An Introduction to CDISC: Available CDISC Standards and Models and How SAS Supports These

Dave Handelsman
Principal Strategist, Clinical R&D, SAS

October 2005
Agenda

- Part I: What is CDISC and Why Standards?
- Part II: Available CDISC Standards and Models and How SAS Supports These
- Part III: CDISC Collaborations and Standards in Progress
Clinical Data Interchange Standards Consortium

CDISC is an open, multidisciplinary, non-profit organization committed to the development of worldwide industry standards to support the electronic acquisition, exchange, submission and archiving of clinical trials data and metadata for medical and biopharmaceutical product development.

The mission of CDISC is to develop and support global, platform-independent data standards that enable information system interoperability to improve medical research and related areas of healthcare.
CDISC Standards

- **Operational Data Model (ODM)**
- **Submission Domain Standard (SDS)**
  - Study Data Tabulation Model (SDTM)
  - Standard for Exchange of Non-clinical Data (SEND)
- **Analysis Data Model (ADaM)**
- **Laboratory Data Model (Lab)**
- Case Report Tabulation Data Definition Specification (CRTDDS - aka define.xml)
- Protocol Representation Standard
- Biomedical Research Integrated Domain Group (BRIDG)
ODM – Operational Data Model

- Written in XML
- Clinical trials data recording / transfer / archive
- Contains both study metadata and clinical data

Model Features:

- Framework can be used by any clinical data management (CDM) system or clinical application
- Supports 21 CFR Part 11 audit trail and electronic signatures
- Supports archiving & data retention guidance
- Provides for incremental transfers
- Includes vendor extension capability
Interchange cycle – move to ODM

- **Traditional approach**
  - MetaData Xmit = Excel spreadsheet
  - Data Xmit = SAS Datasets

- **CDISC approach**
  - MetaData Xmit = ODM XML
  - Data Xmit = ODM XML

- **CDISC Benefits**
  - More information exchange
  - Increased automation
ODM & Audit Trail

- **What**: ItemData
  - AuditRecord
    - DateTimeStamp
  - UserRef
    - LocationRef
    - ReasonForChange
      - SourceID
  - CryptoBindingManifest
  - MeasurementUnitRef
    - Annotation
      - Flag
  - Comment
- **Who**: UserRef
- **Why**: LocationRef
  - DateTimeStamp
- **When**: UserRef
SDS, SDTM and SEND – Submission Standards

- **SDS: Submissions Data Standards**
  - Refers to the *implementation* of the model for *clinical submissions* (e.g. SDS v3.1.1)

- **SDTM: Study Data Tabulation Model**
  - Refers to the *model* that supports *clinical and preclinical submissions* data

- **SEND: Standard for Exchange of Nonclinical Data**
  - Defines *domains and variables* for *submitting* all data generated from *animal toxicity studies*
SDS V3.1.1: CDISC Standard Domain Models Observation Classes

- Interventions
  - Exposure
  - ConMeds
  - Subst Use

- Events
  - AE
  - Disposition
  - MedHist

- Findings (80% of data)
  - Labs
  - Vitals
  - PhysExam
  - ECG

- InclExcl
- SubjChar
- ConMeds
- AE

- Other
  - Demog
  - Relates
  - Supp Qual
  - Comments
  - Study Design

Special Purpose Tables

Study Observations
SDTM Basic Concepts

- Utilizes standard variable names, metadata, and data structure
- Each domain has required, expected, and permitted variables
- No redundancy (other than identifying variables)
- SDS team has developed standards, containing specific metadata for >20 domains
SDTM

- 3 general observation classes
  - Interventions
  - Events
  - Findings

- Special domains
  - Demographics
  - Comments
  - Relationship between records in multiple domains
  - Supplemental qualifier domains
Specifications For Organizing The Datasets

SDTM Benefits

- **Increase in efficiency**
  - Data transfers
  - Data reviews
  - Creation of analysis files

- **Flexible**
  - Supplemental domains
  - Can accommodate a mixture of raw and derived
ADaM – Analysis Dataset Model

Uses for Analysis Datasets:
- Provide clear communication of the scientific results to regulatory reviewers
- Replicate or verify the sponsor’s analyses, results, and conclusions
- Test the validity and robustness of the sponsor’s analyses and assumptions (what if…)
- Audit the data for inconsistencies and errors
- Status: Analysis Dataset Models Completed
  - Change from baseline
  - Time to event
  - Categorical analysis
- Working on Additional Dataset Models
  - Linear model
  - Safety Analysis
- Ensuring compatibility with SDS Version 3.1 (SDTM)
Lab – Clinical Laboratory Data Model

- Core model designed to handle “simple lab data”: one test, one result structure
- 10 levels of content includes support of administrative information for transfer of samples, reporting of results
- Production Version 1.0.1 - on CDISC web site
  - Specification, data fields, reference ranges
  - Includes Microbiology extension
- Currently standard approaches are included for
  - Bar delimited ASCII (default representation)
  - SAS
  - XML
  - HL7 RIM message
- More complex tests - handled by extensions
  - Work on pharmacogenomics beginning
Case Report Tabulation Data Definition Specification (CRTDDS aka define.xml)

- Machine readable replacement for define.pdf currently used in submission process
- Organizes and describes submission content
- Extends ODM metadata for submission content
  - Enhances general column descriptive information
  - Adds item-level metadata support
  - Supports linkage to external folders and files
Protocol Representation Standard

- Initiated as an HL7 project to develop standard representation of clinical trial protocol elements
- CDISC team formed to provide domain expertise
- Spreadsheet of elements, with glossary definitions open for public review and comment
- Initial Clinical Document Architecture (CDA) model developed and balloted through HL7
CDISC Biomedical Research Integration Domain Group (BRIDG)

- Follows the HL7 Development Framework
- Initiated in January 2004 by CDISC Board
- Developed through numerous modeling sessions with domain experts and reviewed by CDISC, industry, FDA, NCI/NIH groups; led by HL7 RIM expert

Purposes and Anticipated Benefits:
- To help ensure harmonization among CDISC models (present and future)
- To provide the industry with a standard model to represent the clinical research domain
- To enable an HL7 implementation of the CDISC ODM
- To help harmonize the CDISC and HL7 standards
- To help enable interoperability between clinical research and healthcare systems
Evolution of CDISC Standards: Started with the end in mind...

One Harmonized Standard

- **Data Sources**
  - Site CRFs
  - Laboratories
  - Contract Research Organizations
  - Development Partners

- **Operational Database**
  - Study Data
  - Audit Trail
  - Metadata

- **Submission Data**
  - CRT/Domain Datasets
  - Analysis Datasets
  - Metadata

ODM = Operational Data Model/Std
LAB = Laboratory Data Model/Std
SDS = Submission Domain Standards
ADaM = Analysis Data Models
SEND = Standards for the Exchange of Non-Clinical Data
How SAS Supports CDISC
“At first glance, SAS might appear to be the loser in any standards shift. After all, it owns the format in which clinical data are analyzed….

... But [SAS] declares: *Standards do not scare us. We embrace them.*

... The company is both acutely sensitive to its large user base and attuned to the shifting winds of industry, where the Clinical Data Interchange Standards Consortium (CDISC) has been gaining traction.”

- Mark Uehling, Senior Science Editor, Bio-IT World | eCliniqua
  Thursday, Dec 9, 2004
Support for the CDISC Organization

- SAS is a corporate sponsor
- Ed Helton, of SAS, recently elected Chairman of the Board for CDISC
- Dave Handelsman is SAS’ representative on the Industry Advisory Board
- SAS has hosted the Industry Advisory Board meeting in the past, and periodically hosts Board of Directors’ meetings
- SAS participates in several of the modelling teams
SAS Software Developed for CDISC

- PROC CDISC
  - ODM (v1.2.1): Data conversion between SAS and XML
  - SDTM (v3.1): Data structure verification
- XML Engine ODM Native mode (v1.2.1)
- XML Engine and XMLMap Extensions
- New base SAS formats/informats for ISO-8601
- SAS CDISC Viewer (ODM 1.2.1)

See Tony Friebel’s talk at 16:05 today for more details: SAS Dataset Content Conversion to CDISC Data Standards
PROC CDISC Example

libname results 'SASEnvironment\SASCode\CDISC Demo\data' ;
FILENAME XMLINP "\CDISC Demo\xml\ae.xml";

PROC CDISC
    MODEL=ODM       READ=XMLINP
    formatActive=YES formatNoReplace=NO;

    ODM
        ODMVersion = "1.2"
        ODMminimumKeyset=NO;

    CLINICALDATA
        OUT = results.AE
        SASDATASETNAME = "AE" ;

RUN;
SAS Integration Based on CDISC ODM

- ODM v1.2.1 used to bridge DataTrak EDC with SAS Drug Development
- Incremental data is transferred from EDC server to SAS Drug Development
  - Scheduled, or
  - Ad hoc
- Data is then:
  - Extracted from XML (ODM) into normalized SAS data sets
  - Compared to master files
  - Updated according to ODM transaction types
- Once loaded, data sets are:
  - Versioned
  - Audit trailed
  - Available for further transformation and analysis
SAS ETL Studio
ETL Studio

- Source Designer
- Target Designer
- Process Designer
ETL Studio Desktop

- Customizable Trees
- Many Object types
  - Document
  - Source Table
  - Target Table
  - Job
  - External Files
  - Groups
  - Libraries
  - Notes
  etc

Repositories
- Foundation
  - CDISC Data Standards
  - Clinical Domain Models
  - Misc
- Projects
  - Cardiovascular
    - Nicardipine
      - Aggregated Data
    - Ncshah Documents
      - CRF Images
        - Adverse Events CRF
        - Demography CRF Page 1
        - Demography CRF Page 2
        - Demography CRF Page 3
        - Demography CRF Page 4
        - Demography CRF Page 5
        - Laboratory Results CRF
    - Protocol
    - SAP
    - Ncshah Study 001
    - Ncshah Study 002
      - Ncshah 002 Notes
      - Ncshah 002 Jobs
        - Copy of Load Ncshah Study 001 Concomitant Medications (CM)
        - Load Ncshah Study 002 Adverse Event Data (AE)
        - Load Ncshah Study 002 Demography Data (DM)
        - Load Ncshah Study 002 Disposition (DS)
        - Load Ncshah Study 002 Vasospasm Data (VA)
    - Ncshah Study 002 Rawdata
    - Ncshah Study 002 SDTM 3.1 Data
      - Ncshah Study 002 Adverse Events - AE
      - Ncshah Study 002 Concomitant Medications - CM
      - Ncshah Study 002 Demographics - DM
      - Ncshah Study 002 Disposition - DS
      - Ncshah Study 002 Vasospasm - VA
- How-to Docs
Source Designer

The following metadata will be created:

ADVERSE, DEATH, RANDOM FILE, REGISTER

The following tables will be created in the Project: E7Luser repository:
Target Designer

The following metadata will be created:

- **Name**: New Study Adverse Events (AE)
- **SAS Dataset Name**: AE
- **SAS library**: NCS/AH Study 001
- **Number of Columns**: 45

- **Column Name**: STUDYID
  - **Description**: Study Identifier
  - **Column Type**: C
  - **Column Len**: 40
  - **Informat**: 
  - **Format**: 

- **Column Name**: DOMAIN
  - **Description**: Domain Abbreviation
  - **Column Type**: C
  - **Column Len**: 40
  - **Informat**: 
  - **Format**: 

- **Column Name**: NCS/AH Study 001
Process Designer
/* Job: Load Demography */
* Description: *
* Repository: Sandbox *
* Server: SASMain *
* Source Table: Demo - s_pr.Dem *
* Target Tables: Demographics - DM - s_ta.DM *
* Supplemental Qualifiers - SUPQUAL - s_ta.SUPQUAL *

* Generated on: Tue Mar 08 13:38:05 EST 2005 *
* Generated by: vamp-1@seadog *
* Version: 9.1.20041201.4014 *

/*****************************************************************************/

/* Setup to capture return codes */
%let job_rc = 0;
%let trans_rc = 0;
%let sqlrc = 0;
%global $saerr;
CDISC SDTM Data Model

- Implemented in the metadata
- Implements all currently defined models
- Contains table definitions and model metadata
- Facilitates a workflow to define new models
- Provides supplemental tables for validation steps
Reference Tables (controlled terminology)

<table>
<thead>
<tr>
<th>#</th>
<th>ECG TEST Codes</th>
<th>ECG Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FINDING</td>
<td>General ECG Finding</td>
</tr>
<tr>
<td>2</td>
<td>HR</td>
<td>Heart Rate</td>
</tr>
<tr>
<td>3</td>
<td>INTP</td>
<td>ECG Interpretation</td>
</tr>
<tr>
<td>4</td>
<td>PR</td>
<td>PR Interval</td>
</tr>
<tr>
<td>5</td>
<td>GRS</td>
<td>GRS Interpretation</td>
</tr>
<tr>
<td>6</td>
<td>GT</td>
<td>GT Interval</td>
</tr>
<tr>
<td>7</td>
<td>GTCB</td>
<td>GTc Interval Bazett</td>
</tr>
<tr>
<td>8</td>
<td>GTCF</td>
<td>GTc Interval Friedreich</td>
</tr>
<tr>
<td>9</td>
<td>RR</td>
<td>RR Interval</td>
</tr>
<tr>
<td>10</td>
<td>VRATE</td>
<td>Ventricular Rate</td>
</tr>
</tbody>
</table>
SAS and SDTM as the Emerging Standard

OCTAGON RESEARCH SOLUTIONS IDENTIFIED AS A PRIMARY TRAINER FOR CLINICAL DATA INTERCHANGE STANDARDS CONSORTIUM (CDISC)

Corporate Sponsor to Provide Training Curriculum on Emerging Data Standards

April 4, 2005

OCTAGON RESEARCH SOLUTIONS IMPLEMENTS SAS TECHNOLOGY [SAS® Enterprise ETL Server] TO REVOLUTIONIZE CDISC SDTM LEGACY CONVERSION OFFERINGS

April 11, 2005
On the SAS Radar....

- define.xml
- Lab, SEND, ADaM
- Other emerging standards
The Move to Standards…

“Faced with rapid changes, the nation’s healthcare system has **fallen short of its ability to translate information into knowledge** that can be used in practice, and to apply new technology safely and appropriately.

The results are exactly what you would expect. Everyone who uses the current system **constantly confronts large information gaps**, whether it’s at the doctor’s office, on the hospital ward or at government agencies charged with protecting the public health. That goes for the FDA -- we’re no exception.”

*Dr. Mark McClellan, former FDA Commissioner*
*CDISC Interchange, October 2003*

“**Innovation depends upon standardization.**”

*Dr. Bob O’Neill, Director, Office of Biostatistics, CDER, FDA*
Benefits of Standards – Regulatory View

“The importance of a standard for the exchange of clinical trial data cannot be overstated. FDA reviewers spend far too much valuable time simply reorganizing large amounts of data submitted in varying formats. Having the data presented in a standard structure will improve FDA’s ability to evaluate the data and help speed new discoveries to the public.” -Lester Crawford, Former Commissioner, FDA
“... But [SAS] declares: Standards do not scare us. We embrace them.”

- Mark Uehling, Senior Science Editor, Bio-IT World | eCliniqua
Thursday, Dec 9, 2004