CDASH Standards and EDC CRF Library

Guang-liang Wang
September 18, 2014
2014 Q3 DCDISC Meeting
Disclaimer

• The content of this presentation does not represent the views of my employer or any of its affiliates. The content represents solely my personal opinions, and also presents information obtained from public sources, such as CDSIC publications.
Agenda

• Purpose
• Regulatory Expectations
• CDASH Background
• Implementation Considerations
• Benefits of Standardization
Regulatory Requirements

• “…develop standardized clinical data terminology through open standards development organizations (i.e., CDISC)”

• “…periodically publish final guidance specifying the completed data standards, formats, and terminologies that sponsors must use to submit data in applications.”

- PDUFA (Prescription Drug User Fee Act (2012))
Regulatory Guidance

• FDA
  – Providing Electronic Submission in Electronic Format – Standardized Study Data
    • Initial Draft: Feb. 2012 (eStudy Guidance)
    • Revision 1: Feb. 2014
  – Study Data Technical Conformance Guide
    • Initial Draft: Feb. 2014 (Conformance Guide)

• PMDA
  – Submission in standard SDTM format by 2016.
eStudy Guidance

• FD&C Act section 745A(a) authorizes FDA to implement electronic submission requirements by specifying the format for such submission in guidance.
• Covers amendment, supplement and report to original submission, even if the original submission is not in standard data format.
• Data should be of standards, formats and terminologies in FDA Data Standard Catalog
  – SDTM (Study Data Tabulation Model)
  – CDISC Controlled Terminologies, MedDRA
• “When planning a study (including the design of case report forms, data management systems, and statistical analysis plans), the sponsor or applicant must determine which FDA-supported standards, formats and terminologies to use…”
• FDA will not provide waiver to non-conformance data, except version of standards used.
CDISC Foundational Standards

Foundational Standards

PLANNING
- Protocol

DATA COLLECTION
- CDASH
- Lab

DATA TABULATIONS
- SDTM
- SEND

ANALYSIS
- ADaM

XML Data Exchange
- SDM-XML
- ODM-XML

Define-XML
- Dataset-XML

Therapeutic Area Standards

Healthcare Link

Semantics
- Glossary
- Controlled Terminology-Questionnaires
- BRIDG
CDISC Standards in Clinical Study

CDASH: Clinical Data Acquisition Standards Harmonization
SDTM: Study Data Tabulation Model
ADaM: Analysis Data Model
Regulatory Expectations on Data Collection

• FDA Critical Path Initiative: Streamlining Clinical Trials
  – Opportunity 45: Consensus on Standards for Case Report Forms

Clinical trial data collection, analysis, and submission can be inefficient and unnecessarily expensive. A wide array of different forms and formats are used to collect clinical trial information, and most data are submitted to the FDA on paper. **Differences in case report forms across sponsors and trials creates opportunities for confusion and error.** Standardization of the look and feel of case report forms could reduce these inefficiencies and also help accelerate progress toward electronic data capture and submission.
Regulatory Expectations

• Upfront implementation of standards
  – eStudy Guidance

• Avoid “map” to data standard at time of submission
  – Perform compliance check while study is on-going
  – Provide transparency by using standards

• Traceability from data collection to submission
  – Traceable from CRF data collection to SDTM datasets in eCTD
  – Straight forward relation between T&L/CSR to site source

“We cannot improve efficiency or innovate without standards.“
– Chuck Cooper, MD, FDA/CDER, Office of Translational Sciences
Purpose of CDASH

- Develop CRF content standards for a basic set of global industry-wide CRF fields to support clinical research
  - Initial scope limited to most commonly collected data
  - These CRF standards apply across most therapeutic areas and phases of clinical development (I-IV)

- Maximize re-use of data, CRFs, programming, etc.

- Increase transparency and traceability in the data

- Support data repository and data sharing

- Support integrating research into clinical care workflow
  - CDASH used as a content standard to harvest data from electronic health records (Healthcare Link)
CDASH Principles

- Standardize the questions/fields on CRFs
- Standardize the variables and harmonize with SDTM (CDASH is a subset of SDTM)
- Use standard CDISC controlled terminology that maps to SDTM
- Limit variables to required and necessary
- Reduce redundancies
- Comply with regulatory requirements
- Be appropriate for use for all phases of studies
- Allow consistent and efficient data collection, storage, transmission and analysis
CDASH Documents

• CDASH Standard v1.1
  – Domains, conformance rules and best practices

• CDASH SAE Supplement v1
  – Map data collected on CRF and/or SAE form to ICH E2B safety reporting requirements

• CDASH User Guide v1
  – Contains implementation examples, including CDASH to SDTM mappings, CDASH ODM files and a library of example CRFs that have been created in several different data collection systems, including paper examples.
**CDASH Domains**

- **Existing CDASH Domains:**

<table>
<thead>
<tr>
<th>Adverse Events (AE)</th>
<th>Inclusion and Exclusion Criteria (IE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comments (CO)</td>
<td>Lab (LB)</td>
</tr>
<tr>
<td>Concomitant Medications (CM)</td>
<td>Medical History (MH)</td>
</tr>
<tr>
<td>Demographics (DM)</td>
<td>Physical Examination (PE)</td>
</tr>
<tr>
<td>Disposition (DS)</td>
<td>Protocol Deviations (DV)</td>
</tr>
<tr>
<td>Drug Accountability (DA)</td>
<td>Subject Characteristics (SC)</td>
</tr>
<tr>
<td>ECG (EG)</td>
<td>Substance Use (SU)</td>
</tr>
<tr>
<td>Exposure (EX)</td>
<td>Vital Signs (VS)</td>
</tr>
<tr>
<td>Common Identifier Variables</td>
<td>Common Timing Variables</td>
</tr>
</tbody>
</table>

- Additional domains & data fields should be added as required by protocol.
- CDASH to be update with SDTM IG 3.2 domains
- TA-specific domains are under development
Relationship Between CDASH Forms and SDTM Domains

• One to one
  – Vital Signs form -> VS

• Many to one
  – Informed Consent, Screen failure, Randomization, Completion Status, Follow-up -> DS

• One to many
  – Demographics including Informed Consent -> DS, DM
### CDASH Specifications

<table>
<thead>
<tr>
<th>Question Text</th>
<th>Prompt</th>
<th>SDTM or CDASH Variable Name</th>
<th>BRIDG</th>
<th>Definition</th>
<th>CRF Completion Instructions</th>
<th>Information for Sponsors</th>
<th>Core</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Did the subject meet all eligibility criteria?</td>
<td>Met Criteria</td>
<td>IEYN</td>
<td>Derived*</td>
<td>Response for whether the subject met all the eligibility requirements for this study at the time the subject was enrolled. Did subject meet all eligibility criteria? (NY) (See Section 2.2).</td>
<td>Record “Yes” if all eligibility criteria were met for the study. Record “No” if subject did not meet all criteria at the time the subject was enrolled.</td>
<td>This is a Yes/No question that is intended to be used primarily as a monitoring and/or data management tool to verify that the investigator/site recorded any entry criteria that were not met. Must be recorded on the CRF. This field does not map directly to an SDTM CRF. *See the BRIDG model for complete path.</td>
<td>HR</td>
</tr>
<tr>
<td>2 What was the category of the criterion?</td>
<td>Criterion Type</td>
<td>IECAT</td>
<td>DefinedActivity. categoryCode*</td>
<td>Specifies whether the criterion is inclusion or exclusion.</td>
<td>Record whether the criterion that this subject did not meet was “Inclusion” or “Exclusion”. Checkbox: Check “Inclusion” or “Exclusion”.</td>
<td>Only records for criteria that are not met appear in the IE SDTM IG domain, and for those records IECAT must also be populated. This criterion category may be collected on the CRF in a checkbox format, or, it may be included as part of the Criterion Identifier (e.g., IOI, E01), or derived from the inclusion/exclusion criteria in the Trial Inclusion (TI) trial dataset, or other protocol definitions external to the clinical database when a clinical database is used.</td>
<td>O</td>
</tr>
</tbody>
</table>
## Metadata Tables

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question Text</td>
<td>Prompt</td>
<td>SDTM or CDASH Variable Name</td>
<td>BRIDG</td>
<td>Definition</td>
<td>CRF Completion Instructions</td>
<td>Information for Sponsors</td>
<td>Core</td>
</tr>
</tbody>
</table>

- **Full question text for the data collection field.** Either question text or prompt can be used on the CRF.
- **Short prompt for the data collection field; could be used as the CRF label.**
- **SDTM Variable Name OR CDASH Variable Name.**
- **BRIDG Mapping.**
- **Defines the Data collection field.**
- **Instructions for the clinical site on how to enter data on the CRF. Includes controlled terminology.**
- **Information/rationale and instructions on how to implement the CRF data collection fields.**
- **Designations HR Rec/Cond Optional.**
Core Designations of Fields

• Highly Recommended (HR): A data collection field that must be on the CRF (e.g., a regulatory requirement)
• Recommended/Conditional (R/C): A data collection field that should be on a CRF if applicable for the study.
• Optional (O): A data collection field that is available for use.
Conformance

• Tier 1
  – Addresses site-facing aspects of CDASH at CRF level
    • All Highly Recommended & applicable Recommended/Conditional fields are present on CRF
    • All Codelists displayed in the CRF use or map to current published CDISC Controlled Terminology
    • CDASH Question Text or Prompt is used
  – Conform to Best Practices recommendations
## Best Practices of CRF Design

<table>
<thead>
<tr>
<th>Considerations for CRF Design</th>
<th>Rational and Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only collect necessary data. All data points should be mapped to end points</td>
<td>Time and cost to collect and clean data.</td>
</tr>
<tr>
<td>Establish CRF review process and version control</td>
<td>Document all changes and amendments</td>
</tr>
<tr>
<td>Involve cross-functional team member in CRF review</td>
<td>Ensure alignment with protocol and SAP</td>
</tr>
<tr>
<td>Consider the workflow at the site</td>
<td>Easy for data entry</td>
</tr>
<tr>
<td>Employ standards</td>
<td>Increase efficiency and data quality</td>
</tr>
<tr>
<td>Ask questions neutral and avoid leading questions to influence site</td>
<td>Reduce bias and error</td>
</tr>
<tr>
<td>CRF questions should be clear and unambiguous</td>
<td>Reduce query and data cleaning costs.</td>
</tr>
</tbody>
</table>
Conformance

• Tier 2
  – Conform to the User Guide for implementation
  – Addresses operational level conformance
  – Data collection fields are defined using CDASH naming conventions in the operational database
    – Use SDTM variable or follow SDTM naming conventions
  – Pre-define mapping from operational database to SDTM datasets
  – Follow CDASH recommendations for creating fields that do not exist in CDASH
  – Follow the CDASH rules to define additional domains
Create New Fields

• Prior to adding any new fields to current domain models, review the CDASH to see if there is a similar field that can be adapted. When adding fields not already defined in a CDASH, there can be 3 categories:
  – A field that is used for data cleaning purposes only, and is not submitted in SDTM.
  – A variable with a response that can be mapped directly into SDTM with no change from how it was collected.
  – A variable that will eventually be mapped into SDTM, but is collected in such a way that it requires some change from the way it was collected.

• The process for creating new data collection fields and the corresponding database variables:
  – If a value can be collected exactly as it will be reported in SDTM, the SDTM variable should be used in the operational database to streamline the mapping process, and its label and meaning should not be modified for data collection.
  – If a study requires a field that is not identical to an SDTM field, the operational database should use a different name from the SDTM variable into which it will be mapped.
Create New Domains

- Ensure standard CDASH domain is not exist
- If SDTM domain is exist, model CDASH domain against the SDTM domain
- Then, create ‘new’ custom domains based on one of the three General Observation Classes as defined in the SDTM.
  - Use the SDTM variable and label in the corresponding Class as field name and question text if possible.
  - Use any specified controlled terminology for that SDTM variable
  - If there is no direct match, look in other CDASH domains in that Class for a data collection field that meets the definition of the data needed
  - If there is no corresponding SDTM variable or similar CDASH field available for what is being collected, create a custom field using the variable naming fragments and conventions available in CDASH and/or the SDTMIG
Industry Standard

- CDASH focus on CRF content
  - Ask the same question for the same data to be collected
  - Provide the same choices for answer
- Independent of data collection tools
- Ensure that all SDTM ‘required’ elements are addressed
- Be ‘standard’, but flexible, allowing customization within defined limit
- Define conformance rules and principles
- Support regulatory submission (SDTM, ADaM) & protocol design
CRF Global Library

• Ensure consistent implement of data content in EDC design
  – Utilizing system feature and functionalities
• Provide consistent layout
  – Portrait vs. landscape; single form vs. log line; dynamic vs. static, etc.
• Pre-validate EDC screens and edit checks to facilitate re-use of library objects
• Pre-define data mapping from EDC to SDTM
• Provide programing instructions
• Same look and feel of CRF screens for all studies
• Define eCRF completion instructions
Data Collection Considerations

• Check box, Radio buttons, Dropdown list, Search list
• Single form, Log form, semi-log form
• Dynamic fields and dynamic forms
• Data field format, text vs. numeric
• Define field length and precision
• Auto-populate, integration vs. manual entry
  – Site ID
  – Subject ID
  – Screen failure vs. enrolled subject
Demographics

Other specify field is next to the parent field. After saving the form, the specify field answer will be in parenthesis.
RACE and ETHNICITY

• How to handle multiple race.
  – Should we allow multiple choice?
  – Should we use ‘Other, Specify’?

• How to provide detailed race and ethnicity classification
  – Store detailed ethnicity information in main DM or SUPPDM?
Collection Date (dd-MMM-yyyy)

Birth Date (dd-MMM-yyyy)

Birth Time (24 hr clock)

Age

Age Unit

Hours

Days

Weeks

Months

Years

Sex

Female

Male

Undifferentiated

Childbearing Potential

Yes

No

Race

American Indian or Alaska Native

Asian

Black or African American

Native Hawaiian or Other Pacific Islander

White

Other
<table>
<thead>
<tr>
<th>Field Name</th>
<th>Data Type</th>
<th>Units</th>
<th>Values</th>
<th>Pre-Filled Values</th>
<th>Include Field OID</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMDAT</td>
<td>dd MMM YYYY</td>
<td></td>
<td></td>
<td>DMDAT</td>
<td></td>
</tr>
<tr>
<td>BRTHDAT</td>
<td>dd MMM YYYY</td>
<td></td>
<td></td>
<td>BRTHDAT</td>
<td></td>
</tr>
<tr>
<td>BRHTIM</td>
<td>HH:nn</td>
<td></td>
<td></td>
<td>BRHTIM</td>
<td></td>
</tr>
<tr>
<td>AGE</td>
<td>2</td>
<td></td>
<td></td>
<td>AGE</td>
<td></td>
</tr>
<tr>
<td>AGEU</td>
<td>$10</td>
<td></td>
<td>HOURS = Hours, DAYS = Days, WEEKS = Weeks, MONTHS = Months, YEARS = Years</td>
<td>AGEU</td>
<td></td>
</tr>
<tr>
<td>SEX</td>
<td>$2</td>
<td></td>
<td>F = Female, M = Male, UN = Undifferentiated</td>
<td>SEX</td>
<td></td>
</tr>
<tr>
<td>CHILDBR</td>
<td>$1</td>
<td></td>
<td>Y = Yes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Codelists & SDTM Terminologies

• Use of Codelists
  – Include all SDTM terminologies
  – Include only most commonly used terminologies, and use ‘Other, Specify’ for the rare terms
  – How to capture the Specified text? Is there a need to standardize?

• Understand use of SDTM terminologies prior to design eCRF screen to ensure consistency in mapping
  – Vital Signs: Heart Rate v.s. Pulse Rate
  – ECG: Heart Rate v.s. Ventricular rate
Normalized vs. De-normalized

• Benefits of de-normalized or portrait layout:
  – Facilitate the flow of data entry
  – Avoid not applicable data entry field
  – Able to define data type and precision of the data field

• Cautions:
  – Ensure use controlled terminologies as variable names, e.g., TEST and TESTCD
  – Follow variable naming conventions in SDTM IG

• Example: Vital Signs, Questionnaires
Date Format

• Separate Date and Time into different fields
• Handling of Partial Date
  – Always allow just in case, or only selected date fields
  – Enter UNK or leave blank for the missing part(s) of date
EDC Global Library

- eCRF Forms
  - Variable name, data format, questions
  - Codelists/Terminologies
- Form-level Edit Checks
- EDC system configurations
  - PDF profiles
- eCRF Completion Guide
- Annotated CRFs
- SDTM Mapping Specification
- SDTM dataset output
Supporting Documents

• Standard Library Management Process
• eCRF Specification Document
• eCRF Design Guidelines
• Edit Specifications
• Mapping Specifications
• Validation Documents
• Library Object Proposal and Approval form
EDC Global Library

Sponsor

Central Global Library

CRO 1

Local Global Library

Study Build

CRO 2

Local Global Library

Study Build
## TA Standards Projects

### CFAST Therapeutic Area Standards Project Pipeline as of June 2014

<table>
<thead>
<tr>
<th>2012-2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Asthma v1</strong></td>
<td>Traumatic Brain Injury</td>
<td>Rheumatoid Arthritis</td>
<td>Duchenne Muscular Dystrophy*</td>
</tr>
<tr>
<td><strong>Alzheimer’s v2</strong></td>
<td>Hepatitis C</td>
<td>Solid Organ Transplant</td>
<td>Oncology v3*</td>
</tr>
<tr>
<td><strong>Multiple Sclerosis v1</strong></td>
<td>Breast Cancer</td>
<td>Post-Menopausal Osteoporosis</td>
<td>Irritable Bowel Syndrome*</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Influenza</td>
<td>Oncology v2*</td>
<td>Skin and Skin Structure Infections*</td>
</tr>
<tr>
<td>Cardiovascular Endpoints</td>
<td>Schizophrenia</td>
<td>Tuberculosis v2*</td>
<td>Complicated UTIs*</td>
</tr>
<tr>
<td>QT Studies</td>
<td>Dyslipidemia</td>
<td>Psoriasis</td>
<td>HIV*</td>
</tr>
<tr>
<td>COPD</td>
<td>Major Depressive Disorder</td>
<td>Cardiovascular v3*</td>
<td></td>
</tr>
<tr>
<td>CV Imaging (Echo)</td>
<td>Parkinson’s Disease v2*</td>
<td>Oncology v4*</td>
<td></td>
</tr>
<tr>
<td>Diabetic Nephropathy</td>
<td>Maintenance Update TBD*</td>
<td>Maintenance Update TBD*</td>
<td></td>
</tr>
</tbody>
</table>

Projects in Italics are candidates to be scheduled and may be subject to change.

*Indicates Project Proposal Summary and Approval Pending

### CDISC Therapeutic Area Data Standards (published prior to CFAST):

<table>
<thead>
<tr>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer’s v1</td>
<td>Pain</td>
<td>Polycystic Kidney Disease</td>
</tr>
<tr>
<td></td>
<td>Tuberculosis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Parkinson’s Disease</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Virology</td>
<td></td>
</tr>
</tbody>
</table>
Therapeutic Area Standards

Scientific and Technical Input via SC and SAC

Medical TA Experts Participate in Scoping

Create Data Element Concepts

Develop SDTM, CDASH, ADaM metadata

Develop & Review TA User Guide prior to Publication

Assist in Advocating Adoption; Provide Feedback for Updates to User Guides
Increase Operational Efficiency

- Reduce EDC setup time
  - CRO can copy pre-validated forms & associated edit checks
- Reduce cost for EDC build, edit check programing, data transfer mapping
- Reduce number of queries
- Avoid data conversion
- Minimize re-mapping at time of submission
- Reduce EDC training & learning time
- Better Site and CRA experience
Improve Data Quality

- Suitable for the purpose of regulatory submission
- Build quality into design of eCRF
  - Consistent user interface & question text
- Data consistency across all CROs
  - Sponsor and CRO access common library
- Pre-validate: reduce common data & mapping errors
- Promote the use of common reporting tools
  - Facilitate data review and trending
- Data traceability from submission back to source
End-to-end Implementation of CDISC
Comments & Questions