Public Review Webinar: Study Data Tabulation Model Metadata Submission Guidelines v2.0

Richard Lewis, SDTM Consultant, Data Standards Consulting Group (A division of Talent-Mine)
Mike Hamidi - Head of Data Science, CDISC
Marcelina Hungria, Managing Director/Owner, DIcore Group

Thursday, 24 SEP 2020
11:00AM – 12:30PM EDT

Today’s Agenda

1. Housekeeping
2. Presenter Introductions
3. Feature Presentations
4. Question & Answer Session
5. Upcoming Learning Opportunities + Resources

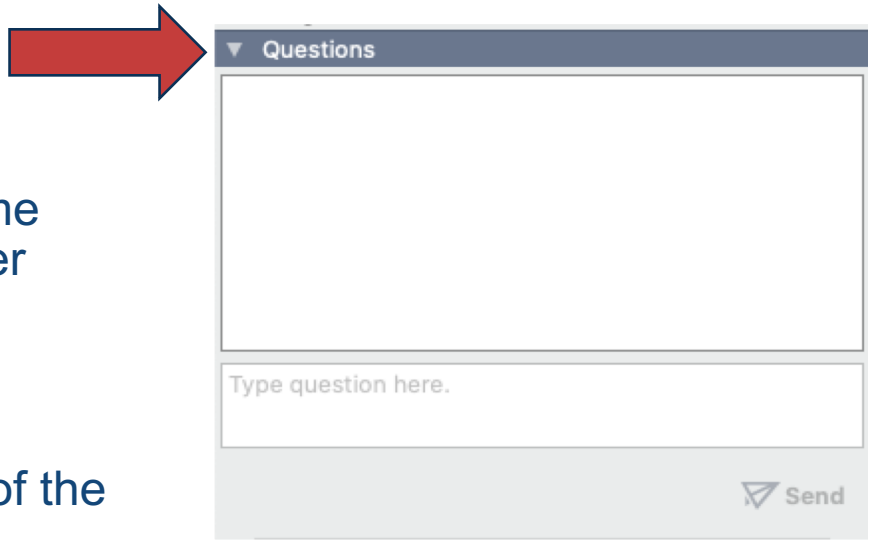
Housekeeping

Housekeeping

- You will remain on **mute** for the entirety of the call
- There will be a Q&A after all of the presentations are finished
- Audio issues? Shut down and restart the GoToWebinar app
- The slides from the presentation and a recording of this webinar will be available in the Members Only section of the CDISC website
  - To access – make sure that you create a login for the CDISC website if you haven't already
  - If you are employed by a CDISC member organization, please ensure you use your employer-issued email address with your employer's domain name, so we can verify membership for the purpose of applying discounts to purchasing event tickets, online courses, and more!

Submitting Questions

- To send a question, use the “QUESTIONS” function on your GoToWebinar app. (See red arrow)
- You can submit questions at any time during the presentation, we’ll answer them during the Q&A.
- If you have a question for a specific presenter, please indicate the presenter’s name at the beginning of the question
  - Examples:
    - John: ‘Question’
    - Alana: ‘Question’



Content Disclaimer

- The purpose of this webinar is to provide examples of implementation and should not be considered official recommendations by CDISC unless otherwise stated in the presentation.
- This webinar is not an authorized CDISC course, is not developed or delivered under CDISC Operating Procedures, and should not replace a published standard. Please refer to the latest published standards for the most authoritative implementation information.

Our Presenters

- Richard Lewis, SDTM Consultant, Data Standards Consulting Group (A division of Talent-Mine)
- Mike Hamidi - Head of Data Science, CDISC
- Marcelina Hungria, Managing Director/Owner, DIcore Group

MSG 2.0 For Public Review

Presented by:

Richard Lewis, SDTM Advisor, Data Standards Consulting Group

Mike Hamidi, Head of Data Science, CDISC

Marcelina Hungria, Consultant / Owner, DIcore Group

2020-09-24

Agenda

1. Introduction
2. MSG Purpose
3. MSG 2.0 Summary of Changes
4. MSG 2.0 aCRF and cSDRG Details
5. MSG 2.0 Define-XML Document and Other Details
6. MSG 2.0 Public Review - Logistics and Expectations
7. Questions

Introduction

SDTM MSG History

Impact

Purpose

Study Data Tabulation Model Metadata Submission Guidelines (SDTM-MSG): Human Clinical Trials

Introduction

SDTM MSG History

- MSG v1.0 developed from 2006-2011
- Draft released in May, 2010
- Final released end of 2011
  - MSG Document
  - Sample Submission Package
    - Datasets
      - SDTMIG v3.1.2
    - Define-XML v1.0
    - Annotated CRF
    - Reviewer's Guide

Impact of v1.0

- Submissions became more uniform in nature
  - Same “Look and Feel”
- Reviewer’s Guides became common in submissions

MSG Purpose

To provide guidance for preparing the components of the International Conference on Harmonisation (ICH) electronic Common Technical Document (eCTD) Module 5 (M5) Clinical Study Reports "sdtm" folder

- Recommendations beyond what is in the standards documents
  - Not requirements
- Not a repeat of content from the SDTM, SDTMIG or Define-XML Specification
  - Does not contradict or supersede their content
- ADaM and SEND MSG packages may be developed separately
  - Until then the SDTM MSG can be used where appropriate

MSG 2.0 Summary of Changes

MSG 2.0 Summary of Changes

MSG document completely rewritten

- Less content on the Define-XML Standard
  - References provided instead

Define-XML Document

- Based on the Define-XML v2.1 Standard

All new CRF created by MSG team

New sample data

- Based on the data from the original SDTM/ADaM Pilot Project
  - Adjusted to comply with SDTM v1.7/SDTMIG v3.3
  - More robust than the v1.0 data, more domains and subjects
- Provided in SAS V5 and Dataset-XML transport formats

Sample Reviewer's Guide included

MSG 2.0 Details

Specific Data Examples Added

Domain intended for which no data were present

- NV
- Several SUPPQUALs

Variables intended to have data collected but no values were

- SUPPDM race variables

Example of domain split by sponsor decision

Example of domain split due to file size

QS, FT and RS (non-oncology)

- To illustrate the new domains for the different data which were contained in the QS domain in previous IGs

Specific Data Examples Added – Cont.

IE Criteria of excessive length

- Original criteria as long as 2199 characters long
- Amendment changes occurred towards end of text
- Changes could not be represented in meaningful way
- Criteria references the Protocol or the Amendment

USUBJID values included which do not use conventional STUDY-SITE-SUBJID format

- Used “CDISCnnn” format instead

Different data collected on replacement CRF pages

Overall CRF Recommendations

Defined “Unique” when referencing CRF pages

- v1.0 used the term but did not define
- As in 1.0, recommend to only submit unique pages

Clarified that CRF pages are not limited to traditional CRF pages

- Non-traditional collection devices
- ePRO

Replacement Pages

Simple example of replaced CRF pages included

Shows example of how that situation can be handled

- Indicated in TOC and bookmarks
- Did not recommend one way to handle replacement pages

Raised the topic of including replacement pages in the aCRF

- Not always done in practice
- Often only last, or first, version is submitted

Replacement Pages – Cont.

Recommends against "See page…" incomplete annotations

- Only one or several data points are changed on a CRF page
- Common practice
  - Only annotate the changed data points
  - Make reference to page where the other data points are annotated
- Recommendation is to fully annotate all versions
- Also applies to non-unique pages

MSG 2.0 Non CRF-Collected Annotations

Many sponsors/CROs expect a 1:1 relationship between CRF-collected and annotations

- “CRF” or “Collected”, depending on Define-XML version

Others do not necessarily treat annotations as collected

The result is inconsistencies in handling origins

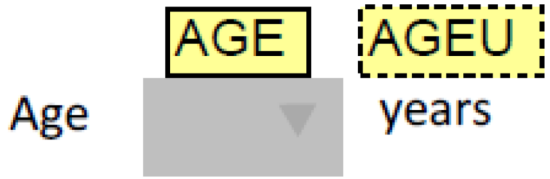
These annotations are often helpful to understand the collected data

- Knowing --METHOD or --SPEC might help clarify what data were collected on lab page
- MHCAT often annotated, usually not actually collected though

MSG 2.0 Non CRF-Collected Annotations – Cont.

Vary annotation slightly for those variables

- Dotted border around annotation
- Reviewer can easily see which variables are CRF collected, and which are not, by looking at the CRF
  - No need to cross-reference Define-XML document each time
- Additional informative annotations can be added without ambiguity



MSG v1.0 Annotation Text Style

MSG v1.0 did not explicitly recommend any font color or style

Annotation style implied through sample annotated CRF

- No size preference

Domain annotations in black, bold text, Arial font

VS=Vital Signs

Variable annotations in red, bold, italic text, Arial font

VSORRES when VSTESTCD = DIABP

MSG v2.0 Annotation Text Style

Annotation style preference included in MSG

Domain annotations in bold, black text, Arial font

- Using “( )” format to differentiate from variable annotations

VS (Vital Signs)

Variable annotations in black text without bold formatting, Arial font

VSORRES when VSTESTCD = DIABP

MSG v2.0 Annotation Text Style Comparison

Domain Annotation

- V1.0

VS=Vital Signs

- V2.0

VS (Vital Signs)

Variable Annotation

- V1.0

VSORRES when VSTESTCD = DIABP

- V2.0

VSORRES when VSTESTCD = DIABP

MSG 2.0 Annotation Background Colors

MSG v1.0 recommended new color per domain on CRF page

- Color choice was sponsor defined
- Some colors chosen made annotations more difficult to read
- Probably no consideration for color-blind reviewers

MSG 2.0 Annotation Background Colors – Cont.

Recommends new color per domain on a page

- Recommended color order defined
- Colors confirmed through Coblis (Color Blindness Simulator)

BLUE	DM (Demographics)	191, 255, 255
YELLOW	DS (Disposition)	255, 255, 150
GREEN	SC (Subject Characteristics)	150, 255, 150
ORANGE	VS (Vital Signs)	255, 190, 155

MSG RELREC Annotations

V1.0 had example of:

“Linked to related AE record via RELREC”

V2.0 explicitly indicates the use of:

“RELREC when <collected variable>=<related domain variable>”

- “RELREC when DDLNKID = AELNKID”
- Shorter for crowded CRF pages
- Conveys more information than previous method

MSG 2.0 Additional Methods for Annotating

Arrows, lines and boxes

- Recommends against overuse, to be used only if necessary
- Several examples included in aCRF
- No recommendation on exactly when to use or colors involved
  - Too many possibilities, needs to be sponsor decision

MSG 2.0 Bookmarking aCRF

Recommendations for dual bookmarking the same

- By Chronology (by Visit)
- By CRF Topics or Form
  - Top bookmark changed from “Domain”, as recommended in v1.0, to “Form”
- “Running Records” indicated as CRF pages which are independent of visits
  - Were included in v1.0 but not explained

MSG 2.0 Clinical Study Data Reviewer's Guide (cSDRG)

The PHUSE template already has a template and completion guidelines

- The MSG only added recommendations to common shortcomings seen in sponsor cSDRGs

Indicated that redundant documentation should be minimized in the cSDRG

- The Define-XML document should include all relevant documentation

MSG 2.0 Define-XML Document

MSG 2.0 Define-XML Document (Define-XML v2.1 compliant)

Planned metadata

- hasNoData mostly relevant at data exchange with exceptions
  - Comments are necessary to explain each case
- Origin definitions – Variable and VLM
  - References to aCRF not only for Collected origin type
  - Aimed at helping the reviewer

Codelists definitions – Variable and VLM

- All values possible/planned, not just present in the data
- Subsets illustrated extensively

VLM

- Several cases illustrated
  - DM (Race), LB (ORRES/ORRESU, STRES/STRESU) , VS, ...
- Must not contradict definitions at Variable level

MSG 2.0 Sample Stylesheet

No stylesheet version recommended

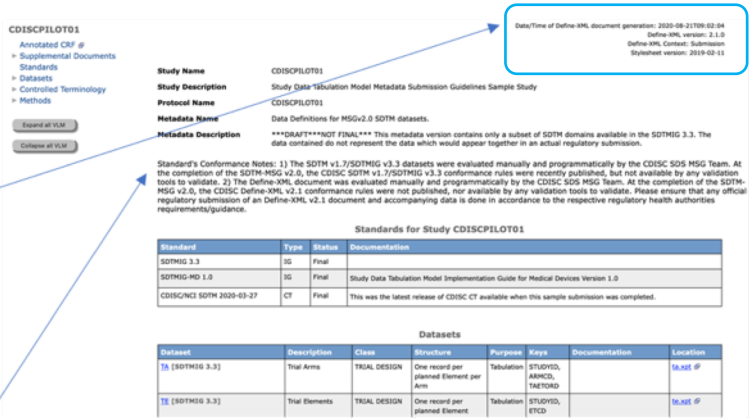
- Stylesheets improve over time
- Any recommendation would be outdated soon

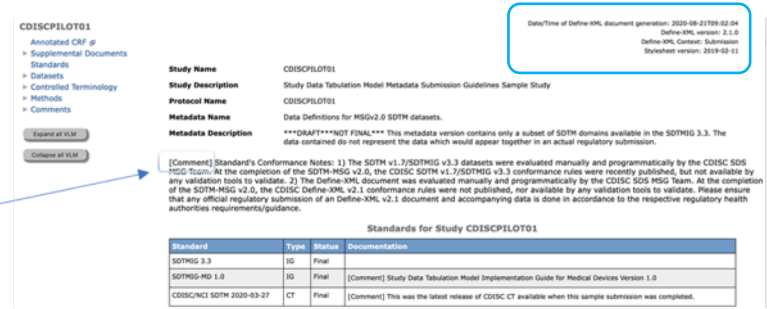
MSG 2.0 Define-XML Document & Stylesheet Views

```
1  <?xml version="1.0" encoding="UTF-8"?>
2  <?xml-stylesheet type="text/xsl" href="define2-1.xsl"?>
23 <xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
191   <xsl:variable name="STYLESHEET_VERSION" select="'2019-02-11'"/>

15 <ODM
16   xmlns="http://www.cdisc.org/ns/odm/v1.3"
17   xmlns:xlink="http://www.w3.org/1999/xlink"
18   xmlns:def="http://www.cdisc.org/ns/def/v2.1"
19
20   ODMVersion="1.3.2"
21   FileOID="www.cdisc.org/StudyMSGv2/1/Define-XML_2.1.0"
22   FileType="Snapshot"
23   CreationDateTime="2020-08-21T09:02:04"
24   Originator="CDISC MSG Team"
25   SourceSystem="M.Hungria-System"
26   SourceSystemVersion="2.1-A1"
27   def:Context="Submission"
28   >
```

```
11709     <def:CommentDef OID="COM.MDV">
11710       <Description>
11711         <TranslatedText xml:lang="en">Standard's Conformance Notes:
11712 1) The SDTM v1.7/SDTMIG v3.3 datasets were evaluated manually and programmatically by the CDISC SDS MSG
11712 Team. At the completion of the SDTM-MSG v2.0, the CDISC SDTM v1.7/SDTMIG v3.3 conformance rules were
11712 recently published, but not available by any validation tools to validate.
11713 2) The Define-XML document was evaluated manually and programmatically by the CDISC SDS MSG Team. At
11713 the completion of the SDTM-MSG v2.0, the CDISC
11714 Define-XML v2.1 conformance rules were not published, nor available by any validation tools to
11714 validate. Please ensure that any official regulatory submission of an Define-XML v2.1 document and
11714 accompanying data is done in accordance to the respective regulatory health authorities
11714 requirements/guidance.</TranslatedText>
11715       </Description>
11716     </def:CommentDef>
```





MSG 2.0 Conformance, Validation and Tools

A Define-XML document must be a valid document

- Compliant with the indicated Define-XML Standard version

The data must be validated using a valid Define-XML document

- Compliant with the indicated model and IGs Standard versions

Errors and warnings should be carefully explained

- Don’t simply repeat the message back
- No generic explanations

MSG 2.0 Public Review - Logistics and Expectations

MSG 2.0 Public Review

New Reviewers

Dear CDISC Community,

CDISC is pleased to announce version 2.0 of Study Data Tabulation Model Metadata Submission Guidelines (SDTM-MSG) is now entering the Public Review period to solicit your feedback.

Description

The purpose of SDTM-MSG v2.0 is to provide guidance for preparing the components of the International Conference on Harmonisation (ICH) electronic Common Technical Document (eCTD) Module 5 (M5) Clinical Study Reports "sdtm" folder. This document and the associated example submission package illustrate the components recommended for electronic submission of SDTM data. These guidelines are based on Define-XML v2.1, SDTM v1.7/SDTMIG v3.3, and SDTM Terminology 2020-03-27.

- View the draft: SDTM Metadata Submission Guidelines Home
- Instructions for providing comments: Instructions for Reviewers
- Public Review closes **20 November 2020**.

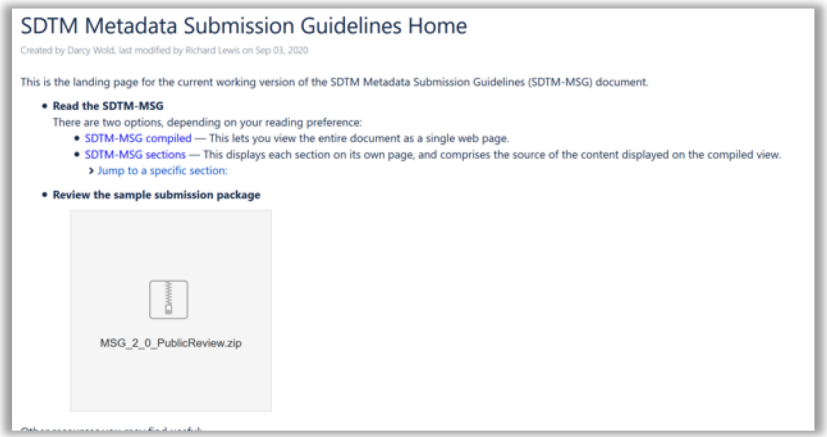
You will need to log in or register for the CDISC Wiki to provide comments.

- Register for the Wiki. If you already have an account on Wiki or JIRA, our issue-tracking system, simply log in to your account; Wiki and JIRA use the same login credentials. CDISC Wiki is a different login from www.cdisc.org.

Public review is a key quality step in our Standards Development Process. CDISC relies on your input to ensure neutral, consensus-based data standards are developed and adopted by a diverse global community interested in improving research processes and quality for the benefit of all.

Thank you for contributing your time and expertise.

MSG 2.0 Public Review



CDISC Wiki

SDTM Metadata Submission Guidelines

PAGE TREE

- Instructions for Reviewers
- SDTM-MSG compiled
- SDTM-MSG sections
  - 1 Introduction
    - **1.1 Purpose**
    - 1.2 Organization of this Document
  - 2 Define-XML Document
  - 3 Annotated CRF
    - 3.1 Basic Principles for Annotations
    - 3.2 Bookmarking CRFs/eCRFs
    - 3.3 Table of Contents for the Annotated
  - 4 Clinical Study Data Reviewers Guide
  - 5 Submission Datasets
  - Appendices

- Sample submission package included as ZIP file
  - "README" file included with additional information

MSG 2.0 Components for Public Review

MSG Document (on the wiki)

Define-XML document

- HTML versions included for those with problems displaying through stylesheet

Annotated CRF

Sample Data

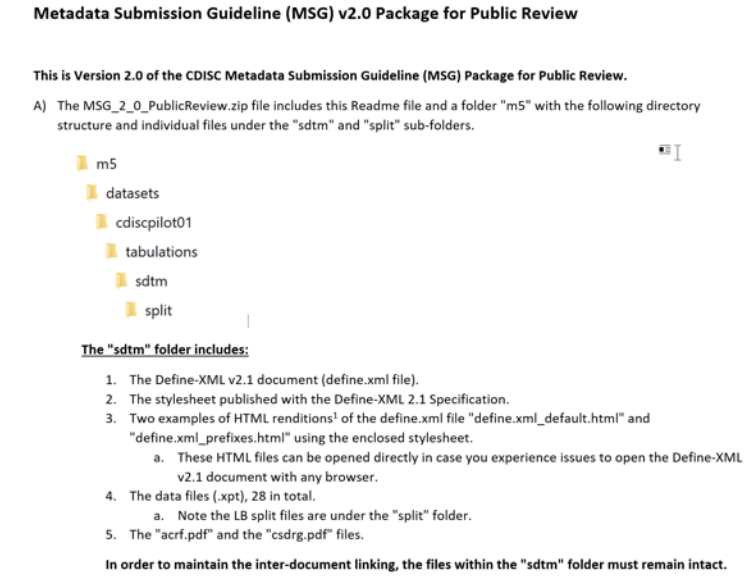
- Represented in two transport formats

Reviewer's Guide

- Based on PHUSE cSDRG template v1.3

Sample Stylesheet

ReadMe.pdf file



Metadata Submission Guideline (MSG) v2.0 Package for Public Review

This is Version 2.0 of the CDISC Metadata Submission Guideline (MSG) Package for Public Review.

A) The MSG_2_0_PublicReview.zip file includes this Readme file and a folder "m5" with the following directory structure and individual files under the "sdtm" and "split" sub-folders.

m5
datasets
cdiscpilot01
tabulations
sdtm
split

The "sdtm" folder includes:

1. The Define-XML v2.1 document (define.xml file).
2. The stylesheet published with the Define-XML 2.1 Specification.
3. Two examples of HTML renditions[1] of the define.xml file "define.xml_default.html" and "define.xml_prefixes.html" using the enclosed stylesheet.
   a. These HTML files can be opened directly in case you experience issues to open the Define-XML v2.1 document with any browser.
4. The data files (.xpt), 28 in total.
   a. Note the LB split files are under the "split" folder.
5. The "acrf.pdf" and the "csdrg.pdf" files.

In order to maintain the inter-document linking, the files within the "sdtm" folder must remain intact.

MSG 2.0 Components for Public Review – Cont.

MSG Document (on the wiki)

Define-XML document

- HTML versions included for those with problems displaying through stylesheet

Annotated CRF

Sample Data

- Represented in two transport formats

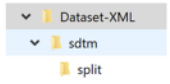
Reviewer's Guide

- Based on PHUSE cSDRG template v1.3

Sample Stylesheet

ReadMe.pdf file

B) The MSG_2_0_PublicReview_Dataset-XML.zip file includes a folder " Dataset-XML" with the following directory structure and individual files under the "sdtm" and "split" sub-folders. It contains an equivalent data package based on the published CDISC Dataset-XML v1.0 Standard[2].



Users may opt to use the CDISC Dataset-XML Standard for a range of business use cases. While regulatory agencies do not accept submissions in the Dataset-XML Standard at the time of publication of the MSG 2.0, that may change in the future. Refer to Appendix D Additional Disclaimers on the MSG 2.0 publication.

The "sdtm" folder includes:

1. The Define-XML v2.1 document (define.xml file), where the only differences compared to the Define-XML v2.1 document under the MSG_2_0_PublicReview.zip file are:
   a. References the data in Dataset-XML transport format instead of XPT format.
   b. The Define-XML document's context is "Other" instead of "Submission".
      i. "Other" Define-XML document context refers to any other data exchange format used; for example, between a sponsor and its CROs when the parties agree to use the Dataset-XML transport format.
      ii. If a data package based on the CDISC Dataset-XML Standard is used for submission, the Define-XML document context will need to be adjusted accordingly.
2. The stylesheet published with the Define-XML 2.1 Specification.
3. Two examples of HTML renditions of the define.xml file "define.xml_default.html" and "define.xml_prefixes.html" using the enclosed stylesheet.
   a. These HTML files can be opened directly in case you experience issues to open the Define-XML v2.1 document with any browser.
   b. Refer to footnote 1 on Page 1.
4. The data files (.xml), 28 in total.
   a. Note the LB split files (.xml) are under the "split" folder.
5. The "acrf.pdf" and the "csdrg.pdf" files.

In order to maintain the inter-document linking, the files within the "sdtm" folder must remain intact.

MSG 2.0 Public Review Logistics and Expectations

Commenting / Jira tickets

https://wiki.cdisc.org/display/MSG/Instructions+for+Reviewers

- Need specific reference of component or item for the feedback
  - In-line (in general for the MSG document on the wiki)
  - Directly in the Jira app (in general for the contents of the zip file)

Timelines

- 60 days ending 20-Nov-2020

Feedback Expectations

- Readability
- Usability
- Issues, inconsistencies

MSG 2.0 MSG_2_0_PublicReview.zip

Navigating the content and reporting a sample issue in Jira.

Thank You!

Public Review Webinar: Study Data Tabulation Model Metadata Submission Guidelines v2.0

Richard Lewis, SDTM Consultant, Data Standards Consulting Group (A division of Talent-Mine)
Mike Hamidi - Head of Data Science, CDISC
Marcelina Hungria, Managing Director/Owner, DIcore Group

Thursday, 24 SEP 2020
11:00AM – 12:30PM EDT

Audience Questions

Question for Richard: why not expending the scope to ADAM? many things in the guidance apply same to ADaM

Audience Questions

Richard: You mentioned the screen shot of ePRO should be included in the CRF, is that correct? Can we not include it in the CRF package?

Audience Questions

Is there any particular reason you have chosen Arial as a "standard" for acfr annotations?

Audience Questions

My company is submitting SDTMIG v3.2 and Define-XML v2.0. Should we use MSG v1, or v2 after it is finalized?

Audience Questions

Will all sponsors have the tools to be able to implement all of these proposed changes?

Audience Questions

What is we cannot implement some of these features due to our SOPs?

Audience Questions

For richard - latest define.xml specifications recommends using 'where' for conditions but MSG recommends 'when' e.g. RELREC when. Is there any specific reason for this?

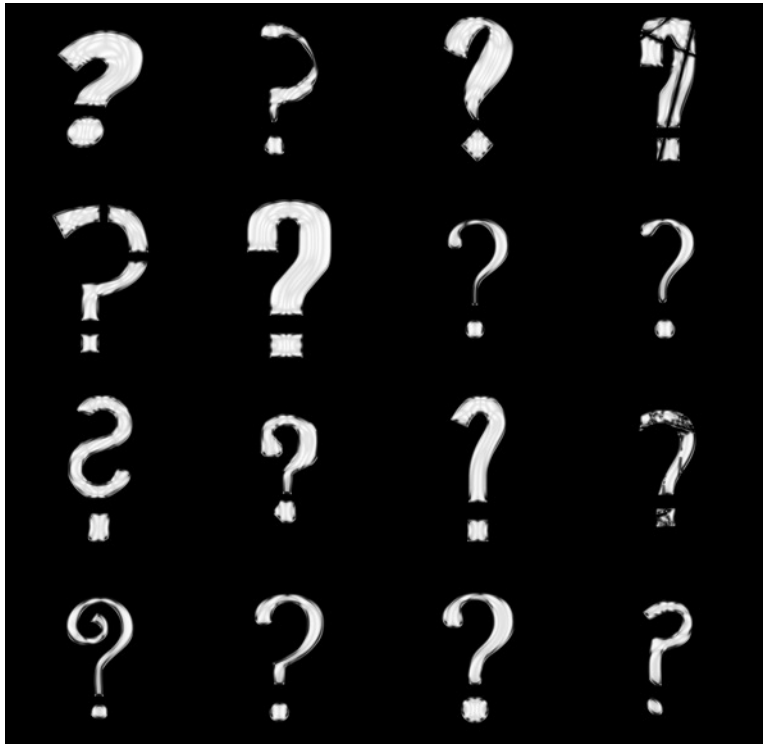
Audience Questions

can the font name for annotaitons be more flexible? e.g. would prefer a monspace font (courier) as can more easily programmitcally determine size of annotaiton box

Audience Questions

How do I get involved in the public review?

Audience Questions



Where/how can I leave a public comment on the guide?

Audience Questions

Can the list of domain colours for annotations be extended? have in the past had pages with more than 4 domain

Audience Questions

Do you have any recommendation in the MSG about whether annotations should be flattened or not (i.e. flattening annotations using PDF pre-flight feature so that annotations are not editable in submission).

Audience Questions

Can you please let us know how we can know if the define.xml is valued? Could you please specify the Validation procedure?

Can you please also explain how define.xml can be validated against data?

Audience Questions

Marcelina: Are define.xml_default.html and define.xml_prefixes.html needed in the submission package? I didn't understand what they are for.

Audience Questions

Audience Questions

Audience Questions

Upcoming Learning Opportunities

2020 US Virtual Interchange - 7-8 October

- Attendees

- Sponsors and Exhibitors

- Early Bird Rates Available through Friday, August 4

- Sign Up Now and Save!

Events Coming Your Way Soon!

- 2020 US Virtual Interchange
  - Registration Now Open!
  - Presenter, Sponsor and Exhibitor Opportunities Available**
  - Launching NEW Enhanced Virtual Conference Platform
    - 1:1 Virtual Networking
    - F2F Online and On-screen Meetings
    - 3D Exhibitor Experience

**Abstract Submissions Due 24 July

Newly added:
A full schedule of Virtual Courses including:

- ADaM Core Theory & Application
- CDISC for Newcomers
- CDASH Implementation
- CDASH Advanced Topics
- Controlled Terminology
- Define-XML
- ODM Implementation
- SDTM Theory & Application
- SDTM for Medical Devices
- SDTM Advanced Topics
- SEND Implementation

https://www.cdisc.org/events/interchanges

Preview Webinar – 1 September

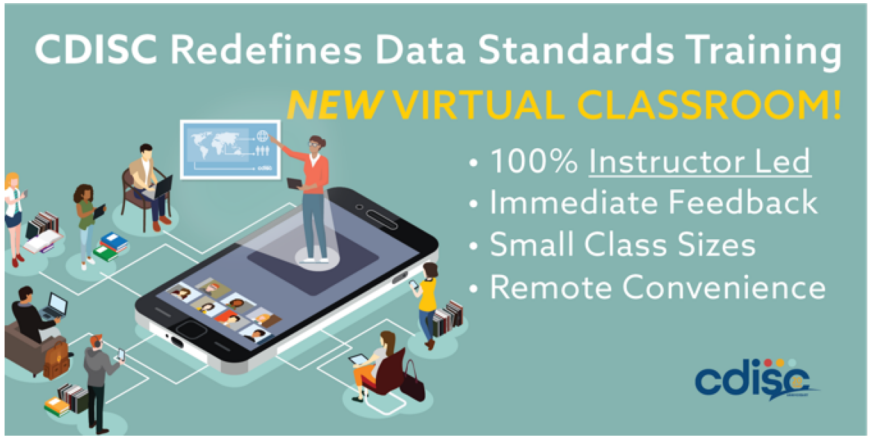
2020 Webinars

Date	Webinar Title
1 OCT	Public Review: TAUG for Crohn's Disease
13 OCT	Controlled Terminology Updates for Q4
20 OCT	Linking Data in SDTM
22 OCT	Introducing the Next Generation CDISC Library

Visit https://www.cdisc.org/events/education/webinars for information on additional Public Training events.

New Virtual Training Methods

- CDISC Provides Many Ways to Begin or Continue Growing Your Standards Knowledge.
  - Popular self-paced training plus new Blended Learning and Virtual Classroom settings.



Thank You!

Questions, comments, concerns? Email bklinke@cdisc.org

Questions about the Interchange specifically? Email events@cdisc.org

Don’t forget to fill out the feedback survey!