Case Report Form
Harmonization and Standardization:
Creation of a Library of CRFs that Implement caBIG® and CDASH Data Standards

Dianne M. Reeves, RN, MSN
Center for Biomedical Informatics and Information Technology

November, 2009
Agenda

- Background
- Initiative Status
- Application of Lessons Learned
  - Inventory and Traceability of Submitted CRFs
  - Group Dynamics
- CRF Guide Book
- CDASH Collaboration
- Next Steps
  - Implementation
  - Measures of Success
  - Return on Investment
  - Making content available
CTWG Vision

*Enhance the best of all the components of the NCI-supported clinical trials system to develop a cooperative enterprise built on a strong scientific infrastructure and a broadly engaged coalition of critical stakeholders.*

Report of the Clinical Trials Working Group of the National Cancer Advisory Board

Restructuring the National Cancer Clinical Trials Enterprise

June 2005
CRF Harmonization and Standardization

- Establish a Comprehensive Community-Accessible NCI Clinical Trial Database
- Promote System Harmonization & Interoperability between NCI’s Clinical Trials systems
- **Establish Core Library of Standardized Case Report Forms (CRFs) through Stakeholder Consensus**
- Develop a Credentialing System for Investigators and Sites
Workflow and Processes

Harmonization
- Inventory Content
- Prioritize and Partition Content
- Consensus-based Harmonization
- Community review and approval
- Stakeholder approval

Standardization
- Community vetting of standard
- Registration in metadata registry
Process for caBIG® CRF Standardization

CRF Question Partition Categories

Mandatory:
A data collection variable that must be on the CRF (e.g., a regulatory requirement if applicable). Must be used.

Conditional:
A data collection variable that must be collected on the CRF for specific cases that may be dictated by local or sponsor defined business rules.

Optional:
A data collection variable that is available for use if needed. Usually collected due to the design of the protocol, like in instances where marital status or educational level are important indicators in a study.

Non-harmonized:
A data collection variable that is, by consensus, primarily belong to a different CRF module or is not belonging to any defined module.

Done!!

1. Review sample case report forms (CRFs) submitted. Assemble into master spreadsheet of forms, questions, definitions, values, and instructions.**

2. WG review CRFs and partition questions as Mandatory, Conditional, Optional or Non-harmonized in master spreadsheet. Make note of condition rules and special instructions

3. Aggregate and identify common questions across forms

4. WG Re-Review and harmonize (resolve discrepancies in partition decisions)

5. Match CDEs for all mandatory, conditional and optional questions. Work with caDSR Master Curator. Workgroup consensus on final CDE list.

6. WG prepares, review and finalizes final deliverable, folding in vocabulary WG recommendations if applicable

7. caBIG Community Review (feedback reviewed and responded to by WG)

8. Review of content by NCI CBIIT to insure collection due to the design of the protocol, like in instances where marital status or educational level are important indicators in a study.

9. CTROC Review and Approval

10. VCDE approval of CDEs as caBIG standards

11. WG Provides continual CRF Maintenance

**In some modules, such as Agents, separate parallel workgroups may be formed to develop recommendations around vocabularies that insure unique and unambiguous identification of key data (such as study agent name).
Hypothetical Assembly of a *Disease History and Prior Therapies* CRF (for paper) using caBIG® standardized CRF Modules

**Header module**
- Insert Mandatory, Conditional CDEs (if applicable), select Optional CDEs as desired.

**Staging/Extent of Disease**
- Insert Mandatory, Conditional CDEs (if applicable), select Optional CDEs as desired.

**Agents (Prior/Post Therapies)**
- Insert Mandatory, Conditional CDEs (if applicable), select Optional CDEs as desired.

**Outcomes**
- Insert Mandatory, Conditional CDEs (if applicable), select Optional CDEs as desired.

**Disease History and Prior Therapies**
- **Study ID**
- **PI**
- **Course#**
- **Patient ID**
- **Institution ID**
- **Pg#**

- **Disease status**
- **Stage**
- **TNM**
- **Primary Site**
- **Metastatic Sites**

- **Prior Agent Name**
  - *(date type) Therapy Start Date***
  - *(date type) Therapy Stop Date***
  - *(dose type) Total Dose Unit***

- **Disease response**
  - *Date of response assmt***

**Sign Off / Footer**
- Insert Mandatory, Conditional CDEs (if applicable), select Optional CDEs as desired.

- **CRC Name**
- **Signature**
- **Date**
- **PI Reviewed**
- **Signature**
- **Date**
Round 1: Demography

- Total of 15 CDEs
  - 13 vetted and approved as caBIG® Standards
  - 2 Legacy CDEs for replacement

4 Mandatory elements
- Race
- Ethnicity
- Date of Birth
- Gender

6 Conditional elements

5 Optional elements
**Round 2: Seven Modules of Content**

- Total of 78 CDEs; 20 Mandatory
- Now to caBIG VCDE for Standards review

<table>
<thead>
<tr>
<th>Module</th>
<th>Mandatory</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protocol Deviations</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Adverse Events</td>
<td>9</td>
<td>28</td>
</tr>
<tr>
<td>Participant Identifier</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Enrollment</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Registration</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>Medical History</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Physical Examination</td>
<td>1</td>
<td>6</td>
</tr>
</tbody>
</table>
Round 2: Adverse Events

Mandatory – 7 Elements

SAE – 2 Additional Elements

Conditional – 12 Elements

Optional – 7 Elements

CTCAE v4.0 elements will be added to v3.0 content

• Will replace v3.0 content for prospective trials
## Round 3: Seven Modules of Content

- Total of 131 CDEs; 46 Mandatory
- Next to NCI review
- caBIG VCDE for Standards review after Round 2

<table>
<thead>
<tr>
<th>Module</th>
<th>Mandatories</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lab Tests</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>Staging/Extent Disease</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Study Agent</td>
<td>11</td>
<td>26</td>
</tr>
<tr>
<td>Prior/Post Agent</td>
<td>11</td>
<td>21</td>
</tr>
<tr>
<td>Agent Account/Compliance</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>Concomitant Agents</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>Study Outcomes</td>
<td>5</td>
<td>27</td>
</tr>
</tbody>
</table>
Round 3: Outcomes

Of Note:

- Call for CRFs did not specifically identify self-reported measures such as QOL, Pain, FACT, other tools
- Participants felt that self-reported measures and surveys constituted additional content beyond the scope of the present workgroup
- NCI Leadership identified the need to include Patient Reported outcomes as additional content
- Glossary initiated to add terms to Instruction Manual
Summary: Rounds 1 through 3

- 15 Modules of Content
- 188 Data Elements Harmonized; 70 Mandatory
- Round 1 Module Standardized September 2009
- Rounds 2 – Standardization review November 2009
- Round 3 – NCI review November 2009
- Round 4 – To finish late November 2009
- Round 5 – Early 2010
Round 4: Seven Modules of Content

- Imaging/Radiology
- Non-Agent Intervention Treatment
- Diagnosis/Pathology
- Vital Signs
- Eligibility
- Header/Footer
- Off Study/Off Treatment
Round 4: Application of Lessons Learned

Creation of Cross-Round Inventory of CRF Content

- Limitation: Will slow down beginning of Round 4
- Benefit: Eliminate need to collect content each Round
  - Permit Traceability of Content
  - Permit Estimate of Effort to bring CRFs into compliance
  - Create CRF guidance for contributing groups
- Benefit: Use Metrics to measure outcomes
  - Number of CRFs
  - Number of Variables
  - Number of Resulting Harmonized variables
Round 4: Progress to Date

- 7 Workgroups nearing completion
- Attendance 6 – 25 per meeting
- Improved efficiency to partition content
- Improved ability to work cross-Workgroup
- Rework of CRF Guide Book
  - Header CRF has identified ‘Footer’ need
  - Non-Agent Intervention groups using template approach
  - Rework of the Imaging and Radiology content
<table>
<thead>
<tr>
<th>Critical Mass: Core for Trial Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concomitant Agents</td>
</tr>
<tr>
<td>Adverse Events</td>
</tr>
<tr>
<td>Course Assessments</td>
</tr>
<tr>
<td>Eligibility</td>
</tr>
<tr>
<td>Extent of Disease</td>
</tr>
<tr>
<td>Enrollment/Registration</td>
</tr>
<tr>
<td>Followup</td>
</tr>
<tr>
<td>Lab Results</td>
</tr>
<tr>
<td>Medical History</td>
</tr>
<tr>
<td>Off Study/Off Treatment</td>
</tr>
<tr>
<td>Physical Examination</td>
</tr>
<tr>
<td>Prior Treatment</td>
</tr>
<tr>
<td>Response</td>
</tr>
<tr>
<td>Study Agent/Intervention</td>
</tr>
</tbody>
</table>
## Critical Mass: Core for Trial Implementation

<table>
<thead>
<tr>
<th>Category</th>
<th>Round</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concomitant Agents</td>
<td>Round 3</td>
</tr>
<tr>
<td>Adverse Events</td>
<td>Round 2</td>
</tr>
<tr>
<td>Course Assessments</td>
<td>Assess need</td>
</tr>
<tr>
<td>Eligibility</td>
<td>Round 4</td>
</tr>
<tr>
<td>Extent of Disease</td>
<td>Round 3</td>
</tr>
<tr>
<td>Enrollment/Registration</td>
<td>Round 2</td>
</tr>
<tr>
<td>Followup</td>
<td>Not Done</td>
</tr>
<tr>
<td>Lab Results</td>
<td>Round 3</td>
</tr>
<tr>
<td>Medical History</td>
<td>Round 2</td>
</tr>
<tr>
<td>Off Study/Off Treatment</td>
<td>Round 4</td>
</tr>
<tr>
<td>Physical Examination</td>
<td>Round 2</td>
</tr>
<tr>
<td>Prior Treatment</td>
<td>Round 3</td>
</tr>
<tr>
<td>Response</td>
<td>Round 3 (Outcomes)</td>
</tr>
<tr>
<td>Study Agent/Intervention</td>
<td>Round 3</td>
</tr>
</tbody>
</table>
Critical Next Steps

Implementation

- When Round 2 content is standardized, finalized CRFs can be built
- When Round 3 content is standardized, 15 modules of content will be final and ready to add to CRFs

Targeted Early Adopters

- Groups that report internally
- Commercial partners
Implementation: Expanding Scope to CTEP

Implementation within CTEP-reported trials

- Create an Implementation plan to move content into the community in a phased approach
- Propose two phases of implementation
  - Rounds 1 through 3
  - Rounds 4 and beyond

- Instruction manual, training
- Publicized rollout timeline and expectation
- Change Management Strategy
Implementation: Costs and Level of Effort

Cost of implementation of new CRF variables will include:

- Rework of CRFs
- Changes in software validation rules
- Training for users (including Instruction Manual)
- Changes in data extraction procedures to report
- Registration of new CRFs as Templates for widespread re-use
Implementation: Early Adopter Actual Costs

CRF: Demography

Change in CRF: About 2 hours, minimal changes

Change in Validation: Minimal

Training: Insignificant due to content

Change in Extraction: Significant

Registration in Library: Minimal in C3D

Other Costs: Need for Builders to learn about new requirements
### Implementation: Early Adopter Actual Costs

<table>
<thead>
<tr>
<th>CRF:</th>
<th>Adverse Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in CRF:</td>
<td>4 hours</td>
</tr>
<tr>
<td>Change in Validation:</td>
<td>Significant</td>
</tr>
<tr>
<td>Training:</td>
<td>Significant</td>
</tr>
<tr>
<td>Change in Extraction:</td>
<td>Very significant – 3 days</td>
</tr>
<tr>
<td>Registration in Library:</td>
<td>Minimal in C3D</td>
</tr>
<tr>
<td>Other Costs:</td>
<td>Need to create new Global Library objects, learn the new requirements</td>
</tr>
<tr>
<td>Summary:</td>
<td>Complex CRFs will take 5 days; less complex will take 2 days</td>
</tr>
</tbody>
</table>
Harmonization with Industry/FDA

- Clinical Data Interchange Standards Consortium (CDISC) initiative called Clinical Data Acquisition Standards Harmonization (CDASH) has a common goal with ours: harmonizing and standardizing data collection for clinical research
  - CDASH focus: elements that are common to all clinical studies
  - caBIG® focus: oncology studies
- caBIG® modules will include all CDASH “mandatory” questions, plus additional oncology content
- Review of content of first forms with CDASH standard: 84% match on first pass
  - All areas of disparity reconciled
caBIG® and CDASH

- Both projects harmonize and standardize the collection of data in clinical trials
  - CDASH focus is all clinical trials
  - caBIG is focused on oncology trials
- NCI/caBIG CRF modules will include all CDASH “mandatory” questions

NCI Elements

CDASH “mandatory”
Return on Investment

How much we have promoted into the caDSR in the first 15 modules in Rounds 1-3?

- 188 Elements in final delivery
- Thousands of Elements assessed (approximately 3500) from CRFs
  Distillation of Elements to <1% of total assessed

How much more will we decompose, vet and add in modules 4 – 5 or 6?

<table>
<thead>
<tr>
<th>Decomposed</th>
<th>Added</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 Contributing Organizations</td>
<td>Projected Estimate: &gt;3500 elements being added to inventory from rounds 1-3</td>
</tr>
<tr>
<td>325 Case Report Forms</td>
<td></td>
</tr>
<tr>
<td>&gt;10,000 Elements in Inventory</td>
<td></td>
</tr>
</tbody>
</table>
Data Management Efficiencies

- How much we will be able to retire as a result of the Harmonization effort?
  
  - Use Demography elements as an example. Reviewing content in caDSR that is semantically similar, an estimated 42 elements can be marked for retirement, thus improving data quality within the caDSR.

- How much existing caDSR content has been reused as a result of the Harmonization of elements?
  
  - Analysis of each standard CRF area shows an average 66% reuse rate for underlying data elements, which drives semantic consistency and significantly decreases analysis cycle time during the stakeholder confirmation process.
Measure Success

1. Track adoption of the new variables
   - Maintain focus and development of core CRF areas
2. Monitor use of CRF templates, training, Instruction manual
3. Track Reuse of CRF Harmonization artifacts
   - Foster repeatability and consistency
4. Retire redundant caDSR content
   - Improve data quality and accessibility
5. Solidify CRF standards and Traceability
6. Capture entry and utilization of new elements
   - Extensible design allows controlled changes
Location of CRF Standards

Search for Data Elements

caDSR Contexts

- Exact phrase
- All of the words
- At least one of the words

Tip: This is an exact match search. To search for partial words or phrases use the "*" as a wild card. Default settings exclude field and naming context values from the tree and contain "X" statuses. Click the "Search Preferences" link above to view or change the exclusion criteria.

Search
Clear

User Public User

Version 4.0.1 Build 4
### Module Status Table

<table>
<thead>
<tr>
<th>Round of Harmonization &amp; Standardization</th>
<th>CRF Modules</th>
<th>Current Status of Module</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Demographics</td>
<td>Harmonization complete; Standards presentation 9/3/09, pending final approval</td>
</tr>
<tr>
<td>2</td>
<td>Adverse Events</td>
<td>Harmonization complete; preparing for standards review</td>
</tr>
<tr>
<td>2</td>
<td>Enrollment</td>
<td>Harmonization complete; preparing for standards review</td>
</tr>
<tr>
<td>2</td>
<td>Medical History</td>
<td>Harmonization complete; preparing for standards review</td>
</tr>
<tr>
<td>2</td>
<td>Participant Identifier</td>
<td>Harmonization complete; preparing for standards review</td>
</tr>
<tr>
<td>3</td>
<td>Physical Basis</td>
<td>Harmonization complete; preparing for standards review</td>
</tr>
<tr>
<td>3</td>
<td>Protocol Deviations</td>
<td>Harmonization complete; preparing for standards review</td>
</tr>
<tr>
<td>3</td>
<td>Registration</td>
<td>Harmonization complete; preparing for standards review</td>
</tr>
<tr>
<td>3</td>
<td>Agents Vocabulary</td>
<td>Harmonization nearing completion</td>
</tr>
<tr>
<td>3</td>
<td>Study Administration</td>
<td>Harmonization nearing completion</td>
</tr>
<tr>
<td>3</td>
<td>Drug Accountability</td>
<td>Harmonization nearing completion</td>
</tr>
<tr>
<td>3</td>
<td>Concomitant Medications</td>
<td>Harmonization nearing completion</td>
</tr>
<tr>
<td>3</td>
<td>Pre &amp; Post Treatment</td>
<td>Harmonization nearing completion</td>
</tr>
<tr>
<td>3</td>
<td>Laboratory Tests &amp; Results</td>
<td>Harmonization nearing completion</td>
</tr>
<tr>
<td>3</td>
<td>Outcome Measures</td>
<td>Harmonization nearing completion</td>
</tr>
<tr>
<td>3</td>
<td>Outcome Measures Glossary</td>
<td>Harmonization nearing completion</td>
</tr>
<tr>
<td>3</td>
<td>Staging &amp; Extent of Disease</td>
<td>Harmonization nearing completion</td>
</tr>
<tr>
<td>3</td>
<td>Disease &amp; Pathology</td>
<td>Harmonization in progress, in workgroup meetings</td>
</tr>
<tr>
<td>3</td>
<td>Eligibility Criteria</td>
<td>Harmonization in progress, in workgroup meetings</td>
</tr>
<tr>
<td>4</td>
<td>Header Information</td>
<td>Harmonization in progress, in workgroup meetings</td>
</tr>
</tbody>
</table>
## Location of CRF Standards

### Demography CRF Module

The Demography module was the only group of variables included in the first round of the CRF project. This was considered a pilot for the standardization and harmonization of CRF variables. The Demography module includes harmonized variables related to demographic information recommended for collection on all NCI-sponsored clinical trials.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Variable Definition</th>
<th>Rule</th>
<th>Conditionality</th>
<th>CRF Public ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>The patient's self-declared racial origin, independent of ethnic origin, using OMB approved categories</td>
<td>Mandatory</td>
<td>N/A</td>
<td>106s:5.1</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>The patient's participant's self-declared ethnic origin, independent of racial origin, based on OMB approved categories</td>
<td>Mandatory</td>
<td>N/A</td>
<td>2002440v:4.2</td>
</tr>
<tr>
<td>Gender</td>
<td>The classification of the sex or gender role of the patient</td>
<td>Mandatory</td>
<td>N/A</td>
<td>62v:6.0</td>
</tr>
<tr>
<td>Patient's Date of Birth</td>
<td>The month, day, and year on which the patient/participant was born</td>
<td>Mandatory</td>
<td>N/A</td>
<td>798s:4.0</td>
</tr>
<tr>
<td>Census Tract Source for 1970/80/90/2000</td>
<td>The origin of the Census Tract set used to identify coded boundaries</td>
<td>Mandatory</td>
<td>N/A</td>
<td>2681456v:1.0</td>
</tr>
<tr>
<td>Census Tract Codes for Years 1970/80/90</td>
<td>The numbers that describe the Census Tract or Block Numbering Area for a location as defined by boundaries identified in 1970, 1980, or 1990</td>
<td>Mandatory</td>
<td>N/A</td>
<td>2681555v:1.0</td>
</tr>
<tr>
<td>2000 Census Tract Code</td>
<td>The number that represents the Census Tract for a location as defined by boundaries identified in the year 2000</td>
<td>Mandatory</td>
<td>N/A</td>
<td>2681552v:1.0</td>
</tr>
<tr>
<td>Country of Residence</td>
<td>The name of the country in which the patient resides</td>
<td>Conditional</td>
<td>N/A</td>
<td>316v:3.0</td>
</tr>
<tr>
<td>Zip Code</td>
<td>The string of characters used to identify the five-digit Zip Improvement Plan (ZIP) code and the four-digit extension code (if available) that represents the geographic segment that is a subset of the ZIP code, assigned by the U.S. Postal Service to a geographic location to facilitate mail delivery; or the postal zone specific to the country, other than the U.S., where the mail is delivered</td>
<td>Conditional</td>
<td>N/A</td>
<td>2178605v:2.0</td>
</tr>
<tr>
<td>Method of Payment</td>
<td>Primary payer/insurance carrier information at the time of treatment on a protocol</td>
<td>Conditional</td>
<td>N/A</td>
<td>2003330v:2.0</td>
</tr>
<tr>
<td>Date Completed</td>
<td>The date on which a data capture form (CRF or case report form) was completed</td>
<td>Conditional</td>
<td>N/A</td>
<td>2003745v:5.0</td>
</tr>
<tr>
<td>CDC Race Code</td>
<td>The unique identifier code for reporting information about race based on the Center for Disease Control and Prevention (CDC) categories</td>
<td>Optional</td>
<td>N/A</td>
<td>2200286v:1.0</td>
</tr>
<tr>
<td>CDC Ethnicity Code</td>
<td>The unique identifier for reporting information about ethnicity based on the Centers for Disease Control and Prevention (CDC) categories</td>
<td>Optional</td>
<td>N/A</td>
<td>2200282v:2.0</td>
</tr>
<tr>
<td>Education Level</td>
<td>A subdivision of the demographic parameter indicating the highest level of education attained by a human being</td>
<td>Optional</td>
<td>N/A</td>
<td>2674076v:1.0</td>
</tr>
<tr>
<td>Educational Attainment</td>
<td>The subdivisions that summarize the highest level of education attained by a person</td>
<td>Optional</td>
<td>N/A</td>
<td>2681556v:1.0</td>
</tr>
<tr>
<td>Marital Status</td>
<td>Text term to describe a self-reported demographic parameter indicating a person's current conjugal status</td>
<td>Optional</td>
<td>N/A</td>
<td>2188005v:2.0</td>
</tr>
<tr>
<td>Place of Birth</td>
<td>The words that describe the geographic location where a person was born</td>
<td>Optional</td>
<td>N/A</td>
<td>2188002v:2.0</td>
</tr>
</tbody>
</table>
Ongoing Maintenance

- Every calendar year the entire library of CRFs will be reviewed for changes
- Older versions will be appropriately archived, yet accessible as required
- caBIG® will provide support for updates and release with version control
- Modifications will be requested and changes will be vetted by community and NCI
- All CRF versions will be available on NCI website
- The first CRF module (Demography) will become a candidate for its first review in October 2010
In Conclusion

• Multiple benefits drive the harmonization and standardization of data collection

• NCI and CDISC have synergized complementary efforts

• Results are strengthened through consensus building and community approval
caBIG® and CDASH

CDASH basic data collection elements +
caBIG Cancer specific data collection elements =
a standardized CRF for the oncology workspace that bridges the research and Pharma worlds

A Winning Solution for our Patients!