Use and Evaluation of Standards for Investigator-Initiated Studies: Preliminary Results

M. Theresa Perry
RTRN Data and Technology Coordinating Center (DTCC)

2010 AMIA Summit on Clinical Research Informatics
March 12-13, 2010
Parc 55 Hotel San Francisco
San Francisco, California
The CRNFA Project

Overall goal is to determine what steps (and related costs) would be needed to map from non-regulated, non-standardized data collection to standards-compliant.

- Select Standards
- Select Studies to Map
- Conduct Independent Mappings for Each Study
- Conduct Secondary Independent Mappings for Each Study
- Determine Conformance to Standards for Mappings
- Identify Challenges to Conformance and Mapping Processes
- Conduct Full Adjudication
- Complete Economic Analysis

{Final Results Pending Completion at time of Submission}
CRNFA Team

1. Duke University, Durham, North Carolina - LEAD
   James Tcheng, Eric L. Eisenstein, Meredith Nahm, Anita Walden

2. Jackson State University, Mississippi
   M. Edwina Barnett, M. Theresa Perry, Andrew Dent

3. University of Puerto Rico, Puerto Rico
   Jose Conde

4. Vanderbilt University, Nashville, Tennessee
   Paul Harris

5. Meharry Medical School, Nashville, Tennessee
   Ahmed Fadiel

6. Clinical Data Interchange Standards Consortium
   Chris Tolk
Standards Selection

• Standards Selected to
  – Cover content for academic studies
  – Provide data in a format facilitating interoperability
  – Constrain data element representation to ensure a unique mapping from the study to the standard

Our Selection:

• CDISC Study Data Tabulation Model (SDTM) V1.2
  – Well established standards
  – Covers the research domain
  – Developed for clinical trials use
  – Implementation Guide (IG) V3.1.2 as reference
Studies Selected For Mapping

- Availability, varying study design, and not all clinical trials
- RTRN Vitamin D Study
  - Ongoing investigator-led multi-institutional study of the effects of Vitamin D on cardiovascular function in African Americans
  - Casebook is relatively small (21 pages)
  - All but blinded data is collected through an EDC system
- Jackson Heart Study
  - Largest longitudinal national study on cardiovascular disease among African Americans, including genetic and environmental risk factors
  - Large casebook (> 100 pages)
  - Most data is collected via paper questionnaires and entered at the JHS Data Center
- Meharry A-HeFT Study
  - Phase III study of African American patients with NYHA Class III-IV chronic, stable HF
  - On-site visits are augmented with monthly IVRS phone visits and quality of life questionnaires
Mapping Methods

- No subject data was accessed or used
  - Only blank CRFs were used
  - Domains and fields mapped from CRFs to the Standards
- Each team independently mapped CRFs per SDTM
- Each CRF was mapped a second time for comparison to the initial mapping
- Teams kept lists of items that required help from CDISC for appropriate mapping
- Questions posed to CDISC as needed throughout process
- Final adjudication complete for one set of CRFs
Initial Independent Mappings

- CDISC workshop at DTCC in May
- Partial mapping of Vitamin D just after the workshop, involving total of 9 people
  - For illustrative methods
  - Confirm understanding of mapping methods
- Complete initial mapping of Vitamin D at DTCC
- Initial mapping of JHS at DTCC
- Initial mapping of A-HeFT at Duke
Secondary Mappings

- Secondary Mappings conducted independently (Duke or DTCC)
- Quality Control via Independent Reviews
  - Mappings from DTCC reviewed at Duke
  - Mappings from Duke reviewed at DTCC or Duke third party
Agreement

• All three sets of findings were submitted for evaluation of domain and variable level agreement between sites

• The most common disagreements were when selecting domains, however once domains were in agreement, variable agreement was almost always 100%
## Annotator Agreement

<table>
<thead>
<tr>
<th>CRF Module/ Section</th>
<th>% Domain Agreement</th>
<th>% Variable Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Header</td>
<td>44.4 %</td>
<td>100.0 %</td>
</tr>
<tr>
<td>Informed Consent</td>
<td>20.0 %</td>
<td>100.0 %</td>
</tr>
<tr>
<td>Inclusion/Exclusion</td>
<td>100.0 %</td>
<td>100.0 %</td>
</tr>
<tr>
<td>Demographics</td>
<td>30.8 %</td>
<td>100.0 %</td>
</tr>
<tr>
<td>Medical History</td>
<td>71.4 %</td>
<td>100.0 %</td>
</tr>
<tr>
<td>CRF Lab</td>
<td>96.0 %</td>
<td>100.0 %</td>
</tr>
<tr>
<td>Vital Signs</td>
<td>71.4 %</td>
<td>100.0 %</td>
</tr>
<tr>
<td>Randomization, ABPM status</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>Dexa scan</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>ECG</td>
<td>100.0 %</td>
<td>100.0 %</td>
</tr>
<tr>
<td>Stored Specimens &amp; Biomarkers</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>Study Drug</td>
<td>40.0 %</td>
<td>100.0 %</td>
</tr>
<tr>
<td>Completion</td>
<td>57.1 %</td>
<td>75.0 %</td>
</tr>
<tr>
<td>Protocol Violations</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>Concomitant Medications</td>
<td>66.7 %</td>
<td>75.0 %</td>
</tr>
<tr>
<td>Adverse Events</td>
<td>80.0 %</td>
<td>83.3 %</td>
</tr>
</tbody>
</table>
Conformance

All three sets of findings were submitted to CDISC for conformance review

- Adjudication for Vitamin D is complete
- Full adjudication is pending
- Conformance to standards was relatively high, but not at 100%
  - Percent conformance for domain was 81.2% for annotator 1 and 89.5% for annotator 2
  - Variable % conformance was even higher
    - 98.3% for annotator 1
    - 99.4% for annotator 2
Results

1) All data elements could be represented in the SDTM standard

2) Data logically grouped in study data collection forms often required mapping across more than one domain

3) There were instances where two representations of the same data element were both conformant and different
Overcoming Challenges

• Although we used the same standards, we did not achieve a unique mapping of data elements to the standard
  – This is critical for automating the process of mapping to standards across studies
  – Full adjudication may present a unique mapping
  – For industry studies unique mappings are more likely
  – Use of CDASH (CDISC Standard that directly maps to SDTM) would increase both conformance and consistency
Suggestions for Success

• For CDISC qualifiers that are in the protocol but not on the CRFs (such as position for BP), use ‘pre-loaded’ fields prior to entry… maybe give a flag for items not included when CRFs are ‘loaded’ (i.e. ask for protocol-derived data as needed)

• To have standards that meet the needs of academic studies, academic institutions will need to become more active in standards organizations
This project is funded by grant #8059-S06 Clinical Research Network Feasibility Award (CRNFA) as a subcontractor to Westat’s NIH Inventory and Evaluation of Clinical Research Networks contract with the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH).

The RTRN Data and Technology Coordinating Center is supported by grant number U54RR022762 from the National Center for Research Resources, National Institutes of Health (NIH). This grant also receives co-funding from the National Center on Minority Health and Health Disparities, NIH.

Presentation contents are solely the responsibility of the authors and do not necessarily represent the official views of RTRN, Westat, NCRR, or NIH.