Pathway into the Future for Standards Development and Delivery

April 2010

Bron W. Kisler, CDISC Senior Director
bkisler@cdisc.org
CDISC Mission

To develop and support global, platform-independent data standards that enable information system interoperability to improve medical research and related areas of healthcare

Strength through collaboration.
CDISC Snapshot

• Global standards organization initiated in 1997 as a volunteer group; incorporated as non-profit in 2000
• Established worldwide standards to support the electronic acquisition, exchange, analysis and submission of clinical research data
• 200+ corporate sponsors and member organizations (Global Pharmas, Service Organizations, Technology Vendors, Academic Institutions)
• Established Coordinating Committees and annual events in North America, Europe and Japan
• Expanding activities in China, India and Brazil
• Liaison A organization to ISO TC-215; member of Joint Initiative Council
CDISC Standards and Data Flow

Healthcare Standards

Health Level 7 (HL7) Reference Information Model (RIM)

Legend

- ODM XML (transport)
- SDTM & ADaM (content)
- Protocol Information (content)
- Source Data (other than SDTM/eCRF data)

Clinical Research Standards (CDISC)

- Electronic Health Record
- Clinical Study Data
- Patient Info
- CDASH - eCRF Study Data (defined by SDTM)
- LAB Data
- Study & Analysis Data
- Operational & Analysis Database
- Study Reports
- Study & Analysis Data
- Reports or Regulatory Submissions

Protocol Representation
- Study Design Analysis Plan
- Study Protocol
Opening Remarks by CDISC Chair, Dr. Edward Helton (SHARE presentation Oct. 2009)

- The Clinical Data Interchange Standards Consortium (CDISC) is a global standards development organization.
- CDISC has developed a suite of standards to support the clinical research process from protocol representation through analysis and reporting, with a focus to date on safety data domains.
- These encompass protocol-driven research, including regulated research.
- **CDISC and its members want additional content standards (to support efficacy, eligibility, clinical content) to be developed more rapidly.**
- CDISC standards are open and freely available ([www.cdisc.org](http://www.cdisc.org))
What is CDISC SHARE?

A globally accessible electronic library built on a common information model, which (through advanced technology) enables precise and standardised data element definitions that can be used in studies and applications to improve biomedical research and its link with healthcare.
Why SHARE?

• Integrate current CDISC standards and extend them with efficacy-related domains / content and other needed standards in an electronic, accessible format

• Help to ensure terminology for research standards is consistent as much as possible with that needed for other purposes (e.g. EHRs, quality reporting, public health, safety monitoring)

• Strengthen the link between clinical care and research such that research results can inform healthcare more quickly

• Develop a ‘reference’ or target set of standards without duplication of existing standards (e.g. target for mapping legacy/retrospective data)
SHARE Project: Inception Phase
(March – December 2009)

• Goal: Evaluate feasibility; determine future path

• Scope & Vision Document
  – Business Requirements
  – Governance Process, Workflow and Requirements
  – Detailed Stakeholder Assessment
  – Business Models

• Pilot
  – Technology and tools (Mayo/NCI LexGrid)
  – Terminologies and vocabularies for harmonization (n=7)
  – New/comparable content (Oncology) from 5 sources (Mayo, MD Anderson, GSK, Genzyme, Lilly)
Within an organization the extraction and formatting of existing definitions for loading into the CDISC SHARE environment.

Pilot used an excel spreadsheet
Also used a Lilly ODM file exported from their internal system, which was then converted to Excel spreadsheet form
CDISC SHARE: Link

Attempt to align an organization’s internal world with the outside world

BRIDG & CDISC
NCIt & Meta
SNOMED CT
UCUM, ICD 10

Cloud of Knowledge
CDISC SHARE: Merge

By using the links the “machine” can group similar source definitions
Group like data elements and definitions submitted by different organizations and worked on by the community to align and agree to the consensus version.
CDISC SHARE: Harmonize

- Value Set = A,B,C
- Value Set = X,Y
- Value Set = A,X,Y
- Value Set = N,M,O,P,Q,R
- Value Set = ???

Align value sets. Source these from standard terminologies if possible.
Harmonize – “Clumps”

Form “logical” groupings (e.g. BP = SYSBP, DIABP, Body Location, Body Position etc.)
New Definition

- Long Name
- Short Name
- Definition
- Data Type
- Value Set
- Concepts Linked To
- Sources of Definition
Pilot Aims

• Primary
  – Determine whether definitions taken from multiple sources can be merged into a single reference version agreed to by all parties, and can this be done within a timeframe that makes business sense
  – Determine whether high-quality definitions can be created, and can existing ontologies help in ensuring such while avoiding duplicate definitions from being created

• Secondary
  – Provide any relevant lessons to future development work
  – **Refine user, business and governance requirements**
CDISC: Focus on Standard Content

CDISC Partners
( Technology )

Software

Hosting

Inception

Elaboration

Construction

Transition

Production

CDISC & Collaborators
( Content )
CDISC: Focus on Standard Content

**NCI – tool development**

- **Technology**
  - Servers, Repository, Software, etc. = NCI
  - Hosting = Central and/or Federated

- **Content** = CDISC and Collaborators
  - **Content Governance** = CDISC and Collaborators

- **Business Model and Funding**
  - CDISC - Core Standards
  - External - Therapeutic Area Standards

**CDISC SHARE – content**
NCI Partnership

- NCI has produced and hosted CDISC terminology through its Enterprise Vocabulary Services (EVS) since early 2003.
- NCI also produces and hosts terminologies for FDA, NIH, and a broad array of research and healthcare organizations nationally and internationally through EVS.
- NCI is migrating to new tools for semantic management including a next generation metadata repository (MDR).
- NCI has offered to include all CDISC SHARE requirements in its repository development process.
- NCI has requested direct CDISC participation in the development team.
- NCI wishes to support tools for CDISC SHARE that will be customizable, global, and cover all therapeutic areas.
NCI Partnership (2)

- NCI’s MDR development will be based requirements from a wide variety of groups to include research and broader healthcare standards; SDOs - HL7, CDISC and other; regulatory entities; pharmaceutical; providers and vendors
- The new ISO 11179 standard MDR will be based on a federated, distributed architecture; meaning it will be decentralized, allowing for multiple peer repositories
- A platform independent model, to be openly shareable
- Allows for modular development of many and varied customized applications and services for different users, but with a common foundation and generic API
- **Open, Platform & Vendor Neutral, Distributable, Shareable**
Why Partner with NCI?

• Proven terminology partnership with rapid turn-around service and support for accelerating standards timelines

• Reliable and globally accessible Best-of-Breed infrastructure for standards publication (DE’s, Terminology, DAMs)

• Ability to “harmonize” real-time with key partners without costly standards mapping

• NCI plans align well with CDISC’s mission as well as near and long-term strategy (e.g. new disease areas)

• Offer to roll-in requirements (CDISC, HL7 CIC, others) and to provide extensive resources and expertise is unprecedented

• Open, Platform & Vendor Neutral, Distributable, Shareable where all CDISC standards can be aligned and published
NCI-EVS NCIt content stakeholders

<table>
<thead>
<tr>
<th>Terminology Subsets</th>
<th>NCI-GLOSS</th>
<th>RAND</th>
<th>SEER</th>
<th>Swiss-Prot</th>
<th>ULCUM</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRCH</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDISC</td>
<td>6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDISC</td>
<td>6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTCAE</td>
<td>9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTCAE</td>
<td>9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTRM_ID</td>
<td>6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTRM_ID</td>
<td>6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FDA</td>
<td>19%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FDA</td>
<td>19%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICH</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICH</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICSR</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICSR</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISO</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISO</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JAX</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KEGG_ID</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KEGG_ID</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTH</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTH</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTH</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTH</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCI Only</td>
<td>42%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCI Only</td>
<td>42%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCI Only</td>
<td>42%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCI Only</td>
<td>42%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCI-GLOSS</td>
<td>3%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCI-GLOSS</td>
<td>3%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCI-GLOSS</td>
<td>3%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCI-GLOSS</td>
<td>3%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAND</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAND</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAND</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAND</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEER</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEER</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEER</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEER</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swiss-Prot</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swiss-Prot</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swiss-Prot</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swiss-Prot</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ULCUM</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ULCUM</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ULCUM</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ULCUM</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
NCI-EVS Terminology Services

• Subject Matter Expertise
• Definition writing and analysis
• Terminology tagging and sub-setting
• Terminology coding that ensures cross-harmonization with key partner organizations
• Terminology requests and maintenance
• Links to other controlled terminologies as needed (e.g. FDA, MedDRA, ISO, UCUM etc.)
• Extending into new disease areas
Existing CDISC data standards currently coded in NCI-EVS NCIt.

<table>
<thead>
<tr>
<th>CDISC Controlled Terminology</th>
<th>Numbers</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDTM</td>
<td>3360</td>
<td>Published</td>
</tr>
<tr>
<td>CDASH</td>
<td>82</td>
<td>Published</td>
</tr>
<tr>
<td>ADaM</td>
<td>26</td>
<td>Published</td>
</tr>
<tr>
<td>SEND</td>
<td>1928</td>
<td>Pilot</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CDISC (SDTM+CDASH) Metadata</th>
<th>Numbers</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domains</td>
<td>45</td>
<td>Published</td>
</tr>
<tr>
<td>Observation Classes</td>
<td>200</td>
<td>Harmonization</td>
</tr>
<tr>
<td>Variables</td>
<td>654</td>
<td>Harmonization</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CDISC Disease Standards</th>
<th>Numbers</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiovascular</td>
<td>25</td>
<td>Mapping in EVS</td>
</tr>
<tr>
<td>Tuberculosis</td>
<td>130+</td>
<td>Mapping in EVS</td>
</tr>
<tr>
<td>Polycystic Kidney Disease</td>
<td>200+</td>
<td>Mapping in EVS</td>
</tr>
</tbody>
</table>

Published in .xls and .txt for download (.xml coming soon!)
Current NCI Semantic Infrastructure

Information Models

Data Elements

Vocabulary

https://cabig-kc.nci.nih.gov/Vocab/KC/index.php/SI_Conop_Existing_caBIG_Semantics_Implementation
Future Semantic Infrastructure

CDISC SHARE Project Plan

TEAM 1 - SAFETY CONTENT
TEAM 1 - EFFICACY CONTENT
TEAM 1 & 2 - REQUIREMENTS
TEAM 2 - GOVERNANCE & METADATA MODEL
NCI BASED TECHNOLOGY DEVELOPMENT
USE OF THE NCI’S “REPOSITORY”
EDUCATION & USER SUPPORT
SHARE COMMUNICATION PLAN

Team 1 Tasks - 1) Safety - CDASH & SDTM initially then all applicable CDISC production standards prepared for EVS inclusion. 2) Efficacy - TB & Cardiovascular standards initially and then all applicable CDISC therapeutic area production standards prepared and included in the NCI thesaurus (NCIt).

Team 2 Tasks - 1) Refine requirements and send to NCI in 2 packages (high-level Requirements and detailed requirements). 2) Establish governance for entering CDISC production standards into NCIt and refine draft governance. 3) Develop a metadata model for the creation of multi-dimensional data elements.

All new CDISC safety and efficacy production standards entered into the NCIt post the CDISC consensus process.

Rhonda Facile, CDISC
Next Steps

• Complete CDISC SHARE use cases and 1st phase of requirements (high level)
• Continue gathering and refining requirements throughout 2010 and beyond
• Align existing CDISC standards – CDASH, SDTM, BRIDG, Controlled Terminology
• Continue developing new disease-specific standards – Infectious Diseases, Cardiovascular Disease, Oncology, Neurological Disorders
• Accelerate CDISC standards development process
Strength through collaboration.

As a catalyst for productive collaboration, CDISC brings together individuals spanning the healthcare continuum to develop global, open, consensus-based medical research data standards.

Special Thanks to: Margaret Haber (NCI EVS), Erin Muhlbradte (NCI EVS), Rhonda Facile (CDISC), Dave Ibersen-Hurst (CDISC)