CDISC SDTM and ADaM Real World Issues

Washington DC CDISC Data Standards User Group Meeting

Sy Truong
President
MXI, Meta-Xceed, Inc.
http://www.meta-x.com
Agenda

• CDISC SDTM and ADaM Fundamentals
• CDISC Real World Implementation Examples, 7 Steps
• Lessons Learned
MXI Profile

• Clinical Analytical Software & Services for Pharmaceutical and Biotech since 1997
• CDISC Solution Provider
• MXI Specializes in:
  - CFR Part 11 Data Standards and CDISC
  - Electronic Submission / FDA Compliance
  - Validation of Statistical Programming and Biostatistics Systems

• Headquarters: Milpitas California US
• Software Development Center: Vietnam
Using CDISC Standard Saves

The Standards Advantage

**Figure 1.** Cycle time savings with CDISC standards.

**Source:** CDISC and Gartner.

**Analysis/Reporting**

**Study Conduct**

**Study Startup**

---

**Industry Average**

**CDISC Standards Impact**

8 Months**

Cycle Time Reduction/Trial

---

**Cycle Time in Months**

---

*Industry averages from PAREXEL’s *Source Pharmaceutical R&D Statistical Sourcebook 2004/2005* for benchmark studies. These figures are based upon average benchmark studies. Estimates for specific individual study saving will require calculations based upon estimated percentage cost and time savings from the Business Case.

**Source:** CDISC and Gartner.
# Time and Resource Savings

<table>
<thead>
<tr>
<th>Current Industry Average (Months)</th>
<th>Study Startup</th>
<th>Study Conduct</th>
<th>Analysis and Reporting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Savings Reported From Using CDISC Standards (%)</td>
<td>5</td>
<td>4*</td>
<td>5</td>
</tr>
<tr>
<td>Revised Time Estimate When Using CDISC Standards (Months)</td>
<td>80</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>Time Savings (Months)</td>
<td>1</td>
<td>2.4</td>
<td>2.5</td>
</tr>
<tr>
<td>4</td>
<td>1.6</td>
<td>2.5</td>
<td></td>
</tr>
</tbody>
</table>

*Does not include subject participation.

Industry averages from PAREXEL’s *Pharmaceutical R&D Statistical Sourcebook 2004/2005*.

Note: These figures are based upon benchmark average studies. Specific individual study savings will require calculations for time and cost savings based upon estimated percentage cost and time savings from the Business Case.

**Source:** CDISC and Gartner.

**Table 1.** Time and cost savings realized when CDISC standards were implemented.
SDTM 3.1.1 Tabulations Dataset

- **General Use** - Standardize clinical trial tabulation datasets for FDA submissions

- **Tabulations Datasets**
  - Observations of data collected in clinical trials organized into domains
  - One of four data types submitted to FDA also including: Patient Profiles, Listings and Analysis Files.
SDTM Standard Guidelines

- **Data Name** - Dataset name are a short two letter name corresponding to domain, (AE for Adverse Events)
- **Variable Attributes**
  - Variable Name, 8 characters
  - Label, 40 characters
  - Type, character or numeric
  - Controlled Terminology or Format
  - Origin, source of each variable
  - Role - how the variable will be used
  - Comments - relevant text information about variable
SDTM Control Terminology (CT)

- **Upper Case** - Text upper case unless referencing external terms (i.e. MedDRA LOINC) or Units (i.e. mg/dL)
- **Use CT** Rather than Numeric Codes
- **CDISC SDTM** Recommends Codes

<table>
<thead>
<tr>
<th>Keyword(s)</th>
<th>Fragment</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTION</td>
<td>ACN</td>
</tr>
<tr>
<td>ADJUSTMENT</td>
<td>ADJ</td>
</tr>
<tr>
<td>ANALYSIS DATASET</td>
<td>AD</td>
</tr>
<tr>
<td>BASELINE</td>
<td>BL</td>
</tr>
<tr>
<td>BIRTH</td>
<td>BRTH</td>
</tr>
<tr>
<td>BODY</td>
<td>BOD</td>
</tr>
<tr>
<td>CANCER</td>
<td>CAN</td>
</tr>
<tr>
<td>CATEGORY</td>
<td>CAT</td>
</tr>
<tr>
<td>CHARACTER</td>
<td>C</td>
</tr>
<tr>
<td>CONDITION</td>
<td>CND</td>
</tr>
<tr>
<td>CLASS</td>
<td>CLAS</td>
</tr>
<tr>
<td>CODE</td>
<td>CD</td>
</tr>
</tbody>
</table>
ADaM 2.0 Overview

• **Purpose** - Used to standardize analysis datasets for statistical results in FDA submission

• **Data Types** - Three kinds of analysis metadata: (analysis dataset metadata, analysis variable metadata, and analysis results metadata)

• **Add Standard** - Built on the existing standard of SDTM 3.1.1
ADaM 2.0 Principles

• Analysis Datasets - Datasets used for statistical analysis and reporting
• Principles - datasets can be analyzed with little programming or manipulations, “analysis ready”, “one proc away”
ADaM 2.0 Data Flow
ADaM Guidelines

• **Name** - starts with AD such as ADSL. ADxxxx where xxxx is sponsor defined.
• **Structure** - Optimize for analysis with little programming
• **Attributes** - maintain SDTM attributes if matches with SDTM
• **Variables** - Follow SDTM or sponsor defined but be consistent
# ADaM Date Imputation

<table>
<thead>
<tr>
<th>Missing Elements</th>
<th>ADaM Date Variables</th>
<th>SDTM V3.1.1 Date-Time Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>YYYY-MM-DD</td>
<td>blank</td>
</tr>
<tr>
<td>Day</td>
<td>YYYY-MM-**</td>
<td>D</td>
</tr>
<tr>
<td>Month (and Day)</td>
<td>YYYY-<strong>-(</strong>)</td>
<td>M</td>
</tr>
<tr>
<td>Year (and M, D)</td>
<td>****-(<strong>)-(</strong>)</td>
<td>Y</td>
</tr>
</tbody>
</table>
## ADaM Numeric Coded Terminology

### Example of Analysis Variable Metadata for Numeric Code Variables

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Variable Label</th>
<th>Type</th>
<th>Controlled Terms or Format[^1]</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>AESEV</td>
<td>Adverse Event Severity</td>
<td>Char</td>
<td>Mild, Moderate, Severe, Life-threatening</td>
<td>AE.AESEV</td>
</tr>
<tr>
<td>AESEVN</td>
<td>Adverse Event Severity Numeric Code</td>
<td>Num</td>
<td>1=Mild, 2=Moderate, 3=Severe, 4=Life-threatening</td>
<td>Coded from AE.AESEV</td>
</tr>
</tbody>
</table>

[^1]: Indicates the numeric codes and their corresponding meanings.
CDISC Implementation

- **SDTM** - Study Data Tabulation Model
  version 3.1.1
- **ODM** - Operational Data Model
  Final version 1.2.1 with v1.3 Draft
- **ADaM** - Analysis Data Model
  Version 2.0
- **LAB** - Laboratory Data Model
  version 1.0.1
Implementation Flow

- **CDISC Models for Considerations:** SDTM, ODM, LAB, ADaM
- **Evolving Standard with SDTM 3.1.1 Becoming Prime Time**
MXI 7 Steps Methodologies

- **Step 1:** CDISC Training
- **Step 2:** Implementation Planning - Template SOPs and Domain Knowledge
- **Step 3:** Data Standards Analysis
- **Step 4:** Data Transformation
- **Step 5:** Validation and Testing
- **Step 6:** Analysis and Reporting
- **Step 7:** ESUB Preparations and Documentation
Automate Methodologies

Step 1
- CDISC Training – Traingen™

Step 2
- Implementation Planning – Template SOPs and Domain Knowledge

Step 3
- Data Standards Analysis – CDISC Builder™, Sy/Data, ThesQA™

Step 4
- Data Transformation – Transdata™, CDISC Builder™

Step 5
- Validation – SyValidate™, CDISC Builder™, ThesQA™

Step 6
- Reporting – CDISC Reports™, TrialEx System™, TrialEx Plugin™, SyMap™

Step 7
- ESUB – Definedoc™
Step 1: CDISC Training

- Challenges:
- Lack of CDISC Expertise or Education

Solutions:
Traingen™ - Interactive Training with Performance Metrics
- Creates Quiz to Measure Metrics
- Multi Media Training
Lessons Learned

• Initial Education Invested Results In Accurate and Efficient Production
• Work with Global or Holistic Vision of CDISC Implementation
• Team Works Cohesively with Common Goals
Step 2: Implementation Planning

- Clear Plan and Vision Prior to Implementation
- Obtain Proper Resource and Time
- Get Management and Support from Team
PROJECT MANAGEMENT

• Project Management Steps for Entire Process
  1. Define Scope
  2. Identify Tasks
  3. Project Plan
  4. Validation Plan
  5. Transformation Specification
  6. Applying Transformation
  7. Verification Reports
  8. Handling Special Purpose Domains
  9. Managing Sequence, Order and Lengths
PROJECT MANAGEMENT OBJECTIVES

• Establish Plan to Galvanize Organization in Support
• Provide Clear Understanding
• Avoid Political Pitfalls between Interdepartmental Territories
DEFINE SCOPE

- **Scope is a Document Similar to Requirements Documentation**
- **Short one or two Paragraphs to Cap Project Creep**
- **Identify Initial Pilot versus Production Roll Out**
DEFINE SCOPE

- Identify Audience and Groups of Users for Data Standards including Internal and External Regulatory
- Define Formality of Validation
- Identify Level of Documentation
- Establish Standards Usage for Future Projects
DEFINE SCOPE EXCERCISE

• Type a Word Document (*scope.doc*)

Quantifying Scope of:

- Pilot
- Roll Out
- Target Audiences
- Level of Validation
- Documents
- Future Standards
PROJECT SCOPE USAGE

• Can be Reused in Formal Requirement Documentation
• Communication Tool to Manage Levels of Expectations
IDENTIFYING TASKS

- Identify All Tasks in Implementation for CDISC within Scope
- Initial Plan will be Estimated
- Tasks Modified During and Conclusion of Project
- Estimate Time and Resources for Tasks
IDENTIFYING TASKS EXERCISE

- Type Projected List of Identified Tasks (identify_tasks.doc)
- Identify Your Own Projected Tasks
- Update Units to Reflect Number of Hours
- Assign Identified Team Members by Roles
IDENTIFYING TASKS SUGGESTIONS

• Detail Tasks Enough to Clearly Communicate Expectations for Assigned Team Members or Groups
• Use as Communication Mechanism for Team to Allocate Project Resource and Time Commitments
**PROJECT PLAN**

• **Formalize Previous Documents into Plan**

• **Project Tasks are further Detailed from Identified Tasks**

### Study ABC1234 CDISC Transformation Project Plan

#### Overview
This project plan will detail some of the tasks involved in transforming the source data of study ABC1234 into CDISC SDTM in preparation for electronic submission. The proposed time lines are intended as goals which can be adjusted to reflect project priorities.

#### Project Tasks
The following tasks are organized into groups of tasks which have some dependency. They are therefore organized in chronological order.

1. **Data Review**
   1. Evaluate variable attributes differences within internal data of ABC1234
   2. Evaluate variable attributes between ABC1234 as compared to ACME Standards
   3. Evaluate ABC1234 differences and similarities with CDISC SDS v3.1
   4. Evaluate potential matches of ABC1234 variable names and labels against CDISC SDS v3.1
   5. Initial evaluation of ABC1234 against CDISC evaluation
   6. Generate metadata documentation of the original source data from ABC1234

2. **Data Transformation Specifications**
   1. Perform a thorough review of all data and associated attributes against CDISC SDS v3.1. Identify all recommended transformation requirements. This is documented in a transformation requirement specification.
   2. Create transformation models based on the transformation specifications for each data domain.
   3. Have transformation reviewed for feedback.
   4. Update the specification to reflect feedback from review.

#### Task Assignments

<table>
<thead>
<tr>
<th>Project Tasks</th>
<th>Project Manager</th>
<th>Team Managers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Review</td>
<td>James Brown, Director of Data Management</td>
<td>James Brown, Billy Joel, Joe Jackson</td>
</tr>
</tbody>
</table>

September 13, 2007
PROJECT SCHEDULE

• Schedule Identified Tasks

<table>
<thead>
<tr>
<th>Schedule of Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>August 2005</td>
</tr>
<tr>
<td>Sun</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>Data Review</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>21</td>
</tr>
<tr>
<td>Data Transformation Specifications</td>
</tr>
<tr>
<td>28</td>
</tr>
</tbody>
</table>

September 13, 2007
Step 3: Data Standards Analysis

• Evaluate Legacy Data For Standards Deviations Prior to Transformation
• Review Data for CDISC Compliance
Lessons Learned

- Pre-Transformation Analysis Provides Accurate Project Scoping
- Resourcing and Project Creep Leads to Project Failure
- Understanding Data Prior to Transformation Leads to Efficiencies
Step 4: Data Transformation

- Define Data Transformation Specification Prior to Implementation
- Develop Transformation According to Specification
TRANSFORMATION SPECIFICATION

- Specification Document Details of Transformation Attributes
  - Dataset Name
  - Dataset Label
  - Variable Name
  - Variable Label
  - Variable Type
  - Variable Length
  - Format
TRANSFORMATION TYPE

• Specify Common Transformation Type
  - Yes/No
  - Vertical
  - Combine
  - Drop
  - Same
  - Value Change
TRANSFORMATION SPECIFICATION EXCERCISE

- Edit Spreadsheet (transformations_spec.xls)
- Cut and Paste Attributes from SAS Viewer
- Type a Sample Transformation
TRANSFORMATION SPECIFICATION

• Transformation Specification is a Dynamic Process with Multiple Interactive Reviews
• Plan Extra Time for Programming Updates Since it is Resource Intensive and Time Consuming
• Automation Tools such as Transdata Make Iterative Process More Efficient
PROGRAM TRANSFORMATION

• Transformation Specification is the Functional Specification for Programming

• Data Transformation and Merges

```sql

* Program: trans_ss.sas
* Path: c:\temp
* Description: Transform Adverse Events data
* From DATAMAP.AR to STDMIB.AE
* By: By Truong, 04/21/2006, 3:49:33 pm

set DATAMAP.AR;
retain cke 1;
*** Define new variable: serelinet that combined by old variables: adconatt adsoagtt;
attrib serelinet label="Relationship to Non-Study Treatment" length=110;

serelinet = trim(trim(adconatt) || '' || adsoagtt);

*** Define new variable: adscout that combined by old variables: addes1 addes2 addes3 addes4 addes5
addes6 addes7 addes8 addes9 addes10;
attrib adscout label="Outcome of Adverse Event" length=1000;

adscout = trim(trim(trim(trim(trim(trim(trim(trim(trim(trim(trim(adconatt) || addes1)) || addes2)) || addes3)) || addes4)) || addes5)) || addes6)) || addes7)) || addes8)) || addes9)) || addes10));

drop adconatt adsoagtt adscout adeconnett;
run;
```
SPECIAL PURPOSE DOMAIN

• Three Special Purpose Domains: SUPPQUAL, RELREC and CO
• SUPPQUAL - Nonstandard variables associated with Parent Domain
• RELREC - Related Records used to Describe Relationships in Event, Intervention or Findings
• CO - Comment Text Pertaining to Specified Domains
### SPECIAL PURPOSE DOMAIN

#### SUPPLEMENTAL QUALIFIER

<table>
<thead>
<tr>
<th>Obs</th>
<th>Study Identifier</th>
<th>Related Domain Abbreviation</th>
<th>Unique Subject Identifier</th>
<th>Identifier Variable</th>
<th>Identifier Variable Value</th>
<th>Qualifier Variable Name</th>
<th>Qualifier Variable Label</th>
<th>Data Value</th>
<th>Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23423</td>
<td>AE</td>
<td>1803</td>
<td>sequum</td>
<td>64001</td>
<td>ae</td>
<td>ADVERSE EXPERIENCE</td>
<td></td>
<td>derived</td>
</tr>
<tr>
<td>2</td>
<td>23423</td>
<td>AE</td>
<td>101</td>
<td>sequum</td>
<td>64002</td>
<td>ae</td>
<td>ADVERSE EXPERIENCE</td>
<td>(R) PLEURAL EFFUSION</td>
<td>derived</td>
</tr>
<tr>
<td>3</td>
<td>23423</td>
<td>AE</td>
<td>1102</td>
<td>sequum</td>
<td>64003</td>
<td>ae</td>
<td>ADVERSE EXPERIENCE</td>
<td>(SOB) SHORTNESS OF BREATH</td>
<td>derived</td>
</tr>
<tr>
<td>4</td>
<td>23423</td>
<td>AE</td>
<td>1103</td>
<td>sequum</td>
<td>64001</td>
<td>ae</td>
<td>ADVERSE EXPERIENCE</td>
<td>ABDOMINAL CRAMPING</td>
<td>derived</td>
</tr>
<tr>
<td>5</td>
<td>23423</td>
<td>AE</td>
<td>102</td>
<td>sequum</td>
<td>64008</td>
<td>ae</td>
<td>ADVERSE EXPERIENCE</td>
<td>ABDOMINAL DISTENTION</td>
<td>derived</td>
</tr>
<tr>
<td>6</td>
<td>23423</td>
<td>AE</td>
<td>101</td>
<td>sequum</td>
<td>64004</td>
<td>ae</td>
<td>ADVERSE EXPERIENCE</td>
<td>ABDOMINAL PARODONTAL</td>
<td>derived</td>
</tr>
</tbody>
</table>
SUPPLEMENTAL QUALIFIERS

- Vertical Structure of Data Allows for Storage Flexibility
- Data is not Normally Stored this Way so Transformation is Commonly Required

```plaintext
*** Transpose the for variable TMTCCV ***;
data work.supqual;
set sourcelib.a.e;
*** Convert numeric idVar ***;
idvar="sexbdt";
idvarval = left|trim(putf(idvarval,8F8.0));
qual = tcropy;
qual = left|trim(put quali, best.111);
subiden = putnum;
domain = "Ad";
qual = "otherse";
qlabel = "Other Specify Adverse Event";
run;
*** Append data to the final destination SUPPQUAL ***;
data SDTM13.supqual [label = "Supplemental Qualifiers for Adverse Events"]; set SDTM13.suppqual_a;
run;
if comprss(studyid) = 1 and comprss(subjectid) = 1 then delete;
run;
```
SEQUENCE ORDER AND LENGTH

- **Sequence**: More than one Observation per Subject Requires Proper Sequence
- **Order**: The Variable Appears in a Standard Order, Keys First Followed By Alphabetical
- **Lengths**: Not Strictly Specified by Guidelines but Optimized for Efficiency
Lessons Learned

- Data Transformation Specification Iterative and Time Consuming Process
- Special Purpose Domains such as SUPPQUAL Requires Transformation Programming and Validation Effort
- Maintain Standards through Non Strict Guidelines including Order and Lengths
Step 5: Validation and Testing

- Level of Formality and Documentation Dictated by Scope
- Risk Assessment
- Test Plan
- Summarize Result Reports
VALIDATION EXERCISE

• Perform Risk Assessment by updating
  (risk_assessment.doc)
  - Update the information to reflect performing a rename of AE variables to CDISC Standard
  - Perform a test and Summarize Score
VALIDATION OBJECTIVES

- Risk Assessment Drives the Level of Validation
- Scope Drives the Formality of the Validation Documentation
VALIDATION DEFINITIONS

- Define Testing Approach In Test Plan
- Define Required Steps for the Performance of Tests
- Define Clear Objectives of Success and Failure of Tests
VALIDATION TEST SCRIPTS

• Test Scripts Applies to Each Test Case in Test Plan
• Tester is a Different Person from Person Performing Tasks
VERIFICATION REPORTS

• Execute Test Scripts as Part of Test Plan
• Ensures Integrity and Accuracy of All Transformations
### Duplicate Variable Reports

**Duplicate Variables and Attributes**

<table>
<thead>
<tr>
<th>Obs</th>
<th>Source Dataset Name</th>
<th>Variable</th>
<th>Update To</th>
<th>Transformed Variable</th>
<th>Destination Table Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AE</td>
<td>ADCONATT</td>
<td>aerelinst label=&quot;Relationship to Non-Study Treatment&quot; length=$140</td>
<td>AerelInst</td>
<td>STDMLIB.AE</td>
</tr>
<tr>
<td>2</td>
<td>AE</td>
<td>ADOAGATT</td>
<td>aerelinst label=&quot;Relationship to Non-Study Treatment&quot; length=$140</td>
<td>AerelInst</td>
<td>STDMLIB.AE</td>
</tr>
</tbody>
</table>

Generated on: 01/21/2006, 4:27:24 pm, Sy Truong
Model located at: C:\CDISC\DATA\MODELS
Report located at: C:sasy8\
FREQUENCY REPORTS

- Subset Data for Visual Inspection
- PROC FREQ Checking
Lessons Learned

- Do Appropriate Level of Validation based on Risk Assessment
- Create Consistent Repeatable or Duplicate Purpose Verification Program when Possible
Step 6: Analysis and Reporting

• Perform Analysis using ADaM and SDTM datasets
• Analysis and Reporting as a Form of Validation
Step 7: ESUB Preparations and Documentation

- Document Submission Datasets
- Prepare DEFINE.PDF, DEFINE.XML
- Update Annotated CRF with CDISC Data
DATA DEFINITION DOCUMENTATION

- Roadmap to Your Derived and Source Data Documenting Three Levels:
  - General Information
  - Data Table
  - Variable
## DATA DEFINITION DOCUMENTATION

### GENERAL INFORMATION

<table>
<thead>
<tr>
<th>Metadata</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company Name</td>
<td>This is the name of the organization that is submitting the data to the FDA.</td>
</tr>
<tr>
<td>Product Name</td>
<td>The name of the drug that is being submitted.</td>
</tr>
<tr>
<td>Protocol</td>
<td>The name of the study on which the analysis is being performed which includes this set of data.</td>
</tr>
<tr>
<td>Layout</td>
<td>The company name, product name, and protocol are all going to be displayed on the final documentation. The layout information will describe if it will be in the footnote or title and how it is aligned.</td>
</tr>
</tbody>
</table>
## DATA DEFINITION DOCUMENTATION

### DATA TABLE INFORMATION

<table>
<thead>
<tr>
<th>Metadata</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Library</td>
<td>Library name defines what physical path on which server and where the data is located. This can also be in the form of a SAS LIBNAME.</td>
</tr>
<tr>
<td>Key Fields</td>
<td>Keys usually correlate to the sort order of the data. These variables are usually used to merge the datasets together.</td>
</tr>
<tr>
<td>Format Library</td>
<td>This is where the SAS format catalog is stored.</td>
</tr>
<tr>
<td>Dataset Name</td>
<td>The name of the SAS dataset that is being captured.</td>
</tr>
<tr>
<td>Number of Variables</td>
<td>A count of the number of variables for each dataset.</td>
</tr>
<tr>
<td>Number of Records</td>
<td>Number of observations or rows within each dataset.</td>
</tr>
<tr>
<td>Dataset Comment</td>
<td>A descriptive text describing the dataset. This can contain the dataset label and other descriptive text explaining the data.</td>
</tr>
</tbody>
</table>
## DATA DEFINITION DOCUMENTATION
### VARIABLE INFORMATION

<table>
<thead>
<tr>
<th>Metadata</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable Name</td>
<td>The name of the SAS variable.</td>
</tr>
<tr>
<td>Type</td>
<td>The variable type which includes values such as Character or Numeric.</td>
</tr>
<tr>
<td>Length</td>
<td>The variable length.</td>
</tr>
<tr>
<td>Label</td>
<td>The descriptive label of the variable.</td>
</tr>
<tr>
<td>Format</td>
<td>SAS formats used. If it is a user defined format, it would need to be decoded.</td>
</tr>
<tr>
<td>Origins</td>
<td>The document where the variable came from. Sample values include: Source or Derived.</td>
</tr>
<tr>
<td>Role</td>
<td>This defines what type of role the variable is being used for. Example values include: Key, Ad Hoc, Primary Safety, Secondary Efficacy</td>
</tr>
<tr>
<td>Comment</td>
<td>This is a descriptive text explaining the meaning of the variable or how it was derived.</td>
</tr>
</tbody>
</table>
DATA DEFINITION DOCUMENTATION PROCESS

- Required for Regulatory Reviewers but Also Useful Tool Internally Review
- Start This at the Beginning as a Documentation and Planning Tool
- Iterative Process so Automation is Useful
Lessons Learned

- Document throughout process and not at end leads to accurate and efficient process
- Thorough review among different authors leads to more accurate documentation
- Produce Documentation in Different Format from Centralize Source for consistency and broadens audience
CONCLUSION

- **CDISC Implementation can Save Significant Time and Cost**
- **Lessons Learned**
  - Training has team working on same page with greater efficiencies
  - Pre-Transformation Analysis prevents project creep, accurate scoping
  - Allocate Resource to transformation specification and special domains
  - Replicate validation efforts where possible to save resource
  - Document throughout the process provides insights and results in more accurate documentation

- **Avoid Mistakes and Pitfalls by Following Recommended Steps and Methodologies**
QUESTIONS

Sy Truong
MXI Meta-Xceed Inc.
1751 McCarthy Blvd.
Milpitas, CA 95035
(408) 955-9333
sy.truong@meta-x.com