Automation of SDTM Generation & Artifacts using CDISC 360 enriched standards

Bhavin Busa (Vita Data Sciences), Jianhui [Jimmy] Zhao (Allergan)
CDISC 360: The Journey So Far and the Road Ahead
April 28, 2020
Agenda

1. Workstream 6 Introduction
2. Current State: CDASH to SDTM Execution
3. CDISC 360 Enriched Metadata
4. Future State with Concept-based Standards: CDASH to SDTM Execution
5. Process Flow for CDISC 360 Proof of Concept
7. Learnings so far
Workstream 6 Introduction
CDISC 360 Workstreams

**Workstream 1 - ENHANCE STANDARDS**
Create concepts in knowledge graphs

**Workstream 2 - PUBLISH STANDARDS**
Load into library

Biomedical Concepts
Analysis Concepts
Foundational Standards

API
Extend API's

Transform concepts in machine readable form

**Workstream 3 - STUDY Metadata**

**Workstream 4 - DEFINE**
Identify and select standards specification (Use Case 1)

**Workstream 5 - BUILD**
Configure study specification and create artifacts (Use Case 2)

**Workstream 6 - EXECUTE**
Automatically process and transform data (Use Case 3)
Use Case 3 (Workstream 6): **Execute**

Automatic population of data into artifacts
Workstream 6 & Task Team Leads

**Workstream 6 Lead**

Bhavin Busa, Vita Data Sciences

**SDTM/ADaM Automation Task Team Leads**

Kaja Najumudeen, TalentMine
Jianhui Zhao, Allergan

**TFL Automation Task Team Leads**

Prasanna Murugesan, AstraZeneca
Stuart Malcolm, Frontier Science
Current State - without Concept-based Standards
Current State - without Concept-based Standards: CDASH to SDTM Execution

* Manual Process
+ Manual or Semi-automated Execution
CDISC 360 Enriched Metadata
Machine-readable CDISC 360 Enriched Metadata

Structural

Conceptual

Process

Semantic
Concept-based Standards: Biomedical Concept

- Triple Store
- Linking controlled terminology to the variable - standardize value level metadata
- Linked derivations and algorithms to variable(s)
- Include process metadata (ETL instructions)
- Machine readable definition of validation rules

Reference: ‘CDISC 360 - The Journey so Far and the Road Ahead’, Peter Van Reusel, 28th April 2020
Linked Graph Model: Importing Concept-based Standards

Study Build of ODM.XML and Define.XML

ODM CRF Generated using Biomedical Concepts, Bindings, & Standards

ODM-based Vital Signs (VS) CRF

Stylesheet rendering of ODM VS CRF

Reference: 'CDISC Library: Integrating and Surfacing 360 Content', Sam Hume, October 16, 2019
Future State - with Concept-based Standards
Future State - with Concept-based Standards: CDASH to SDTM Execution

- ODM CRF
- Define XML
- SDTM Specs & aCRF
- SDTM Programs
- SDTM Datasets
- Submission XPT files, Define, aCRF, SDRG

Specify | Build/Execute | Report

= Automated Process
Process Flow for CDISC 360 Proof of Concept (PoC)
Process Flow for CDISC 360 PoC
Machine-readable Mapping Specifications
Essential Elements for Machine-readable Mapping Specifications

We break down the essential elements in 2 dimensions to meet the 4 key aspects of the machine readability

Dimension 1

- **Source**: location (library name), datasets, processing sequence
- **Mapping**: fields needed to describe how source transits to target
- **Target**: location (library name), datasets, processing sequence, attributes (label, class, structure, purpose, etc.)

Dimension 2

- **Dataset Level**: Transit datasets from source to target
- **Variable Level**: Map variables from source to target
- **Value Level**: Map variables from source to target under different conditions
### Mapping Specifications: Dimension 1

<table>
<thead>
<tr>
<th>Source Sequence</th>
<th>Source Library</th>
<th>Source Dataset</th>
<th>Source Variable</th>
<th>Map Sequence</th>
<th>Origin</th>
<th>Method</th>
<th>Comment</th>
<th>Code List</th>
<th>Target Sequence</th>
<th>Target Library</th>
<th>Target Dataset</th>
<th>Target Variable</th>
<th>Target Description</th>
<th>Target Data Type</th>
<th>Target Length</th>
<th>Target Sorting Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CDASH</td>
<td>VS</td>
<td>CDISC380-2</td>
<td>Source</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mapping</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>CDASH</td>
<td>VS</td>
<td>CDISC380-2</td>
<td>Target</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>CDASH</td>
<td>VS</td>
<td>CDISC380-2</td>
<td>Code List</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Source:**
- CDASH
- VS

**Mapping:**
- CDISC380-2

**Target:**
- Source: CDASH
- Library: VS
- Dataset: CDISC380-2
- Variable: Code List
- Origin: Source
- Method: Map
- Comment: Sequence
- Target Variable: Target
- Description: Study Identifier
- Data Type: text
- Length: 10
- Sorting Order: 1
Mapping Specifications: Dimension 2

<table>
<thead>
<tr>
<th>Source Sequence</th>
<th>Source Library</th>
<th>Source Dataset</th>
<th>Subset Condition</th>
<th>Pre Processing</th>
<th>Join Type</th>
<th>Join Timing</th>
<th>Merge Key</th>
<th>Target Sequence</th>
<th>Target Library</th>
<th>Target Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CDASH</td>
<td>VS</td>
<td></td>
<td></td>
<td>USUBJID</td>
<td>PRE</td>
<td></td>
<td>5</td>
<td>SDTM</td>
<td>VS</td>
</tr>
<tr>
<td>2</td>
<td>SDTM</td>
<td>DM</td>
<td></td>
<td>TARGET</td>
<td>PRE</td>
<td>USUBJID, VISITNUM</td>
<td></td>
<td>5</td>
<td>SDTM</td>
<td>VS</td>
</tr>
<tr>
<td>3</td>
<td>SDTM</td>
<td>SV</td>
<td></td>
<td>SORT</td>
<td>PRE</td>
<td>USUBJID, VISITNUM, VSDTC</td>
<td>5</td>
<td>SDTM</td>
<td>VS</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>SORTE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Dataset Level**

<table>
<thead>
<tr>
<th>Source Sequence</th>
<th>Source Library</th>
<th>Source Dataset</th>
<th>Source Variable</th>
<th>Map Sequence</th>
<th>Origin</th>
<th>Method</th>
<th>Comment</th>
<th>Code list</th>
<th>Target Library</th>
<th>Target Dataset</th>
<th>Target Variable</th>
<th>Target Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Assigned</td>
<td>VS</td>
<td>VSTTESTCD</td>
<td>VSTTESTCD</td>
<td>VS</td>
<td></td>
<td></td>
<td></td>
<td>SDTM</td>
<td>VS</td>
<td>VSTTESTCD</td>
<td>Vital Signs Test Short Name</td>
</tr>
<tr>
<td>4</td>
<td>Derived</td>
<td>VS</td>
<td>VSORRES</td>
<td>VSORRES</td>
<td>SOR</td>
<td></td>
<td></td>
<td></td>
<td>SORRES</td>
<td>SORRES</td>
<td>SORRES</td>
<td>Result or Finding in Original Units</td>
</tr>
<tr>
<td>4</td>
<td>Derived</td>
<td>VS</td>
<td>VSORRES</td>
<td>VSORRES</td>
<td>SOR</td>
<td></td>
<td></td>
<td></td>
<td>SORRES</td>
<td>SORRES</td>
<td>SORRES</td>
<td>Original Units</td>
</tr>
<tr>
<td>4</td>
<td>Derived</td>
<td>VS</td>
<td>VISITNUM</td>
<td>VISITNUM</td>
<td>VISITNUM</td>
<td></td>
<td></td>
<td></td>
<td>SORRES</td>
<td>SORRES</td>
<td>SORRES</td>
<td>Standard Units</td>
</tr>
<tr>
<td>4</td>
<td>Derived</td>
<td>VS</td>
<td>VSTTESTCD</td>
<td>VSTTESTCD</td>
<td>VS</td>
<td></td>
<td></td>
<td></td>
<td>SDTM</td>
<td>VS</td>
<td>VSTTESTCD</td>
<td>Numeric Result/Finding in Standard Units</td>
</tr>
<tr>
<td>4</td>
<td>Derived</td>
<td>VS</td>
<td>VSTTESTRES</td>
<td>VSTTESTRES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SDTM</td>
<td>VS</td>
<td>VSTTESTRES</td>
<td>Character Result/Finding in Std Format</td>
</tr>
</tbody>
</table>

**Variable Level**

<table>
<thead>
<tr>
<th>Source Sequence</th>
<th>Source Library</th>
<th>Source Dataset</th>
<th>Source Variable</th>
<th>Map Sequence</th>
<th>Origin</th>
<th>Method</th>
<th>Comment</th>
<th>Code list</th>
<th>Target Library</th>
<th>Target Dataset</th>
<th>Target Variable</th>
<th>Target Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Assigned</td>
<td>VS</td>
<td>VSTTESTCD</td>
<td>VSTTESTCD</td>
<td>VS</td>
<td></td>
<td></td>
<td></td>
<td>SDTM</td>
<td>VS</td>
<td>VSTTESTCD</td>
<td>Vital Signs Test Short Name</td>
</tr>
<tr>
<td>4</td>
<td>Derived</td>
<td>VS</td>
<td>VSORRES</td>
<td>VSORRES</td>
<td>SOR</td>
<td></td>
<td></td>
<td></td>
<td>SORRES</td>
<td>SORRES</td>
<td>SORRES</td>
<td>Result or Finding in Original Units</td>
</tr>
<tr>
<td>4</td>
<td>Derived</td>
<td>VS</td>
<td>VSORRES</td>
<td>VSORRES</td>
<td>SOR</td>
<td></td>
<td></td>
<td></td>
<td>SORRES</td>
<td>SORRES</td>
<td>SORRES</td>
<td>Original Units</td>
</tr>
<tr>
<td>4</td>
<td>Derived</td>
<td>VS</td>
<td>VISITNUM</td>
<td>VISITNUM</td>
<td>VISITNUM</td>
<td></td>
<td></td>
<td></td>
<td>SORRES</td>
<td>SORRES</td>
<td>SORRES</td>
<td>Standard Units</td>
</tr>
<tr>
<td>4</td>
<td>Derived</td>
<td>VS</td>
<td>VSTTESTCD</td>
<td>VSTTESTCD</td>
<td>VS</td>
<td></td>
<td></td>
<td></td>
<td>SDTM</td>
<td>VS</td>
<td>VSTTESTCD</td>
<td>Numeric Result/Finding in Standard Units</td>
</tr>
<tr>
<td>4</td>
<td>Derived</td>
<td>VS</td>
<td>VSTTESTRES</td>
<td>VSTTESTRES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SDTM</td>
<td>VS</td>
<td>VSTTESTRES</td>
<td>Character Result/Finding in Std Format</td>
</tr>
</tbody>
</table>

**Value Level**

<table>
<thead>
<tr>
<th>Source Sequence</th>
<th>Source Library</th>
<th>Source Dataset</th>
<th>Source Variable</th>
<th>Where Class</th>
<th>Condition</th>
<th>Output</th>
<th>Map Sequence</th>
<th>Origin</th>
<th>Method</th>
<th>Comment</th>
<th>Code list</th>
<th>Target Library</th>
<th>Target Dataset</th>
<th>Target Variable</th>
<th>Target Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>CBASH</td>
<td>VS</td>
<td>VSTTESTCD.EL.SID</td>
<td></td>
<td>VSTTESTCD.EL.SID</td>
<td>Y</td>
<td>7</td>
<td>Convert</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SORRES</td>
<td>SORRES</td>
<td>SORRES</td>
</tr>
<tr>
<td>3</td>
<td>CBASH</td>
<td>VS</td>
<td>VSTTESTCD.EL.SID</td>
<td></td>
<td>VSTTESTCD.EL.SID</td>
<td>Y</td>
<td>7</td>
<td>Convert</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SORRES</td>
<td>SORRES</td>
<td>SORRES</td>
</tr>
<tr>
<td>3</td>
<td>CBASH</td>
<td>VS</td>
<td>VSTTESTCD.EL.SID</td>
<td></td>
<td>VSTTESTCD.EL.SID</td>
<td>Y</td>
<td>7</td>
<td>Convert</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SORRES</td>
<td>SORRES</td>
<td>SORRES</td>
</tr>
<tr>
<td>3</td>
<td>CBASH</td>
<td>VS</td>
<td>VSTTESTCD.EL.SID</td>
<td></td>
<td>VSTTESTCD.EL.SID</td>
<td>Y</td>
<td>7</td>
<td>Convert</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SORRES</td>
<td>SORRES</td>
<td>SORRES</td>
</tr>
<tr>
<td>3</td>
<td>CBASH</td>
<td>VS</td>
<td>VSTTESTCD.EL.SID</td>
<td></td>
<td>VSTTESTCD.EL.SID</td>
<td>Y</td>
<td>7</td>
<td>Convert</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SORRES</td>
<td>SORRES</td>
<td>SORRES</td>
</tr>
</tbody>
</table>
Mapping Specifications: Dataset Level

1. data VS1;
   set CDASH.VS;
   /********
   variable level: Source Sequence = 1
   ********/
   run;

2. proc sort data=VS1; by SUBJID;
   proc sort data=CDAHS.DM OUT=DM2; by USUBJID;
   data VS2;
   merge DM2(in=a) VS1(in=b);
   by USUBJID;
   if b;
   /********
   variable level: Source Sequence = 2
   ********/
   run;

3. ... Sequence 3, 4

4. proc sort data=VS4;
   by USUBJID VSTESTCD VISITNUM VSDTC;
   run;
   data SDTM.VS;
   set VS4;
   by USUBJID VSTESTCD VISITNUM VSDTC;
   /********
   variable level: Source Sequence = 6
   ********/
   run;
Mapping Specifications: Variable Level

```sas
data VS1;
  set CDASH.VS;

  **** Variable level processing ;
  ❶ DOMAIN = 'VS';
  ❷ USUBJID = catx('.', STUDYID, SUBJID);
  ❸ VISITNUM = input(put(VISIT, $VISITNUM.), BEST.);

  [origin = Predecessor, do nothing];

  ❹ if not missing(VISDAT) then
    VSDTC = put(VISDAT, E8601DA.);
  else if not missing(VSDAT) then
    VSDTC = put(VSDAT, E8601DA.);

  ❺ if VISIT = "VISIT 2 (WEEK 0)" then VSBLFL = 'Y';
run;
```
Mapping Specifications: Value Level

```plaintext
data VS3;
  set CDASH.VS;
  if DIABP_VSPREF = 'Y' then do;
    VSTESTCD = 'DIABP';
    VSORRES = DIABP_VSORRES;
    VSORRESU = DIABP_VSORRESU;
    VSSTRESN = 'mmHg';
    VSSTRESC = INPUT(VSORRES, BEST.);
    VSPOS = DIABP_VSPOS;
    OUTPUT;
  end;
run;
```

```plaintext
*** CONTINUE ***;
if HEIGHT_VSPREF = 'Y' then do;
  VSTESTCD = 'HEIGHT';
  VSORRES = HEIGHT_VSORRES;
  VSORRESU = HEIGHT_VSORRESU;
  VSSTRESN = 'm';
  VSSTRESC = PUT(VSSTRESN, BEST.);
  VSSTRESC = INPUT(VSORRES, 4.0);
  VSPOS = DIABP_VSPOS;
  OUTPUT;
end;
run;
```
Learnings so far

Machine-readable Metadata
Machine-readable Metadata

- CDISC 360 Enriched Metadata = Structural + Conceptual + Semantic + Process [Key to Automation]
  - Content is part of the standards (CDISC library)
  - ETL Metadata (mapping inference & derivation)

- System agnostic standards, concepts and elements
  - Can be consumed by any tool
  - Organization can build an automation engine their own way

- Iterations are needed to learn and evolve
  - Strong workstream collaboration: CDISC, Industry volunteers & Microsoft
Automation of ADaM & TLF Generation using CDISC 360 enriched standards

Bhavin Busa (Vita Data Sciences), Prasanna Murugesan (AstraZeneca)

CDISC 360: The Journey So Far and the Road Ahead
April 28, 2020
Agenda

1. Current State: Analysis Datasets & TFL Execution
2. CDISC 360 Enriched Metadata
   ☐ CDISC 360 Enriched (Machine-readable) TFL Metadata
3. Future State with Concept-based Standards: Analysis Datasets & TFL Execution
4. Process Flow for CDISC 360 Proof of Concept (PoC)
5. TFL Automation Engine – PoC Design
6. TFL Automation Engine – Live Demo
7. Learnings so far
8. Next Steps
Current State - without Concept-based Standards
Current State - without Concept-based Standards: Analysis Datasets and TFL Generation

Foundational Standards

- SAP & TLF Shells*
- TFL Specs*
- TFL Outputs*
- CSR

- ADaM Specs*
- ADaM Programs*
- ADaM Datasets*
- ADaM Define*

Submission of Datasets* & CSR XPT files, Define, ADRG, SAS codes

Specify → Build/Execute → Report

* Manual Process
+ Manual or Semi-automated Execution
CDISC 360 Enriched Metadata
Selection Summary

Study Endpoint
Analysis of Glycated Hemoglobin
Analysis of the continuous clinical endpoint of HbA1c. Example: a Phase III, parallel-group study designed to determine efficacy of Drug A for patients with Type II diabetes. The primary endpoint defined as the change in HbA1c from baseline.

Analysis
Mean Change from Baseline in HbA1c (% Over Time)
Provides a visual display of the information in the "HbA1c Longitudinal Repeated Measures Analysis" table. Includes additional weeks beyond those in that table. The mean changes shown are based on adjusted changes from baseline from the repeated measures model.

Analysis Datasets
ADSL
Analysis Data Subject Level
- View analysis dataset metadata
- View sample analysis data
- View analysis dataset structure

ADHBA1C
DBS - Structured Dataset
- View analysis dataset metadata
- View sample analysis data
- View analysis dataset structure

Reference: ‘CDISC 360 - The Journey so Far and the Road Ahead’, Peter Van Reusel, 28th April 2020
Analysis Concept

Reference: ‘CDISC 360 - The Journey so Far and the Road Ahead’, Peter Van Reusel, 28th April 2020
Analysis Result Concept

Reference: ‘CDISC 360 - The Journey so Far and the Road Ahead’, Peter Van Reusel, 28th April 2020
CDISC ARM Metadata

CDISC 360 Enriched (Machine-readable) TFL Metadata
Additional TFL Metadata Required for Automation

Table 14.1.1.1
Demographic characteristics (Safety Population)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>METFORMIN (N=XX)</th>
<th>HUMAN INSULIN (N=XX)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>XX</td>
<td>XX</td>
</tr>
<tr>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>XX.X</td>
<td>XX.X</td>
</tr>
<tr>
<td>SD</td>
<td>XX.XX</td>
<td>XX.XX</td>
</tr>
<tr>
<td>Min</td>
<td>XX</td>
<td>XX</td>
</tr>
<tr>
<td>Q25</td>
<td>XX.X</td>
<td>XX.X</td>
</tr>
<tr>
<td>Median</td>
<td>XX.X</td>
<td>XX.X</td>
</tr>
<tr>
<td>Q75</td>
<td>XX.X</td>
<td>XX.X</td>
</tr>
<tr>
<td>Max</td>
<td>XX</td>
<td>XX</td>
</tr>
</tbody>
</table>

Age Group - n (%)  
15 - <30 years  XX (XX.X)  XX (XX.X)  
30 - <45 years  XX (XX.X)  XX (XX.X)  
>=45 years  XX (XX.X)  XX (XX.X)  

Gender - n (%)  
Male  XX (XX.X)  XX (XX.X)  
Female  XX (XX.X)  XX (XX.X)  

Max = Maximum, Min = Minimum, N = Number of subjects in treatment group, n = Number of subjects included in analysis, SD = Standard deviation.

## CDISC 360 Enriched TFL Metadata Tables

<table>
<thead>
<tr>
<th>Metadata View Table</th>
<th>Description</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>The contents and format of each output (which displays, file format, etc.)</td>
<td>One record per Output per Display</td>
</tr>
<tr>
<td>Display</td>
<td>List of all Displays - both generic library Display and study-specific (using in 1 or more Output)</td>
<td>One record per Display per Version</td>
</tr>
<tr>
<td>Result</td>
<td>All result metadata required to describe the analysis and create display in output</td>
<td>One record per Result</td>
</tr>
<tr>
<td>WhereClause</td>
<td>All the component parts of a where clause used to filter data</td>
<td>One record per where clause component</td>
</tr>
<tr>
<td>Style</td>
<td>Stylesheet parameters associated with Outputs</td>
<td>One record per Style per parameter</td>
</tr>
</tbody>
</table>
# CDISC 360 Enriched TFL Metadata Tables – Sample

## Demographic characteristics (Safety Population)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>METFORMIN (N=XX)</th>
<th>HUMAN INSULIN (N=XX)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years) n</td>
<td>XX</td>
<td>XX</td>
</tr>
<tr>
<td>Mean</td>
<td>XX.X</td>
<td>XX.X</td>
</tr>
<tr>
<td>SD</td>
<td>XX.X</td>
<td>XX.X</td>
</tr>
<tr>
<td>Min</td>
<td>XX</td>
<td>XX</td>
</tr>
<tr>
<td>Q25</td>
<td>XX.X</td>
<td>XX.X</td>
</tr>
<tr>
<td>Median</td>
<td>XX.X</td>
<td>XX.X</td>
</tr>
<tr>
<td>Q75</td>
<td>XX.X</td>
<td>XX.X</td>
</tr>
<tr>
<td>Max</td>
<td>XX</td>
<td>XX</td>
</tr>
</tbody>
</table>

**Age Group - n (%)**

<table>
<thead>
<tr>
<th>Age Group</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 - &lt;30</td>
<td>XX (XX.X)</td>
</tr>
<tr>
<td>30 - &lt;45</td>
<td>XX (XX.X)</td>
</tr>
<tr>
<td>&gt;=45</td>
<td>XX (XX.X)</td>
</tr>
</tbody>
</table>

**Gender - n (%)**

<table>
<thead>
<tr>
<th>Gender</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>XX (XX.X)</td>
</tr>
<tr>
<td>Female</td>
<td>XX (XX.X)</td>
</tr>
</tbody>
</table>

---

**Notes:**
- Max = Maximum, Min = Minimum.
- N = Number of subjects in treatment group. n = Number of subjects included in analysis. SD = Standard deviation.
- Datasets used: adol
- Executed by <username> on DOM1NTY: HH:MM

---

## Display View

### DisplayID
- T14111_SAF_DEMOG
- T14111_01_SAF_DEMOG
- T14111_02_SAF_DEMOG
- T14111_03_SAF_DEMOG
- T14111_04_SAF_DEMOG
- T14111_05_SAF_DEMOG

### WhereClause
- Table 14.1.1.1: Demographic characteristics (SAF)

---

## WhereClause View

<table>
<thead>
<tr>
<th>WhereClauseID</th>
<th>Dataset</th>
<th>WhereClause Variable</th>
<th>WhereClause Comparator</th>
<th>WhereClause Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 14.1.1</td>
<td>01 SAF_DEMOG</td>
<td>AGE</td>
<td>eq</td>
<td>15 &lt;= to &lt;30 years</td>
</tr>
<tr>
<td>Table 14.1.1</td>
<td>02 SAF_DEMOG</td>
<td>AGE</td>
<td>eq</td>
<td>30 &lt;= to &lt;45 years</td>
</tr>
<tr>
<td>Table 14.1.1</td>
<td>03 SAF_DEMOG</td>
<td>AGE</td>
<td>eq</td>
<td>&gt;=45 years</td>
</tr>
<tr>
<td>Table 14.1.1</td>
<td>04 SAF_DEMOG</td>
<td>SEX</td>
<td>eq</td>
<td>F</td>
</tr>
<tr>
<td>Table 14.1.1</td>
<td>05 SAF_DEMOG</td>
<td>SEX</td>
<td>eq</td>
<td>M</td>
</tr>
</tbody>
</table>
Future State - with Concept-based Standards
Future State - with Concept-based Standards: Analysis Datasets and TFL Generation

Analysis Concepts
Analysis Result Concepts
Foundational Standards

Design

SAP*
TFL Designer*

ADaM Specs
ADaM Programs
TFL Shells & Metadata
TFL Programs

TFL Outputs + ARM
CSR

Submission of Datasets & CSR
XPT files, Define, ADRG, SAS codes

* = Manual Process
= Automated Process

Specify Build/Execute Report
Process Flow for CDISC 360 Proof of Concept (PoC)
Process Flow for CDISC 360 PoC
TFL Automation Engine – Proof of Concept Design
CDISC 360 – TFL Automation Engine PoC Design

**CDISC 360**
- Enriched ARM Metadata

**R Shiny**
- Select TFL of Interest
- Select TFL Layout (Template)
- Review data
- Generate SAS Program and Define.xml*
- Customize TFL Layout & Metadata

**SAS**
- Execute SAS Program
- Generate Output
- Validate and Deliver

---

* ARM to be combined with ADaM Define
TFL Automation Engine – Live Demo!
Learnings so far
Machine-readable TFL Metadata

• ARM + additional TFL Metadata
  o Use case tested with enriched metadata
  o Can be consumed by any tool

• TFL Automation Engine PoC
  o Demonstrated execution of TFL & generation of Define.xml + ARM
  o Organization can build an automation engine their own way

• TFL Designer
  o Will help build TFL Shells and ADaM Specs

• CDISC currently does not support TFL standards – can templates be developed?
Next Steps

Our plans for remaining part of PoC
Next Steps in Automate Execution
[CDASH → SDTM → ADaM → TLFs]

• Collaborate with other workstreams to develop concept-based standards, ODM CRF, and Define-XML [per CDISC 360 defined scope]

• CDASH to SDTM
  • Test & finalize machine-readable metadata elements for mapping specifications
  • Autogenerate SDTM artifacts from CDASH via CDISC 360 Process Flow for PoC [DM, EX, LB, VS, and trial design domains]

• SDTM to ADaM
  • Define, test & finalize machine-readable metadata elements for mapping specifications
  • Autogenerate ADaM artifacts from SDTM via CDISC 360 Process Flow for PoC [ADSL]

• ADaM to TFL
  • Adjust TFL metadata to meet CDISC ARM v1.0 for Define-XML v2.0 standards
  • TFL Designer – will be conceptualize but team to hold on further development of PoC
Thank You!

Bhavin Busa, Vita Data Sciences
Jianhui [Jimmy] Zhao, Allergan
Prasanna Murugesan, AstraZeneca

Courtesy: Mahi Busa