Building (and Rebuilding) the CDISC Toolbox

Frank Dilorio, CodeCrafters, Inc., Philadelphia PA
Jeff Abolafia, Rho, Inc., Chapel Hill NC

Background

- CDISC data and documentation standards have had a significant impact on pharma and CRO work flow and processes
- The standards have not always been warmly received
- But they can be used not just for regulatory compliance but to improve workflow efficiency
- This presentation describes tools we’ve used and developed for CDISC implementation at Rho over the last 10+ years
- We’ll talk about missteps as well as successes

Organization

- Our Lessons Learned approach is most effective if we follow the timeline as we learned the lessons.
- Shortly pre-CDISC to present day is divided into 6 phases.
- For each phase we discuss changes to the regulatory environment for data standards and models, what tools we developed or adopted, and how workflow was affected.

Overview: Development Phases

Frame the discussion around development phases at Rho.

Note:
- Pretty arbitrary
- Fairly sequential
- Somewhat overlapping

- Early eSub
- First external standards
- Exploit the standards
- New standards
- Revised standards
- Revised documentation standard

But Wait, Haven’t We Been Here Before? Repeatedly?

Another metadata & tools paper? Really ...?

Not on the List: Pre eSub

Actually more complex than this slide implies, but this is the essence of the process.
Prior to FDA’s formal adoption of CDISC data models and standards.
Emergence of metadata and tools to manipulate it.

First external standards
Exploit the standards
New standards
Revised standards
Revised documentation standard

Early eSub: Summary
CDISC
• c. 1997: “a good idea” but no mandate for implementation
• Early ODM, SDTM suggest importance of metadata
• FDA still accepting non-standardized data, define.pdf

Highlights
• Initial (non-CDISC) metadata
• Awareness of need for tools to use the metadata
• Move away from Word docs as programming specs

Early eSub: Comments

Standards
• 1999 regulatory submission guidance
• CRT dataset and documentation format not rigorously defined

Metadata
• Initially, focus on satisfying creation of XPTs, define.pdf
• Seeded from SAS Dictionary Tables
• Store using SAS/Share, Excel
• Key: structured, machine-readable

Tools
• Print specs from m’data rather than Word
• Create %attrib stmts from m’data; also var. renaming, formatting, ...
• Use m’data when building XPTs, define.pdf

Other
• Train users in the basics: meaning, entry of m’data; use of tools
• Key: need for m’data and tools to make it easily accessible

First External Standards: Summary

CDISC
• Initial SDTM: data standard, presented as metadata
• Documentation standard: define.xml
• Relatively little implementation guidance

Highlights
• Expanded skill set (XML, XSL)
• Validation to the standard
• Altered work flow
• Learn as we go. Combine R&D with project work. Not pleasant.

First External Standards: Comments

Standards
• Compliance / validation
• Education / training
• Dual annotation of CRF
• New (SDTM) workflow stream

Metadata
• More to populate
• Create “Gold Standard”
• Refine interface (consult end-users)
• Re-think storage: MDB

Tools
• Updated spec printer (for SDTM)
• Domain programming tools
• Validation (Domains, XML)
• Revise study setup
• Build XML and render as PDF
• Legacy conversions

Skill Set
• Familiarize w/ XML, XSL, SAS XML mapper
• Define and manage revised workflow (for SDTM)
• Heavier use of ODS

2004: FDA accepting SDTM. External standards place greater requirements on content, structure and documentation of eSub deliverables.
Expand earlier use of m’data and tools.
Clients: legacy conversions
### Movement toward mandatory use of SDTM.

Realize advantages of standards: significant effort in tool building, making processes more efficient.

### Exploit the Standards: Summary

#### CDISC
- SDTM adoption by FDA and many sponsors
- Maturing of other standards
- More opportunities for exchange of information: forums, user groups

#### Highlights
- Emphasis on tool development
- Interface: improved; becomes corporate standard
- Metadata in Oracle facilitates development of repository
- CDISC models becoming familiar in-house

### Exploit the Standards: Comments

#### Standards
- SDTM in-house standard for collected data
- Clients requesting SDTM, draft ADaM, results-level
- In-house: standardize on ADaM (like)

#### Metadata
- Improved interface (add CT, value-level, based on user feedback). More QC in the interface.
- Store in Oracle
- Add repository

#### Tools
- Domain splitter
- ISO macros; compare m'data across studies
- Validation (home-grown, OpenCDISC)
- Refine eSub tools

#### Skill Set / Other
- Training
- Oracle
- “Build and buy”
- Handle much larger volume of studies, compared to earlier phases.

### New Standards: Summary

#### CDISC
- Release of ADaM, CDASH
- Early PRM
- Frequent CT releases
- Therapeutic area standards

#### Highlights
- Opportunity for end-to-end standardization
- ADaM implementation
- CDASH “awareness”
- Increased metadata requirements

### New Standards: Comments

#### Standards
- Potential for end-to-end
- Compliance / validation
- Links from CDASH to SDTM, SDTM to ADaM, ADaM to displays
- New workflow streams

#### Metadata
- Yet more to populate
- Display-level: refine proj. mgt. system
- SDTM-ADaM linkage in interface

#### Tools
- TFL library (ADaM std’on makes this feasible)
- Tools to access results-level m’data (titles & f’notes, other)
- Expanded validation

#### Skill Set
- CDASH: workflow (3rd party - Medidata), XML
- eSub validation training
- Project volume ➔ more hands ➔ need for better “onboarding”
Revised Standards: Summary

- Accelerated pace of updates
- Need for identification of changes across non-harmonized standards
- Clients: standardize on a version throughout a study, or use different ones, then adjust for ISS/ISE?

Highlights
- Automate population of new versions’ Gold Standard
- Enhance tools for validation, study setup
- Implications of more people creating eSub deliverables

Revised Standards: Comments

Standards
- New IGs for SDTM, CTs, others
- Maintain version continuity throughout study
- Our, 3rd party tools need to keep pace

Metadata
- Database management to apply different versions
- Plus coding updates for interface
- CDASH m'data library

Tools
- Project setup: assign, maintain SDTM, CT versions for all studies
- OpenCDISC (compare checks; create new version config files; how to handle “semi-false” positives)

Other
- Control the beast: soft dev Best Practices (better doc’on, version control)
- Fewer client-specific accommodations needed, requested

Revised Documentation Standard: Summary

- Define-XML Version 2
- Significant amount of new metadata
- Eventual implementation of Results metadata

Highlights
- Opportunity for retooling:
  - Harmonized SDTM-ADaM interface
  - Updated define.xml/pdf and support programs
  - Need for new documentation, (re-)training
## Revised Documentation Standard: Impact

<table>
<thead>
<tr>
<th>Standards</th>
<th>Metadata</th>
<th>Other</th>
</tr>
</thead>
</table>
| • New Define-xml not just incremental change | • Much more to populate  
• Single SDTM-ADaM GUI  
• Create new tables (study-level, global settings)  
• Look to integrate results-level from other in-house system | • Standardize on CDISC style sheet (modify XSL if client requests)  
• Documentation: updated, new  
• Training (both “old” and new hires) |
| • Will incorporate Results-level  
• More complex validation |                            |                                           |
| • Complete rewrite of older (10+ year-old) tools  
• Validation (3rd party + in-house)  
• Autopopulating some m’data becoming possible  
• Harmonize Results (proj. tracker) with eSub systems |                                           |
### Summary: Workflow, Tools, Metadata

<table>
<thead>
<tr>
<th>dms</th>
<th>CRF program</th>
<th>draft</th>
<th>final</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDASH</td>
<td>CRF XML</td>
<td>draft</td>
<td>“final”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CRT/SDTM</th>
<th>spec program</th>
<th>draft</th>
<th>“final”</th>
<th>eSub</th>
<th>XML docs XPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anl/ADaM</td>
<td>spec program</td>
<td>draft</td>
<td>final</td>
<td>eSub</td>
<td>XML docs XPT</td>
</tr>
<tr>
<td>TFL</td>
<td>SAP spec program</td>
<td>draft</td>
<td>final</td>
<td>eSub</td>
<td>XML docs XPT</td>
</tr>
</tbody>
</table>

**Metadata**
- prj mgt, displays
- study data CT
- CDISC global admin

**Tools**
- config m’data val’on
- gui eSub tfl util

---

**Summary: Workflow, Tools, Metadata**

<table>
<thead>
<tr>
<th>dms</th>
<th>CRF program</th>
<th>draft</th>
<th>final</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDASH</td>
<td>CRF XML</td>
<td>draft</td>
<td>“final”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CRT/SDTM</th>
<th>spec program</th>
<th>draft</th>
<th>“final”</th>
<th>eSub</th>
<th>XML docs XPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anl/ADaM</td>
<td>spec program</td>
<td>draft</td>
<td>final</td>
<td>eSub</td>
<td>XML docs XPT</td>
</tr>
<tr>
<td>TFL</td>
<td>SAP spec program</td>
<td>draft</td>
<td>final</td>
<td>eSub</td>
<td>XML docs XPT</td>
</tr>
</tbody>
</table>

**Metadata**
- prj mgt, displays
- study data CT
- CDISC global admin

**Tools**
- config m’data val’on
- gui eSub tfl util

---

**Validation:**
- OpenCDISC custom
- Gold Stds

**File locations**
- seed study m’data
- create dir structure
- pgm templates

---

**XML header fields**
- %attrib
- %domSplit

**Export m’data**
- %export
- %PDF
- %createZIP

**Gold Stds**
- %results
- %pgm templates

---

**Repository**
- gui
- eSub tfl util

**ISO**
- ttl/’notes

---

**OpenCDISC** BAT, CFG

---

**6 PharmaSUG SDE, Cary NC, September 29, 2014**
Key Lessons Learned

Store data specs in **machine-readable** metadata format (for re-use).

Importance of **seamless access** to metadata: interface, tools

Treat metadata as a **valued resource** – Oracle (or similar), not Excel(!)

**Validation:** complex, time-consuming

The CRO dilemma: exploit the Standards but **maintain flexibility** in tools: every sponsor wants their own version

Tools – built or bought – must be **adaptable**, reacting quickly to changes in standards

**Best soft dev practices** for tools SAS/other tools (doc’on, life cycling, validation)

Looking Ahead: “Phase Next”

**Models**
- Better end-to-end integration of CDISC models, especially PRM, CDASH

**Strategy, Coding**
- Better manage OpenCDISC checks; automate inclusion of output into Reviewer’s Guide
- Extend reach of standard displays (using Results-level metadata)
- Auto-generated code from metadata
- Continually review the market, revisiting subsystem “Build or Buy”

**Training**
- More documentation
- Clearly identify eSub/other m'data needs
- Clear, shared understanding of what ’s required for eSub completeness
- Documentation of validation errors and anomalies

**Constant Change!**
- Building and maintaining systems requires ongoing organizational commitment to R&D
Thanks for coming!

- Questions?
- Comments?

Contact
- jAbolafi@RhoWorld.com
- Frank@CodeCraftersInc.com