2023 EUROPE INTERCHANGE

COPENHAGEN | 26-27 APRIL

Ensuring that a Define-XML is Submission Ready

Presented by Caro Sluijter, Senior Clinical Programmer,
OCS Life Sciences

Meet the Speaker

Caro Sluijter

Title: Senior Clinical Programmer & Proposition Lead Training

Organization: OCS Life Sciences

Caro obtained her masters in Biomedical Science with a major in Epidemiology from the Radboud University in Nijmegen. For the last four and a half years at OCS Life Sciences, Caro, as a project lead, has been involved in multiple CDISC data conversion projects. Part of her work in these projects is to create Define-XMLs which are ready for submission. Furthermore, in her role as the proposition lead training, Caro is also involved in the creation and delivery of multiple trainings, such as Define-XML training.

Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author and do not necessarily reflect the official policy or position of CDISC.*

- *The author has no real or apparent conflicts of interest to report.*

cdisc

Agenda

1. **Introduction**
   What is a Define-XML
   Define-XML creation at OCS Life Sciences
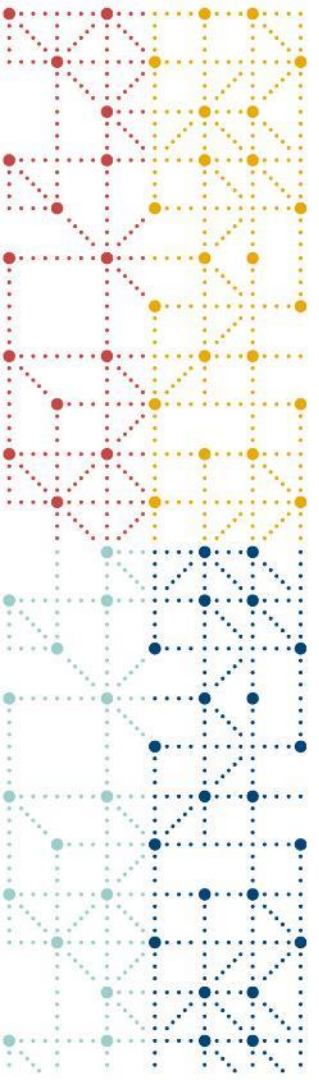
2. **Creation of Define-XML**
   Creating a Define-XML
   Updating a Define-XML

3. **Updating to Define-XML V2.1**
   What's new in V2.1
   Challenges Define-XML V2.0 vs V2.1

Introduction

What is a Define-XML
Define-XML creation at OCS Life Sciences

What is a Define-XML

Definition:

Define-XML transmits metadata that describes any tabular dataset structure. When used with the CDISC Foundational Standards, it provides the metadata for human and animal model datasets using the SDTM and/or SEND standards and analysis datasets using ADaM.

Define-XML Creation at OCS Life Sciences

- CDISC Data Conversion projects in different trial phases
  - Personal experience with phase I trials

- CDISC conversion projects from start to finish
  - Case Report Form (CRF) annotation through Define-XML creation
  - Prior knowledge

- Check and update existing Define-XMLs from clients
  - Not involved in data conversion
  - No prior knowledge
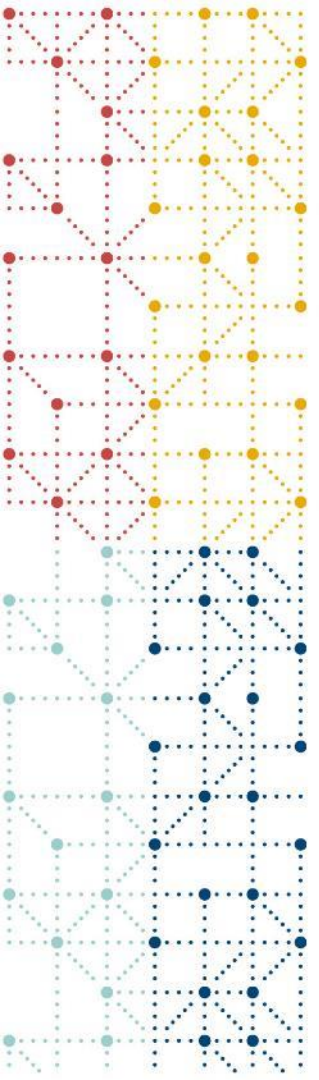
Creation of a Define-XML

Creating a Define-XML

Updating an existing Define-XML

Creating a Define-XML

- ## Starting from 'scratch'

  - ### Involved in the CDISC conversion process
    - Annotating the CRF
    - Creating mapping specifications
    - Do the actual programming

  - ### Have all information needed to create our own Define-XML

  - ### Pinnacle 21 Community Define.xml Generator
    - Generate Spec
    - Populates Datasets, Variables, Lengths

Creating a Define-XML

- Starting from 'scratch'

  - But **P²¹** Community does not complete your Define!
    - Add CRF pages
      - Need to add CRF page when variable is collected within CRF
    - Add value level metadata
      - Need to describe differing metadata attributes for subsets of cells within a column
    - Add derivation methods
      - Needed for all variables that have a derived origin
    - Determine the correct origin

  - Tips to prepare
    - Write mapping specifications that can be used as Define-XML Methods
      - No use of source dataset for SDTM
      - Details how data transitions from SDTM to ADaM
      - Use of sections to be able to split specifications for value level

Creating a Define-XML

- Starting from 'scratch'

  - If Excel is completed

  - can generate the Define-XML
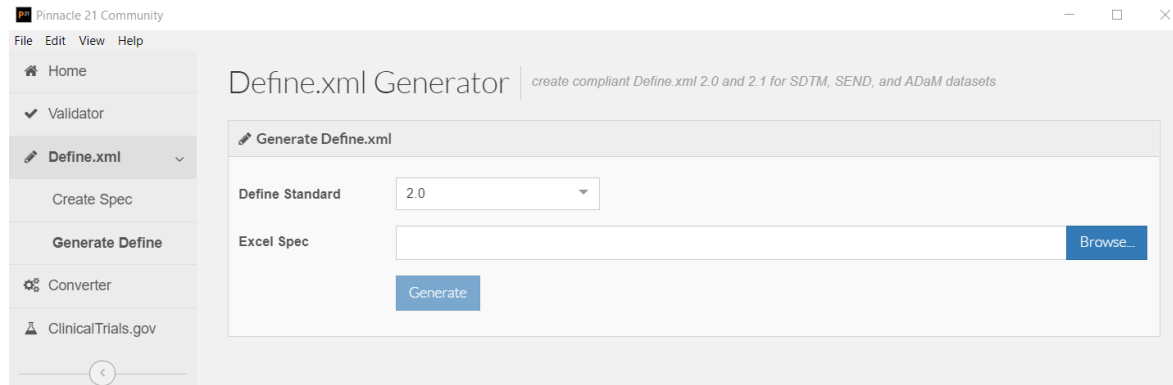
Creating a Define-XML

- Starting from 'scratch'

  - First checks can be done by
    - Variable order and label check
    - Check data types and lengths
    - Correct linking of Variables and their Methods and Comments
    - Correct linking Variables and Value Level
    - Codelist
      - Values
      - NCI Codelist Codes/NCI Value Codes

  - Make sure all resolvable issues are resolved
    - Define-XML
    - Define-XML + SDTM/ADaM datasets

Creating a Define-XML

- Starting from 'scratch'

  - Next step: validation

  - Validator ensures that programmer resolved all P21 issues

  - Validator need to check what P21 cannot

OCS Project Number:	<3 digit code>	Client Project ID:	<client reference>
Client:	<client name>		
Project Name:	<OCS assigned project name>		
Document:	CL0003F04 Version 1		

OCS LIFE SCIENCES
A DIVISION OF OCS CONSULTING

Other issues and comments

Description of other checks or findings:	Pass	Fail	N/A
28. Decoded Value is present for abbreviated terms only and correctly describes them.			
29. All messages in the validation report are as expected.			

Sign-off

Reviewer	Developer
By signing I confirm that I have executed all checks on this form, and that there were no findings or all findings have been resolved.	By signing I confirm that there were no findings or that I have resolved all findings.
Date:	Date:
Signature:	Signature:

Description of issues identified to be resolved, or rationale for selecting N/A

#	Check Id	Description of finding	Action (fix/leave)	Response by programmer <isodate + initials>	Result by validator <isodate + initials>
1					
2					
3					
4					
5					
6					

Description other issues and comments and notes

Other issues and comments

Notes

© OCS Life Sciences 2021

Page 2 of 2

cdisc

Creating a Define-XML

- Check in Excel

  - Key Variables match sorting variabl[...]
    programs
  - Structure logical and in line with ke[...]
    variables
  - Class matches specification in
    Implementation Guide
  - Format of datetime and other numeric
    variables
  - Number of significant digits are correct
  - Mandatory column matches required
    variables in Implementation Guide

PEORRES VLM		Verbatim Examination Finding	text	Result Qualifier	82		
	PECAT = "PHYSICAL EXAMINATION" and PETESTCD = "ABD" (Abdomen)		text		82	Peorres Value For Rest • "NO" = "No" • "YES.NOTE" = "Yes.Note"	Derived Set to "NORMAL" when result collected on acrf is equal to "No". Otherwise, copy from notes when result is equal to "Yes Note", removing "1:" when present.
			text		20	Peorres Value For Breast • "NOT DONE" = "Not Done" • "YES" = "Yes" • "NO.NOTE" = "No.Note"	Derived Set to "NORMAL" when result collected on acrf is equal to "No". Otherwise, copy from notes when result is equal to "Yes Note", removing "1:" when present.
	PECAT = "PHYSICAL EXAMINATION" and PETESTCD = "CNS" (Central nervous system)		text		20	Peorres Value For Rest • "NO" = "No" • "YES.NOTE" = "Yes.Note"	Derived Set to "NORMAL" when result collected on acrf is equal to "No". Otherwise, copy from notes when result is equal to "Yes Note", removing "1:" when present.

- Scroll through HTML page to find 'ugly' things:
  - Same method present at Variable and Value level
  - Lack of interpunction
  - Strange values in codelist

cdisc
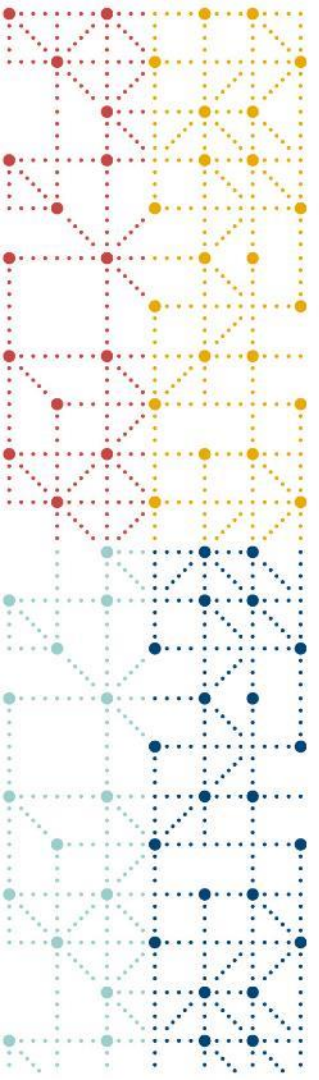
Creation of a Define-XML

Creating a Define-XML

Updating an existing Define-XML

Updating an existing Define-XML

- Make sure an existing Define-XML is ready for submission

- Look at the submission package as a reviewer would do
  - No prior knowledge
  - No access to executable programs
  - Connected with Data Reviewer's Guide (DRG)

- Finding out what is meant / what is done

- Sometimes easier to fix/update what is not yours

Define-XML is Submission Ready

- All P21 resolvable issues resolved

- Remaining P21 issues described in cSDRG/ADRG
- Large derivations are described in cSDRG/ADRG or separate document

- Correct stylesheet is present
- CDISC datasets are present in XPT format
- All referenced documents are named correctly and present
    - aCRF
    - cSDRG/ADRG
    - Protocol/SAP

cdisc
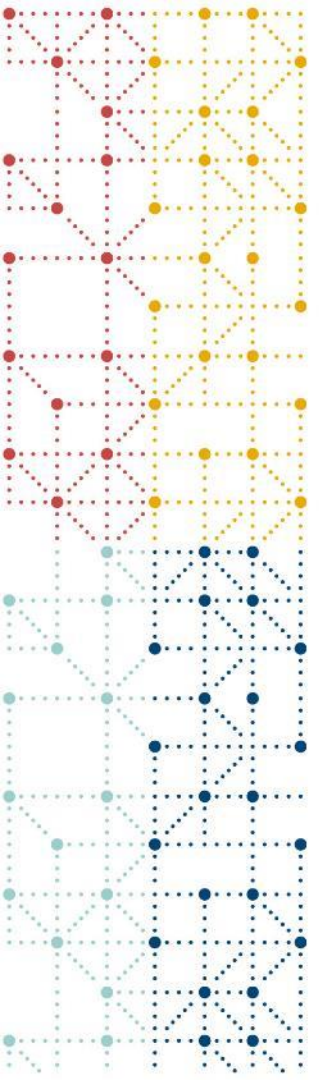
Updating to Define-XML V2.1

What's new in V2.1?

Challenges Define-XML V2.0 vs V2.1

What's new in V2.1?

- SDTM domains annotated but no data collected may be identified

SUPPDM [SDTMIG 3.1.2]	Supplemental Qualifiers for DM (Demographics)	RELATIONSHIP	One record per IDVAR, IDVARVAL, and QNAM value per subject	Tabulation	STUDYID, RDOMAIN, USUBJID, IDVAR, IDVARVAL, QNAM		suppdm.xpt
SUPPVS [SDTMIG 3.1.2] [No Data]	Supplemental Qualifiers for VS (Vital Signs)	RELATIONSHIP	One record per IDVAR, IDVARVAL, and QNAM value per subject	Tabulation	STUDYID, RDOMAIN, USUBJID, IDVAR, IDVARVAL, QNAM		

- SDTMIG version identified and referenced by dataset (domain)
- CDISC controlled terminology identified and referenced by codelist

Datasets

Dataset	Description	Class	Structure	Purpose	Keys	Documentation	Location
TA [SDTMIG 3.3]	Trial Arms	TRIAL DESIGN	One record per planned Element per Arm	Tabulation	STUDYID, ARMCD, TAETORD		ta.xpt
TE [SDTMIG 3.3]	Trial Elements	TRIAL DESIGN	One record per planned Element	Tabulation	STUDYID, ETCD		te.xpt
TI [SDTMIG 3.3]	Trial Inclusion/Exclusion Criteria	TRIAL DESIGN	One record per I/E criterion	Tabulation	STUDYID, IETESTCD, TIVERS		ti.xpt

Age Unit in Years [C66781] [CDISC/NCI SDTM 2020-12-18]

Permitted Value (Code)	Display Value (Decode)
YEARS [C29848]	Years

cdisc

What's new in V2.1?

Datasets

Dataset	Description	Class - SubClass	Structure	Purpose	Keys	Documentation	Location
ADAE [ADaMIG 1.1]	Adverse Events Analysis Dataset	OCCURRENCE DATA STRUCTURE - ADVERSE EVENT	one record per subject per adverse event	Analysis	STUDYID, USUBJID, AETERM, ASTDT, AESEQ	See SAS program adae.sas	adae.xpt

- The Class concept has been adjusted to support the addition of SubClass

- Initial use cases are:

  - Identify Adverse Events Analysis datasets within ADaM Occurrence Data Structure class

  - Identify Time to Event datasets within ADaM Basic Data Structure class

cdisc

What's new in V2.1?

- Representation of SDTM Origin metadata has been enhanced
  - Identify the source in addition to the origin details
  - To support multiple origins

- Origin on variable level is changed
  - CRF and eDT are no longer available → Origin is now 'Collected'

- Source is added to support Origin
  - Subject, Investigator, Vendor or Sponsor

What's new in V2.1?

Origin	Source				Notes
	Subject	**Investigator**	**Vendor**	**Sponsor**	
Collected	ePro	CRF	Lab data, ECG	X	This term should be used for clinical data that were actually observed or recorded by a person or received from an instrument; it should not be used for data that have been interpreted, calculated, or derived from other information.
Derived	X	X	Lab data, ECG	SDTM	Derivation examples include calculations performed during data collection (e.g., --DY). Other derivation examples: calculations within ePRO (e.g., questionnaire section scores) and calculations within EDC (e.g., BMI, BSA).
Assigned	X	X	Adjudicator	SDTM	Examples of this include third-party attributions by an adjudicator, coded terms that are supplied as part of a coding process, and values that are set independently of any subject-related data values in order to complete SDTM fields such as DOMAIN and --TESTCD
Protocol	X	X	X	SDTM	An example would be VSPOS (Vital Signs Position), which could be specified in the protocol and be provided by other means (e.g. CRF, eDT).
Predecessor	X	X	X	X	Use when a value is an exact copy of another value in an SDTM dataset.

NOTES:

o Cells that include **X** indicate that the combination of Origin Type and Source is not valid

o Origin 'Not Available' is not valid

What's new in V2.1?

Origin	Source	Notes
	Sponsor	
Derived	ADaM	For sponsor-performed analysis derivations in ADaM
Assigned	ADaM	
Predecessor	**X**	Use when a value is an exact copy of another value in either SDTM or ADaM dataset.

NOTES:

○ Cells that include **X** indicate that the combination of Origin Type and Source is not valid

○ Origin 'Not Available' is not valid

○ Source is always 'Sponsor' for ADaM Derived and Assigned origin types. For Predecessor, Source is not used.

What's new in V2.1?

- A def:Context attribute has been added to the ODM root element to indicate that Define-XML is for a regulatory submission

    - Define-XML Context: Submission

    - Define-XML Context: Other

```xml
<ODM
    xmlns:xlink="http://www.w3.org/1999/xlink"
    xmlns="http://www.cdisc.org/ns/odm/v1.3"
    xmlns:def="http://www.cdisc.org/ns/def/v2.1"
    ODMVersion="1.3.2"
    FileType="Snapshot"
    FileOID="CDISCPILOT01.SDTMIG.3.3"
    CreationDateTime="2022-09-27T09:18:47+02:00"
    SourceSystem="Pinnacle 21 Community"
    SourceSystemVersion="4.0.1"
    def:Context="Submission">  ⬅
  <Study OID="CDISCPILOT01.SDTMIG.3.3">
```
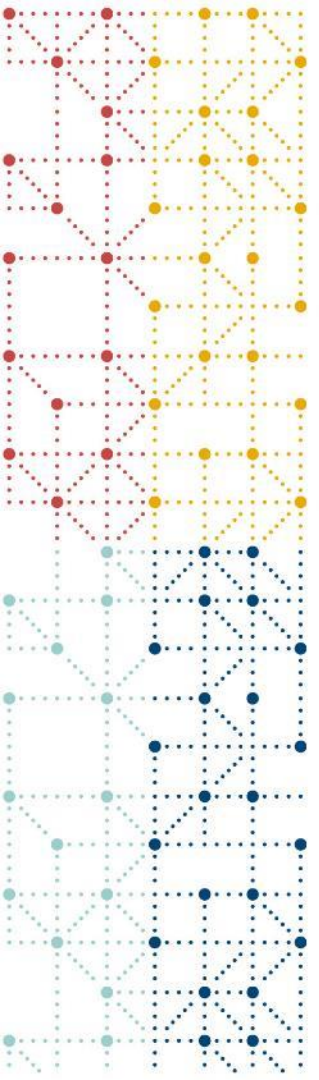
Date/Time of Define-XML document generation: 2022-09-26T11:26:15+02:00
Define-XML version: 2.1.0
Define-XML Context: Submission ⬅
Stylesheet version: 2019-02-11

Study Name	CDISCPILOT01
Study Description	Study Data Tabulation Model Metadata Submission Guidelines Sample Study
Protocol Name	CDISCPILOT01
Metadata Name	Study CDISCPILOT01 Data Definitions

Updating to Define-XML V2.1

What's new in V2.1?

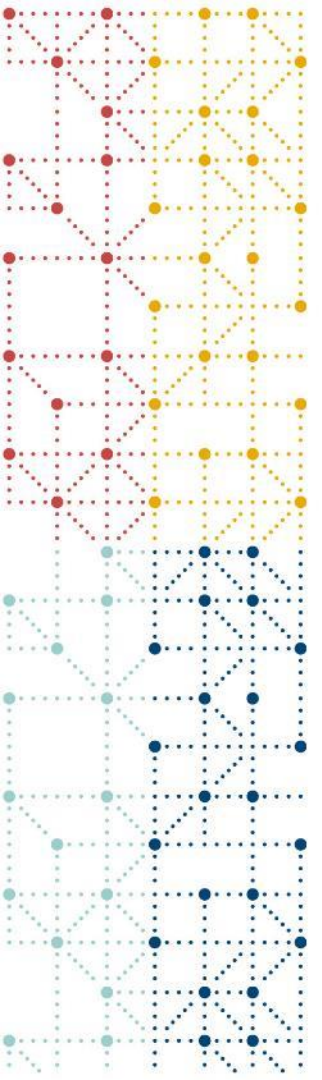Challenges Define-XML V2.0 vs V2.1

Challenges Define-XML V2.0 vs V2.1

- Getting used to the new definition of origin and source
  - Assigned vs Derived
  - Subject vs Investigator vs Vendor

- New Excel sheet no longer contains Where-Clause tab
  - Create own SQL where-clause at Variable-level tab

Tips

Tips

- Developer should also use P21 Validator for
  - Define-XML
  - Define-XML + datasets

- If developer has prior knowledge of CDISC conversion then get an independent validator

- Only use a comment when needed
  - Do not add information from Implementation Guide in comment section

- Support your Define-XML with the Reviewer's Guide

- Make sure which version of the Define-XML is needed

Thank You!

cdisc