant Disability or Incapacity</td>
<td>Refer to Explanation above for &quot;a. Serious Event Type – Congenital Anomaly or Birth Defect.&quot;</td>
</tr>
<tr>
<td>c. Serious Event Type – Death</td>
<td>Refer to Explanation above for &quot;a. Serious Event Type – Congenital Anomaly or Birth Defect.&quot;</td>
</tr>
</tbody>
</table>

**Relationship to study drug:** mild, mod., severe or Toxicity scale like the CTC Grade that is used in onco.studies

**Relationship type:** Where you may need to capture or determine relationship to not just study drug, but study procedures, or study drug and device used to deliver the study drug.

---

Adverse Events (AE)

<table>
<thead>
<tr>
<th>Data Collection Field</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>d. Serious Event Type – Initial or Prolonged Hospitalization</td>
<td>Refer to Explanation on the previous slide for &quot;a. Serious Event Type – Congenital Anomaly or Birth Defect.&quot;</td>
</tr>
<tr>
<td>e. Serious Event Type – Life Threatening</td>
<td>Refer to Explanation on the previous slide for &quot;a. Serious Event Type – Congenital Anomaly or Birth Defect.&quot;</td>
</tr>
<tr>
<td>f. Serious Event Type – Other Serious or Important Medical Events</td>
<td>Refer to Explanation on the previous slide for &quot;a. Serious Event Type – Congenital Anomaly or Birth Defect.&quot;</td>
</tr>
<tr>
<td>Relationship to Study Treatment</td>
<td>This field captures the clinical investigator’s determination of whether the study treatment had a causal effect on the AE. DELETE sentence: Use the appropriate CDISC controlled terminology.</td>
</tr>
</tbody>
</table>

**Relationship to study drug:** mild, mod., severe or Toxicity scale like the CTC Grade that is used in onco.studies

**Relationship type:** Where you may need to capture or determine relationship to not just study drug, but study procedures, or study drug and device used to deliver the study drug.
Adverse Events (AE)

<table>
<thead>
<tr>
<th>Data Collection Field</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Action Taken with Study Treatment</td>
<td>This is the action taken with the study treatment in response to the adverse event. Use the appropriate CDISC controlled terminology.</td>
</tr>
<tr>
<td>(Highly Recommended)</td>
<td></td>
</tr>
<tr>
<td>Other Action Taken (Optional)</td>
<td>Record all other actions taken (does not include action taken related to the study treatment). This field is usually reported as a free text field.</td>
</tr>
<tr>
<td>(Highly Recommended)</td>
<td></td>
</tr>
<tr>
<td>Outcome</td>
<td>Record the appropriate outcome of the AE in relation to the subject's status. Use the appropriate CDISC controlled terminology.</td>
</tr>
<tr>
<td>(Highly Recommended)</td>
<td></td>
</tr>
<tr>
<td>Adverse Event Caused Study Discontinuation (Optional)</td>
<td>Indicates whether the AE(s) caused the subject to discontinue from the study.</td>
</tr>
</tbody>
</table>

### AE Domain Table Example

<table>
<thead>
<tr>
<th>Data Collection Field</th>
<th>Definition</th>
<th>Cost Impact From Completion Termination</th>
<th>Additional Information For Spacing</th>
<th>CDISC Core</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Adverse Event</td>
<td>AE-001-001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Year Start</td>
<td>Data value, for adverse event event.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Next Time</td>
<td>Data value, for adverse event event.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Relationship to study drug:** mild, mod., severe or Toxicity scale like the CTC Grade that is used in onco.studies

**Relationship type:** Where you may need to capture or determine relationship to not just study drug, but study procedures, or study drug and device used to deliver the study drug.
Set up Schema

CRF Design

Data Entry
Using CDM ODM metadata

1. FastTrack ODM loaded into Rave
2. Additional form loaded into Rave
3. Additional behaviour added in Rave Architect

Mapping
Linking the Worlds

Protocol Setup, Capture & Mapping

SDTM Domains

Analysis

5th ANNUAL CLINICAL FORUM BASEL 2011

Mapping

• SDTM does not contain an audit trail
• Mapping between SDTM tables and CRF pages needed for every trial, proprietary within each EDC-CDM system
• Audit trail, signatures, administrative data would have a proprietary format within the EDC-CDM application

SDTM not sufficient for EDC-CDM!

Mapping

• ODM contains audit trail, signatures, internationalisation
• ODM is extremely flexible to adapt to any kind of CRFs
• Mapping between ODM and CRFs is trivial (1:1)
• ODM contains XML-based value-level metadata that can be shared with SDTM
• ODM can integrate the SDTM controlled terminology

ODM
Mapping

ODM = CRF metadata and data = Visits, Forms, ItemGroups, Items

N:M Mapping

SDTM = Case Report Tabulations = different (!) „ItemGroups“, Items

ODM-SDTM Mapping examples

• 1:1 mapping
  – Date of birth on a CRF page → Column “BRTHDTC” in SDTM DM table (horizontal)
  – Sex on a CRF page → Column “SEX” in SDTM DM table (horizontal)
  – Weight and Height on a CRF page → Weight and Height in the column “VSORRES” of the SDTM VS table (vertical)

• 1:N mapping
  – Visit date on one CRF page → Visit date in many SDTM tables

• M:1 mapping
  – Date of FU visit on a CRF page - Date of baseline Visit on another CRF page → Study day in SDTM
Steve’s Disclaimer

Views expressed in this tutorial are those of the speaker and not, necessarily, of the Food and Drug Administration.

Implementing Standards

- FDA: Organization, Mission and The Critical Path
- Laws, Regulations, Guidance and Specifications
  - Food Drug and Cosmetics (FD&C) Act (with amendments)
  - Code of Federal Regulations (CFR)
  - Guidance
  - Specifications
Regulatory Background

- FDA Organization and Mission
- Regulatory Processes and Information Resources

FDA Organization

Office of the Commissioner

- Office of Regulatory Affairs
- Center for Drug Evaluation & Research
- Center for Biologics Evaluation & Research
- Center for Devices & Radiological Health
- National Center for Toxicological Research
- Center for Tobacco Products
- Center for Food Safety & Applied Nutrition
- Center for Veterinary Medicine

FDA – White Oak, MD
FDA’s Mission & Standards

- “protecting the public health”
- “advancing the public health by helping to speed innovations”

Science of drug development: Critical Path
- Improvements: data collection (CDASH), review & submission processes
- Standards: (SDTM, ADaM, SEND, LAB, Terminology, etc)
- Shared repositories
- Electronic Health Records

The Critical Path Initiative

Innovation

Challenge and Opportunity on the Critical Path to New Medical Products

Critical Path -- SDTM for CRTs
The US Regulatory World

- Acts (Laws)
- Regulations
- Guidance
- Specifications
- Public Notice
  - Federal Register
  - FDA/Center Webpages

Acts/Laws

- Passed by U.S. Congress
- Examples include:
  - Food Drug and Cosmetics (FD&C) Act
  - FDA Modernization Act of 1997
    - amended the Federal Food, Drug, and Cosmetic (FD&C) Act relating to the regulation of food, drugs, devices, and biological products.
    - recognized the Agency would be operating in a 21st century characterized by increasing technological, trade and public health complexities.
  - FDA Amendments Act (FDAAA) of 2007 (Including PDUFA IV)
    - FDA is committed to achieve the long-term goal of an automated standards base information technology (IT) environment for the exchange, review, and management of information supporting the process for the review of human drug applications throughout the product life cycle.

Code of Federal Regulations (CFR)

- 21 CFR 314.50
  - Provide general requirements for submitting marketing applications to CDER
  - Subpart B--Applications Sec. 314.50 Content and format of an application. Case Report Tabulations [the observed/raw data -- SDTM]
- 21 CFR 11 –Good practice for all computerized processes – Sponsors and Government
  - Paved way for submission
    - Systems
    - Guidance
    - Procedures
  - “…intended to permit the widest possible use of electronic technology…”
Guidance

- Represents the Agency’s current thinking
- Not binding on FDA or the public
- An alternative approach may be used if such approach satisfies the requirements of the applicable statutes, regulations or both.
- Guidance does not limit the authority of a Center and should not supplant discussions between Centers and sponsors!!!!

Guidance for Industry

- Providing Regulatory Submissions in Electronic Format—Human Pharmaceutical Applications and Related Submissions Using the eCTD Specifications
- April 2006
- eCTD became the only accepted format for electronic submission of an NDA starting January 1, 2008

It is important to make the distinction between the DRAFT GUIDANCE document which we are introducing here today and the Electronic Record; Electronic Signature Rule which I discussed earlier in the way of background. The Rule, 21 CFR Part 11, represents Agency regulation which applies when the choice is made to use electronic records or electronic signatures to meet an Agency record or signature requirement.

FDA’s “Communication” Toolkit

Laws & Guidance &
eCTD Specifications

Many Arnolds…

The Federal Register

http://www.gpoaccess.gov/fr/index.html
ICH Guidelines = FDA Guidance

- E3 (Study Report)
- E6 (GCP)
- E9 (Statistics)
- CTD
- eCTD
- Safety Reporting
Common Technical Document

- ORGANISATION OF THE COMMON TECHNICAL DOCUMENT FOR THE REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE -- M4

- This Guideline has been developed by the appropriate ICH Expert Working Group and has been subject to consultation by the regulatory parties, in accordance with the ICH Process. At Step 4 of the Process the final draft is recommended for adoption to the regulatory bodies of the European Union, Japan and USA.

CTD -- Objectives

- Common format / elements
- Significantly reduce the time and resources needed to compile applications
- Ease the preparation of electronic submissions.
- Facilitate regulatory reviews and communication with the applicant … document of common elements.
- Exchange of regulatory information between Regulatory Authorities will be simplified.

The CTD Triangle/Pyramid
FDA eCTD – Includes the Data

Study Data
File Formats:
• SAS Transport (Version 5)

Organization
• Datasets go in same section in the TOC as its corresponding study report

In Addition to SDTM – We Require the Analysis Files!

- **Data Tabulations**: Observations in SDTM Standard Format
- **Data Listings**: Domain views by subject, by visit
- **Patient Profiles**: Complete view of all subject data
- **Analysis Files**: Custom datasets to support an analysis
Specifications: Analysis Datasets

- Analysis datasets are datasets created to support results presented in study reports, ISS, ISE and other analyses that enable a thorough regulatory review. Analysis datasets contain both raw and derived data.
- Sponsors should therefore augment SDTM with analysis data sets as described in the Analysis datasets section.
- CDISC/ADaM standards for analysis datasets (http://www.cdisc.org/adam) may be used if acceptable to the review division.
- Prior to submission, sponsors should contact the appropriate center’s reviewing division to determine the division’s analysis dataset needs.
Current Situation?

- Recognition that most Sponsors have their own standards
- However, even within an NDA, there may be differences
  - Data structure, variable names, data location, meta-data
- Requires that FDA reviewers spend time to learn each individual data standard at the beginning of each review
- Prevents standard software development
- Inhibits analyses across drug classes
- Makes meta-analysis difficult

We Are “The Problem”
Submission Data Review: Adverse Events

Submission Data Extraction
Assessing Potential Liver Injury
[by Analyzing Increases in Serum Alanine Aminotransferase (ALT) and Total Serum Bilirubin (TBIL)]

Individual Patient Profile:
Linkage of several data tables using the same timeline

Drug experience Data
Adverse Event Data
Concomitant Drugs
Laboratory Data

X-axis: Days into Study

Benefits of Standards

- Efficiency improved
- Effectiveness improved
- Time reduced
- New therapies faster
- Submission reviews more effective/efficient
- Time, money, opportunity

CDISC Data Submission Proposal: The Wrong Way

Sponsor: Oh, and one last item – we would like to submit the data in CDISC

FDA: Sure, we can accept that. That should wrap things up. Thanks for meeting with us!
CDISC Data Submission Proposal:

1. For SDTM: No deviations from SDTM Implementation Guide (SDTMIG)
   - SDTMIG spells out
   - Current production version is SDTMIG 3.1.2 [NOTE]
   - For Analysis Files – Talk to Review Division: May use ADaM [2.1 & 1] or Sponsor-Defined

2. Assessment of analyses of interest and data required for efficacy and safety
   - Determination of data needed for analyses (efficacy and safety)
     - Protocols, Statistical Analysis Plans and Study Reports
     - Safety Analysis Plan (guidance being developed)
   - Safety analyses should be based on what is known from other sources, for example: Pre-clinical information; Drug class information; Phase 1 and 2 data; Other data (foreign post-marketing, other studies etc.)

3. Where the data/variables fit in CDISC format (i.e., SDTM vs. ADaM)
4. Have discussion regarding reviewer needs (Programs, analysis files)
5. Ask whether the Medical Review Division has a checklist or template for data submission
6. Questions about submission of data -- edata@fda.hhs.gov
7. For latest ... always check: http://www.fda.gov/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/ucm248635.htm

Additional Considerations

- Provide sample data sets prior to submission
- Include a list of all included variables as well as all excluded variables
- Allows reviewers the opportunity to request permissible variables that the sponsor was going to exclude
- If SDTM does not meet a specific data need, the answer is not to alter SDTM
  - ADaM
  - SuppQual – use to add new variables
  - New Domains when necessary
- SUPPQUAL vs. new domain – don’t create many variables that really should be in there own domain
- DDD – Due Data Diligence
Study Data Standards for Submission to CDER

- CDER strongly encourages IND sponsors and NDA applicants to consider the implementation and use of data standards.
- Implementation should occur as early as possible in the product development lifecycle.
- This webpage will be updated regularly to reflect CDER’s growing experience in order to meet the needs of its reviewers.

CDER Data Standards Common Issues Document

- Goal to communicate general CDER preferences and experiences regarding the submission of standardized data to aid sponsors in the creation of standardized datasets.
- The document is not intended to replace the need for sponsors to communicate with review divisions regarding data standards implementation approaches or issues, but instead, it is designed to compliment and facilitate the interaction between sponsors and divisions.

Goal: CDISC for Every Submission

- SDTM data
  – All the “raw” (cleaned) data necessary
- ADaM datasets
  – All the analysis datasets
- Define.xml
  – All the metadata descriptions for domains, variables and value sets
SDTM

- Study Data Tabulation Model (SDTM) - Data Tabulations
- Sources:
  - On the paper CRF, updated by queries
  - In the EDC (Electronic Data Capture) database
  - In electronic transfers
- General Rules:
  - Collected/cleaned “raw” or “observed” data
  - Missing values are missing values (dates especially)
  - The sponsor decides what data to submit, based on science and regulation and YOUR conversations with FDA.

SDTM: (Mostly) “Vertical”

**Horizontal Dataset Structure**

<table>
<thead>
<tr>
<th>PATNO</th>
<th>VSDATE</th>
<th>SYSBP in mm</th>
<th>DIABP</th>
<th>PULSE in BEATS/</th>
<th>TEMP in °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1230</td>
<td>2003-02</td>
<td>120</td>
<td>80</td>
<td>65</td>
<td>37</td>
</tr>
</tbody>
</table>

**Vertical Dataset Structure**

<table>
<thead>
<tr>
<th>USUB</th>
<th>VSDTC</th>
<th>VSTEST</th>
<th>VSORRE in mmHg</th>
<th>VSORRE BEATS/</th>
<th>VSORRE °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>12301</td>
<td>2003-02</td>
<td>SYSBP</td>
<td>120</td>
<td></td>
<td>mmHg</td>
</tr>
<tr>
<td>12301</td>
<td>2003-02</td>
<td>DIABP</td>
<td>80</td>
<td></td>
<td>mmHg</td>
</tr>
<tr>
<td>12301</td>
<td>2003-02</td>
<td>PULSE</td>
<td>65</td>
<td>BEATS/</td>
<td></td>
</tr>
<tr>
<td>12301</td>
<td>2003-02</td>
<td>TEMP</td>
<td>37</td>
<td>°C</td>
<td></td>
</tr>
</tbody>
</table>

A sponsor might collect vital signs this way, as one record per subject per visit, to facilitate analysis.
SDTM Documentation
• Study Data Tabulation Model (v1.2) – 35 pages
  – This document describes the Study Data Tabulation Model (SDTM), which defines a standard structure for study data tabulations that are to be submitted as part of a product application to a regulatory authority such as the United States Food and Drug Administration (FDA).
• SDTM Implementation Guide (v3.1.2) – 298 pages
  – V3.1.2 is intended to guide the organization, structure, and format of standard clinical trial tabulation datasets submitted to a regulatory authority such as the US Food and Drug Administration (FDA). The SDTMIG should be used in close concert with the current version of the CDISC Study Data Tabulation Model (SDTM, available at http://www.cdisc.org/standards) that describes the general conceptual model for representing clinical study data that is submitted to regulatory authorities and should be read prior to reading the SDTMIG.

Again, we’ve given you two documents. The first (short) one is the overview. Before you begin your implementation, read this first. It covers the concepts best.
The second is more detailed.

Here is where to refer to web site (note: lets add some screen shots of web site home page).
Here is where to ask class to open documentation. Go thru each section and chapter.
Chapter 8 is relationship among datasets.

SDTM Variable Categories (1)
• Required –
  – **Required** to be included and populated
    – basic to the identification and meaning of a data record.
    – Values **cannot be null**
    – Examples: Study ID, USUBJID, Domain abbreviation
• Expected (TALK to your FDA Review Division!)
  – Required to be included but **does not have to be populated**
    – Necessary to make a record meaningful.
    – Some may be null if unknown or not done.
    – Should still be included even when value is null.
    – Examples: start and stop dates, event date, baseline flag

5th Annual Clinical Forum Basel 2011
SDTM Variable Categories (2)

- **Permissible** *(Review Common Issues Document and TALK to your FDA Review Division!)*
  - **May or may not be included**
  - Variable may be used as appropriate when collected or derived.
  - Examples: timing variables, SAE definition components, toxicity grades
- **Existing variables MUST NOT be renamed or modified in any way**
- **New variables MUST NOT be added except by use of SUPPQUAL**
- Section 4.1.1.5 page 21 of SDTMIG

---

SDTM — see SDTM v1.2

<table>
<thead>
<tr>
<th>Special-Purpose Domains</th>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td>2.2.6</td>
<td>17</td>
</tr>
<tr>
<td>Comments</td>
<td>2.2.7</td>
<td>18</td>
</tr>
<tr>
<td>Subject Elements</td>
<td>2.2.8</td>
<td>19</td>
</tr>
<tr>
<td>Subject Visits</td>
<td>2.2.9</td>
<td>20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Observation Classes</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Interventions</td>
<td>2.2.1</td>
<td>7</td>
</tr>
<tr>
<td>Events</td>
<td>2.2.2</td>
<td>9</td>
</tr>
<tr>
<td>Findings</td>
<td>2.2.3</td>
<td>11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relationships</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SUPPQUAL</td>
<td>4.1.2</td>
<td>27</td>
</tr>
<tr>
<td>RELREC</td>
<td>4.1.1</td>
<td>26</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Trial Design</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Five Tables</td>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>Changes from SDTM v1.1 to v1.2</td>
<td>6</td>
<td>29</td>
</tr>
</tbody>
</table>

---

Special Purpose Domains

- **Demography (DM)** *section 2.2.6 page 17*
  - Defines the subject population
  - One record per subject
  - USUBJID must uniquely identify each subject within a **submission**
    - Ensure that each person has a unique number
    - Ensure that the same person has the same number across studies
- **Comments (CO)** *section 2.2.7 page 18*
  - Can have multiple 200 character comments
  - Can relate comments to various levels of the data
- **Subject Elements (SE)** *section 2.2.8 page 19* and **Subject Visits (SV)** *section 2.2.9 page 20*
  - What happened to a **subject** in each Arm

---

---

Note the importance of having DM set apart from the rest, as a Special Purpose Domain. This domain is like no other. It is critically important because it is the definitive accountability for the clinical trial. Contains one record per subject per study. May have more than one record for one person, but only if this person was enrolled in more than one submitted study protocol.

Note that included variables are limited to these. Other subject information may be included in the Subject Characteristics

---

---

– Data will eventually be stored in FDA Janus warehouse
– Defines specific rules:
  - Domain/variable names
  - Structures
  - Consistency
  - Terminology
– SDTM v1.0 & SDS v3.1 are stable baseline…
  … but will, of course, still evolve
  - SDTM v1.1 minor updates, primarily for SEND
  - SDS v3.1.1 minor updates, new models, “nit-pick” updates
– At 2005–07, our trial already proves 12
General Observation Classes (1)

**Interventions**: investigational treatments, therapeutic treatments, and surgical procedures administered to the subject.

- One record per constant-dosing/treatment interval
- SDTMIG Intervention Domains:
  - Concomitant and Prior Medications (CM)
  - Exposure (EX)
  - Substance Use (SU)
- Interventions Topic and Qualifier variables in SDTM Table 2.2.1
  - *Section 2.2.1 page 7*

General Observation Classes (2)

**Events**: planned protocol milestones and occurrences/occurrences/incidents independent of planned study evaluations occurring during the trial or prior to the trial. Things that happen.

- One record per event
- SDTMIG Event Domains:
  - Adverse Events (AE)
  - Disposition (DS)
  - Medical History (MH)
  - Deviations (DV)
  - Clinical Events (CE)
- Events Topic and Qualifier variables in SDTM Table 2.2.2
  - *Section 2.2.2 page 9*

General Observation Classes (3)

**Findings**: observations resulting from planned evaluations, tests, questions, and measurements.

- One record per finding result or measurement
- SDTMIG Finding Domains:
  - ECG Test Results (EG)
  - Inclusion/Exclusion Exceptions (IE)
  - Laboratory Test Results (LB)
  - Physical Examinations (PE)
  - Questionnaires (QS)
  - Subject Characteristics (SC)
  - Vital Signs (VS)
  - Drug Accountability (DA)
  - Microbiology Specimen
  - Microbiology Susceptibility
  - PK Concentrations
  - PK Parameters
- Findings Topic and Qualifier variables in SDTM Table 2.2.3
  - *Section 2.2.3 page 11*
**Relationships**

- **SUPPQUAL** - Supplemental Qualifiers  
  - Used to represent non-standard variables  
  - Only use when necessary  
  - Allows for data not yet modeled to be included in a submission

- **RELREC** - Relating Records  
  - Used to represent relationships between independent records in separate domains  
  - Used to represent relationships between datasets

Note the importance of having DM set apart from the rest, as a Special Purpose Domain. This domain is like no other. It is critically important because it is the definitive accountability for the clinical trial. Contains one record per subject per study. May have more than one record for one person, but only if this person was enrolled in more than one submitted study protocol. Note that included variables are limited to these. Other subject information may be included in the Subject Characteristics.

---

**Trial Design Overview**

- Planned Trial Elements, Arms, Visits
  - Trial Elements  
  - Trial Arms  
  - Trial Visits

- Trial Inclusion/Exclusion Criteria (Lookup table)

- Trial Summary (Descriptive attributes of trial)  
  Section 3  page 22

---

**SDTM V3.1.2 Domains**

- **Interventions**
  - Con Meds
  - Exposure
  - Substance Use

- **Events**
  - Adverse Events
  - Disposition
  - Medical History
  - Deviations
  - Clinical Events

- **Findings**
  - ECG
  - Incl/Excl Exceptions
  - Labs
  - Physical Exam
  - Questionnaire
  - Subject Characteristics
  - Vital Signs
  - Drug Accountability

- **Special Purpose**
  - Demographics
  - Comments
  - Subject Elements
  - Subject Visits

- **Relationships**
  - SUPPQUAL
  - RELREC

Subject Attributions Tables  
A place holder for additional subject data that does not “fit” within the other models

- ATSUBJ – designed to support subject level linking

- ATRECORD – designed to support record level linking (e.g. subject visit level, subject datetime event level, etc.)

Submission Summary Information Model  
Model that contains...
Remember A Big Goal: The Tools

Individual Patient Profile:
Linkage of several data tables using the same timeline

Still Required: Analysis Files!

Data Tabulations
Observations in SDTM Standard Format

Data Listings
Domain views by subject, by visit

Patient Profiles
Complete view of all subject data

Analysis Files
Custom datasets to support an analysis
ADaM Objective:
- Provide data and metadata models along with examples of how to use analysis datasets to generate statistical results for regulatory submissions
What are Analysis Datasets?

An Analysis Dataset is a collection of:

- Variables that are either represented in SDTM (AGE) or are derived specifically for analysis (AGEGRP)
- Observations that are either represented in SDTM or are derived for an analysis purpose
- Indicator variables to convey information about the use of the variables and/or observations

Purpose of Analysis Datasets

- Combine, in one location, all variables and observations that are needed for an analysis
  - Records/variables that are imputed
  - Records selected for target window
  - Records/variables selected for analysis
- Generate statistical analysis with minimal programming
- Provide metadata to describe source and computational methods used for derived data

ADaM Key Principles

Analysis datasets should:

Be Analysis-Ready

- Allow replication of analysis with little or no programming or complex data manipulation
  - Requires redundancy of variables
- “Analysis-ready” -- eliminate or greatly reduce the amount of programming required by the statistical reviewers
ADaM Key Principles

Analysis datasets should:

Facilitate clear and unambiguous communication and provide a level of traceability

• Providing clear and unambiguous communication of the science and statistics of the trial is essential – but not a replacement for communication
• Communicate about the two processes
  – Analysis Dataset creation
  – Analysis Results generation
• Clearly describe and document each process

ADaM Key Principles

Analysis datasets should:

Be Usable by Currently Available Tools

• Be in format of V5 transport files for now
  – Usable by S+, JMP, SAS, etc.

SDTM and ADaM

• SDTM
  – Source data
  – No redundancy
  – Predominantly character variables
  – Each domain has a specific topic variable
  – Dates are ISO8601 character strings
  – Designed for data transfer and

• ADaM
  – Source & Derived data
  – Redundancy is needed for easy analysis and may contain variables for supportive purposes
  – Numeric variables are often needed for analysis
  – Uses variables from multiple domains
  – Dates are numeric to allow calculation
  – Designed for communication of science

Discussion about derived data in SDTM

1) Baseline flags – different definitions
   compromise – simple last value prior to dosing in SDTM - not finalized
1) Age – different definition
2) Population flags – source of rules?
Metadata: Dataset Name

- ADxxxxxx
  - Limited to 8 characters
- ADaM does not have controlled terminology for dataset names yet (other than ADSL)

Metadata: Structure

- Defines the structure of the data
  - One record per.....
- Information to identify unique records
  Example:
  - One record per subject per day of diary per diary assessment per time-point

ADaM Data Structures

- Subject-Level Analysis Dataset (ADSL) Structure
- The Basic Data Structure (BDS)
- Future ADaM Data Structures
  - Incidence of adverse events (ADAE)
  - Specifications for an ADAE dataset supporting analysis of incidence of adverse events. ADAE may be the first example of a more general structure supporting analysis of incidence data, such as concomitant medications, medical history, etc.
  - Detailed specifications for and examples of applying the Basic Data Structure (BDS) to time-to-event analysis.
Metadata: Variable Name

- Must adhere to CDISC SDTM metadata model conventions
  - 8 character limit, can’t begin with a number, special characters, etc.
- If variable is obtained directly from an SDTM domain and has no potential for change and is used in the same context, then variable name must be retained

ADaM BDS Variables

- Subject Identifier Variables (e.g., SDTM study identifier, SDTM unique subject identifier)
- Treatment Variables (e.g., planned treatment, actual treatment)
- Timing Variables (e.g., period, start date, end date)
- Analysis Parameter Variables (e.g., supine systolic blood pressure, baseline, change-from-baseline)
- Analysis Descriptor Variables (e.g., derivation type, windowing, time-to-event)
- Indicator Variables (e.g., flags)

Subject-Level Analysis Dataset (ADSL)

- One record per subject
- Used to provide the variables that describe attributes of a subject.
- Allows simple merging with any other dataset, including SDTM and analysis datasets.
- Endorsed by Regulatory agency staff
- Required in any CDISC based submission of data from a clinical trial (even if no other analysis datasets are submitted).
- Provides descriptive information about subjects.
  - multiple types of analyses, including descriptive, categorical, and modeling.
  - should not be forced to support all analyses in an attempt to minimize the number of analysis datasets.
  - correct location for key endpoints and data that vary over time during the course of a study is in a BDS dataset.
ADaM Metadata: Examples

- Analysis Dataset Metadata
- Analysis Variable Metadata
- Analysis Parameter Value-Level Metadata
- Analysis Results Metadata

### Analysis Dataset Metadata

<table>
<thead>
<tr>
<th>Dataset Name</th>
<th>Dataset Description</th>
<th>Dataset Location</th>
<th>Dataset Structure</th>
<th>Key Variables of Dataset</th>
<th>Class of Dataset</th>
<th>Documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADAGASA</td>
<td>Data for the ADAGASA model</td>
<td>adagasa.gpt</td>
<td>one recent parameter per analysis set</td>
<td>USPB/BD, RCF, CP, AERST</td>
<td>BOS</td>
<td>ADAGASA SAS Section 14.11 of SAP for adagasa/adagasa scoring algorithm</td>
</tr>
</tbody>
</table>

### Analysis Variable Metadata

<table>
<thead>
<tr>
<th>Analysis Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATASET NAME</td>
<td>The file name of the analysis dataset</td>
</tr>
<tr>
<td>VARIABLE NAME</td>
<td>The name of the variable</td>
</tr>
<tr>
<td>VARIABLE LABEL</td>
<td>A brief description of the variable</td>
</tr>
<tr>
<td>VARIABLE TYPE</td>
<td>The variable type. Valid values are as defined in the Case Report Form: Data Definition Specification Standard (e.g., in version 1.0.0 they include “text,” “integer,” and “float”)</td>
</tr>
<tr>
<td>DISPLAY FORMAT</td>
<td>The variable display information (i.e., the format used for the variable in a table or graphical presentation of results). It is suggested that the display be consistent with the format terminology incorporated in the software package used for analysis (e.g., $16$ or $3$ if using SAS)</td>
</tr>
<tr>
<td>CODELIST/CONTROLLED TERMS</td>
<td>A list of valid values or allowable codes and their corresponding labels for the variable. The field can include a reference to an external nomenclature (identified by name and version) or a hyperlink to a list of values in the code/list/controlled term section of the defining data</td>
</tr>
<tr>
<td>SOURCE/DERIVATION</td>
<td>From the details about the variable’s lineage—what is the predecessor, where the variable comes from—in the source data (CDTA or other analysis dataset) or how the variable was derived. This field is used to identify the immediate predecessor source and/or a brief description of the algorithm or process applied so that one can understand the trend that led to the variable’s creation. The source/description can be as simple as a table name (e.g., ADAGASA) identifying the data file and variable that is the source of the variable (i.e., a variable copied with no change). It can be a simple description of a derivation and the variable used in the derivation (e.g., “integration of ADAGASA”). It can also be a complex algorithm, where the element contains a complete description of the derivation algorithm and a link to a document containing it and/or a link to the analysis dataset creation program</td>
</tr>
</tbody>
</table>
### Analysis Results Metadata

<table>
<thead>
<tr>
<th>Metadata Field</th>
<th>Definition of Field</th>
<th>Metadata</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISPLAY IDENTIFIER</td>
<td>Unique identifier for the specific analysis display</td>
<td></td>
</tr>
<tr>
<td>DISPLAY NAME</td>
<td>Title of display</td>
<td></td>
</tr>
<tr>
<td>RESULT IDENTIFIER</td>
<td>Identifies the specific analysis task within a display</td>
<td></td>
</tr>
<tr>
<td>PARAM</td>
<td>Analysis parameter</td>
<td></td>
</tr>
<tr>
<td>PARAM CODE</td>
<td>Analysis parameter code</td>
<td></td>
</tr>
<tr>
<td>REASON</td>
<td>Reason for performing the analysis</td>
<td></td>
</tr>
<tr>
<td>DATASET</td>
<td>Dataset used in the analysis</td>
<td></td>
</tr>
<tr>
<td>SELECTION CRITERIA</td>
<td>Criteria and cut-off values used for selecting subsets within the analysis</td>
<td></td>
</tr>
<tr>
<td>DOCUMENTATION</td>
<td>Description of the analysis performed</td>
<td></td>
</tr>
<tr>
<td>PROGRAMMING STATEMENTS</td>
<td>The analysis syntax and parameters used in the analysis</td>
<td></td>
</tr>
</tbody>
</table>

#### Display from Analysis-Ready Dataset

**Result of Dose Response PROC GLM**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean Diff</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Low</td>
<td>0.5</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Medium</td>
<td>1.0</td>
<td>0.5</td>
<td>1.5</td>
</tr>
<tr>
<td>High</td>
<td>1.5</td>
<td>1.0</td>
<td>2.0</td>
</tr>
</tbody>
</table>

**Results of Pairwise Comparison PROC GLM**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean Diff</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Low</td>
<td>0.5</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Medium</td>
<td>1.0</td>
<td>0.5</td>
<td>1.5</td>
</tr>
<tr>
<td>High</td>
<td>1.5</td>
<td>1.0</td>
<td>2.0</td>
</tr>
</tbody>
</table>

### ADaM Validation Checks

**CDISC ADaM Validation Checks**

**Version 1.0**

Prepared by the CDISC Analysis Data Model (ADaM) Team

<table>
<thead>
<tr>
<th>Revision History</th>
<th>Date</th>
<th>Version</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Release</td>
<td>June 5, 2010</td>
<td>0.1</td>
<td>Initial draft of the CDISC ADaM Validation Checks</td>
</tr>
<tr>
<td>Final Release</td>
<td>July 20, 2010</td>
<td>1.0</td>
<td>Final production version based on comments</td>
</tr>
</tbody>
</table>
Scope: Validation Checks

- Machine readable checks that can be implemented with software to test rules defined within the ADaM Implementation Guide 1.0.
- Meant to test the structure and certain standardized content of the ADaM data sets.
- This version -- not meant to define the whole spectrum of ADaM compliance including content and well defined metadata.

5th Annual Clinical Forum Basel 2011

Validation Checks: Example

<table>
<thead>
<tr>
<th>Check Number</th>
<th>ADaM Section Number</th>
<th>Test from ADaM (Y)</th>
<th>Machine Testable Failure Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>53</td>
<td>In general, the variable labels specified in the table in Section 5 are required. There are only two exceptions to the rule (1) descriptive text followed at the end of the table of variables whose names contain indexes (0, 1, 0), (2) asterisks (*) and (3) ellipses (...) in or preceded variable labels should be replaced by descriptive text.</td>
<td>Labels for ADaM variables do not match the standard labels for ADaM variables listed in the implementation guide that can be modified (with the exception of variables whose names contain indexes (0), (1), (2), (3) and (4)) and ellipses (...) indicated for descriptive appropriate text.</td>
</tr>
</tbody>
</table>

5th Annual Clinical Forum Basel 2011

“Future Work”

- Examples with data and metadata of using the BDS for analyses such as analysis of covariance
- The ADaM team is pleased to release the ADaM Examples in Commonly Used Statistical Analysis Methods Document for Public Review. (Please provide any comments by November 18, 2011)**
- Detailed description of the ADaM metadata model and its implementation.
- A document defining ADaM compliance.
Define.xml

Define.xml: Specification
www.cdisc.org/define-xml

Case Report Tabulation Data Definition Specification (define.xml)
Prepared by the CDISC define.xml Team

Tentative Edition

Wilson Module 5

Define.XML

- Information or Data about trial Metadata
- Define.XML replaces Define.PDF
- Format for transmitting metadata about the data in a submission
- 3 areas
  - Domain
  - Variable
  - Value
- Machine-readable – facilitates use of transmission data across review tools

Newby and Wilson, 2009
SDTM Domain Metadata

- **Dataset** – 2 character prefix or domain prefix
- **Description** – describes what type of dataset
- **Class** – what type of observation class
- **Structure** – level of detail provided – 1 record/patient
- **Purpose** – Purpose
- **Key Fields** – Used to identify and index records
- **Location** – Folder and filename

----

SDTM Metadata

Domain Metadata

Datasets for Study 1234

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Description</th>
<th>Class</th>
<th>Structure</th>
<th>Purpose</th>
<th>Key Fields</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>DM</td>
<td>Demographics</td>
<td>Demography</td>
<td>Special Purpose</td>
<td>Tabulation</td>
<td>STUDYID, USUBID</td>
<td>crtdatasets/1234/dm.xpt</td>
</tr>
<tr>
<td>EX</td>
<td>Exposure</td>
<td>Intervention</td>
<td>Interventions</td>
<td>One record per event per subject</td>
<td>USUBID, EXTRT, EXSEQ</td>
<td>crtdatasets/1234/ex.xpt</td>
</tr>
<tr>
<td>CM</td>
<td>Concomitant Medications</td>
<td>Intervention</td>
<td>Interventions</td>
<td>One record per event per subject</td>
<td>USUBID, CMTRT, CMSEQ</td>
<td>crtdatasets/1234/cm.xpt</td>
</tr>
<tr>
<td>AE</td>
<td>Adverse Events</td>
<td>Events</td>
<td>Events - One record per event per subject</td>
<td>Tabulation</td>
<td>USUBID, AETERM, AESEQ</td>
<td>crtdatasets/1234/ae.xpt</td>
</tr>
</tbody>
</table>

SDTM Metadata

Variable Metadata

- **Variable Name** – 8 character name
- **Label** – describes what type of dataset
- **Type** – Character String or Numeric
- **Format** – Identifies controlled terminology or presentation
- **Origin** – Indicator of variable origin – CRF or Derived
- **Role** – How variable is used within a dataset (ID, Topic, Timing, Qualifier)
- **Comments** – Used by sponsor to assist reviewer in interpreting the data
### Variable Metadata Example

#### Demographics Dataset (DM)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>Type</th>
<th>Controlled Terms or Format</th>
<th>Origin</th>
<th>Role</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>STUDYID</td>
<td>STUDY IDENTIFIER</td>
<td>text</td>
<td>CRF Page</td>
<td>identifier</td>
<td>Demographics CRF Page 4</td>
<td></td>
</tr>
<tr>
<td>DOMAIN</td>
<td>DOMAIN ABREVIATION</td>
<td>text</td>
<td>CRF Page</td>
<td>identifier</td>
<td>DOMAINT ABREVIATION</td>
<td>Demographics CRF Page 4</td>
</tr>
<tr>
<td>USUBJID</td>
<td>UNIQUE SUBJECT IDENTIFIER</td>
<td>text</td>
<td>CRF Page</td>
<td>identifier</td>
<td>Demographics CRF Page 4</td>
<td></td>
</tr>
<tr>
<td>SUBID</td>
<td>SUBJECT IDENTIFIER</td>
<td>text</td>
<td>CRF Page</td>
<td>Topic</td>
<td>Demographics CRF Page 4</td>
<td></td>
</tr>
<tr>
<td>RFSTDTC</td>
<td>SUBJECT REFERENCE START DATE/TIME</td>
<td>date</td>
<td>CRF Page</td>
<td>Timing</td>
<td>SUBJECT REFERENCE END DATE/TIME</td>
<td></td>
</tr>
<tr>
<td>RFENDTC</td>
<td>SUBJECT REFERENCE END DATE/TIME</td>
<td>date</td>
<td>CRF Page</td>
<td>Timing</td>
<td>SUBJECT REFERENCE END DATE/TIME</td>
<td></td>
</tr>
<tr>
<td>SITEID</td>
<td>STUDY SITE IDENTIFIER</td>
<td>text</td>
<td>CRF Page</td>
<td>Record Qualifier</td>
<td>Demographics CRF Page 4</td>
<td></td>
</tr>
<tr>
<td>INVID</td>
<td>INVESTIGATOR IDENTIFIER</td>
<td>text</td>
<td>CRF Page</td>
<td>Record Qualifier</td>
<td>Demographics CRF Page 4</td>
<td></td>
</tr>
<tr>
<td>BRTHDTC</td>
<td>DATE/TIME OF BIRTH</td>
<td>date</td>
<td>CRF Page</td>
<td>Result Qualifier</td>
<td>Demographics CRF Page 4</td>
<td></td>
</tr>
<tr>
<td>AGEU</td>
<td>AGE UNITS</td>
<td>text</td>
<td>Derived</td>
<td>Variable Qualifier</td>
<td>AGE UNITS</td>
<td></td>
</tr>
</tbody>
</table>

### Variable Metadata -- (Define.xml)

```xml
<def:leaf ID="Location.EX" xlink:href="ex.xpt">
  <def:title>crt/datasets/cdisc01/ex.xpt</def:title>
</def:leaf>

<ItemGroupDef OID="DM" Name="DM" Repeating="No" IsReferenceData="No" Purpose="Tabulation" def:Label="Demography" def:Structure="One record per subject" def:DomainKeys="STUDYID, USUBJID" def:Class="Special Purpose" def:ArchiveLocationID="Location.DM">
  <ItemRef ItemOID="STUDYID" OrderNumber="1" Mandatory="Yes" Role="IDENTIFIER" RoleCodeListOID="RoleCodeList" />
  <ItemRef ItemOID="DOMAIN" OrderNumber="2" Mandatory="Yes" Role="IDENTIFIER" RoleCodeListOID="RoleCodeList" />
  <ItemRef ItemOID="USUBJID" OrderNumber="3" Mandatory="Yes" Role="IDENTIFIER" RoleCodeListOID="RoleCodeList" />
</ItemGroupDef>
```

### SDTM Value Level Metadata

- **Source Variable** – Variable Name
- **Value** – Value entered into the variable name field
- **Label** – Description
- **Type** – Data Type
- **Format** – Controlled term or format list
- **Origin** – CRF, derived
- **Role** – Role of this data
- **Comments** – Comments to help reviewer understand
Progress at CDER

• Computational Science Center (CSC)
• Data Standards Plan
• Study Data Specifications & Analysis
• Transparency and the FDA Track
• PDUFA V: Reauthorization of PDUFA (Prescription Drug User Fee Act)

The CSC

• Build capacity
  – Staffing (data management and analysis)
  – Contracts (support planning, data management)
  – Tools (further development, e.g., i-Review)
  – Training (e.g., CDISC SDTM and ADaM)
• Provide strategic oversight (CSC Board)
  – Develop long-range plans
  – Resource allocation and prioritization
  – Identify needs, measures of performance/value
Resources

- Legacy data conversion (Critical Path, Office of Women’s Health, ARRA)
- Strategic planning
- Training: CDISC SDTM and ADaM for Reviewers
- New Statistical Programming and Data Management Staff

CDER Data Standards Plan

- Examine CDER experience in terms of NDA submissions
- More effective feedback to CDISC for improvements in data standards
- Develop ‘model submission’ standards that includes early consistent communication with sponsors with Sponsors prior to NDA submission
- Develop CDER SDTM and ADaM implementation documents
- Ensure more predictability and transparency from CDER
Development of Disease-Specific Standards Through Collaboration Between FDA and CDISC

- Through Collaboration Between FDA and CDISC (e.g., SHARE Project)
- Identify and Prioritize Therapeutic Areas for Standardization
- Communicate and Coordinate Priorities
- Engage Necessary Stakeholders
- Identify Core Team Leads and Working Group Experts, Identify FDA Team Leads and Working Group Experts

Disease-Specific Data Standards

- Gather Representative Controlled Vocabularies
- Parse Out Unnecessary Data Elements from Data Dictionaries
- Develop and Finalize Draft Set of Data Elements
- Develop and publish draft CDISC products
- Address Public Comments and Publish Standard
Each FDA-TRACK program office collects, analyzes, and reports its performance measures and results via FDA-TRACK dashboards.

These FDA-TRACK dashboards may include one or several program offices which contribute to similar public health objectives or program areas.

The dashboards are published to this site quarterly following the completion of the quarterly briefing.
A. To enhance the quality and efficiency of FDA’s review of NDAs, BLAs, and INDs, FDA shall consult with stakeholders, including pharmaceutical manufacturers and other research sponsors, to issue draft guidance on the standards and format of electronic submission of applications by December 31, 2012.

B. FDA will issue final guidance no later than 12 months from the close of the public comment period on the draft guidance. Such final guidance and any subsequent revisions to the final guidance shall be binding on sponsors, applicants, and manufacturers no earlier than twenty-four months after issuance of the final guidance.

Working on the “Hard Part(s)”

- Terminology – CDISC SHARE
- Beyond Terminology
- Integrated Systems for Data Acquisition: Electronic Capture – CDASH
- Sentinel System, Registrars, Observational Data, Comparative Efficacy, EHRs
- People and Process: Standardizing Good Review Practice/Tools/Training
- Learning Lessons – Continuous Improvement
- Common Understanding -- Collaboration
- Understanding How to Communicate – Analysis, Statistical Programming, etc.

THANK YOU

pierre-yves.lastic@sanofi.com
stephen.wilson@fda.hhs.gov
http://www.cdisc.org