The TransCelerate / CDISC Digital Data Flow Project: Practical Electronic Study Designs

PHUSE EU Connect 2023 (DS02)

Dave Iberson-Hurst, CDISC Product Owner

7th November 2023, Version 3

Meet the Speaker

Dave Iberson-Hurst

Title: **Partner**

Organization: **d4k, Copenhagen**

Dave has over 40 years' experience across several industries with the last 20 years spent in the pharmaceutical industry combining his technology and software development experience with clinical data standards.

During this time, he has served as the CDISC CTO, worked on, and led, several CDISC teams, presented in many forums in Europe, the US, and elsewhere across the globe. He has worked closely with the FDA, EMA, HL7, ISO, and other standards organizations and was was a member of CDISC's Blue Ribbon commission. He is currently the CDISC Product Owner for the Digital Data Flow project.

He is a partner at data4knoweldge in Copenhagen and is focused on getting greater value and utility from clinical trial data.

Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*
- *On contract to CDISC for the DDF work*

cdisc

Abstract

Over the last two years CDISC, in collaboration with Transcelerate, have been working on the Digital Data Flow (DDF) initiative. This initiative aims to *"modernize clinical trials by enabling a digital workflow to allow for the automated creation of study assets and configuration of study systems to support clinical trial execution."*. The work is focused on the protocol and associated study designs and manifests itself in a new CDISC standard, the Unified Study Definitions Model (USDM), and an open-source implementation of the USDM known as the Study Definitions Repository (SDR).

Now coming to the end of the second phase, with the third phase about to commence, the DDF project delivers a new standard that allows for the digitization of study designs and the foundation of the digital protocol.

This presentation will detail:

- The work performed in phases one and two.
- The work planned as part of phase three.
- The use cases supported by the model.
- How the model/standard can enable protocol creation, automated data flow and interoperability between systems.
- How the model/standard can be deployed and implemented today.

Agenda

1. Introduction
2. Digital Data Flow – The Project
3. Use Cases
4. Phase Three: USDM Meets M11
5. Summary

Introduction

Complexity
Silos

© Data 4 Knowledge ApS

Digital Data Flow - The Project

https://www.transceleratebiopharmainc.com/assets/digital-data-flow-solutions/

TransCelerate Digital Data Flow (DDF) Ambition

Write Once, Read Many

TODAY: Document-based paradigm for protocol creation, interpretation, and transcription into consuming systems

TOMORROW: Digital paradigm for protocol creation, with fully automated data flow and interoperability between systems

©2023 TRANSCELERATE BIOPHARMA INC., ALL RIGHTS RESERVED.

TransCelerate BIOPHARMA INC.

CDISC DDF Deliverables

	PHASE ONE *July 2021 – July 2022*	PHASE TWO *Oct 2022 – June 2023*	PHASE THREE *July 2023– Apr 2024*
Unified Study Definitions Model (USDM)	✓	✓	✓
Application Programming Interface (API) Specification	✓	✓	✓
CDISC Controlled Terminology	✓	✓	✓
Reference Architecture Conformance Tests	✓		*Replaced by CORE rules*
Essential Users Stories	✓		
Architecture Principles	✓	*Still applicable*	*Still applicable*
Test Files		✓	✓
Implementation Guide		✓	✓

cdisc

The USDM Standard

CDISC Controlled Terminology

Provides further semantics, complementing the UML model. Includes the definition of classes and attributes along with the definition of value sets

	C174447	Study Arm
	C170984	Study Arm Name
	C93728	Study Arm Descriptio
	C188827	Study Arm Type
nDataOriginDescription	C188828	Study Arm Data Origin Description
riginType	C188829	Study Arm Data Origin Typ
el	CNEW	Study Arm Label
tudyEpoch	C71738	Study Epoch
me	C93825	Study Epoch Name
n	C93824	Study Epoch Descriptio
	C188830	Study Epoch Type
	CNEW	Study Epoch La
ch		

Logical Model

The UML logical model (a class diagram) that provides the basis for the USDM standard.

```
tudyArms": [
  {
    "id": "StudyArm_1",
    "name": "Placebo",
    "label": "",
    "description": "Placebo",
    "type": {
      "id": "Code_61",
      "code": "C174268",
      "codeSystem": "http://www.cdisc.org",
      "codeSystemVersion": "2022-12-16",
      "decode": "Placebo Comparator Arm"
    },
    "studyArmDataOriginDescription": "Data collec
    "dataOriginType": {
      "id": "Code_62",
      "code": "C188866",
      "codeSystem": "http://www.cdisc.org",
      "codeSystemVersion": "2022-12-16",
      "decode": "Data Generated Within Study"
    }
  },
  {
    "id": "StudyArm_2",
    "name": "Xanomeline Low Dose",
    "label": "",
    "description": "Active Substance",
    "type": {
      "id": "Code_63",
      "code": "C174267",
      "codeSystem": "http://www.cdisc
      "codeSystemVersion": "2022-12-
```


Examples

Example protocols implemented in the USDM with associated JSON files and visualisations

API Specification

Provides the means to exchange a single study between machines using a JSON API

Implementation Guide

Guidance on using the USDM model and ensuring conformance with the standard

cdisc

Use Cases

Use Cases: USDM with BCs allows for …

Data Capture	*Automate the setup of data capture systems, incl. RWE, and capture the data.*
SoA	*Use the study design to build the FHIR SoA message.*
Data Import	*Import data from a variety of sources. Can be re-exported thus allowing for conversion across versions.*
Common Protocol Template (CPT)	*Generation of the CPT from a study design.*
Data Decay	*Re-import data using the USDM as a framework to rebuild a study design & data using the SDTM Trial Design Domains.*
Scoring	*The "scoring" of a study for such purposes as site impact, subject impact, environmental impact etc.*
Feasibility	*The use of the design to determine study feasibility including subject recruitment. A study data template.*
CT Registry	*The provision of study information to a CT registry.*
FAIR Data	*The use of the design to aid Findability, Accessibility, Interoperability, and Reusability.*

CTMS, TMF …	*The provision of protocol information to downstream systems needing "study" information.*
Query	*Having multiple studies that have a common structure allows for data export and query across the set of studies.*
SDTM	*Automate the generation of SDTM datasets using the study design and BCs, including the "T" Domains.*

Plus many more …

USDM Adoption

USDM Data Mining Application

Sponsor working with CDISC to develop a retrospective use case

Sponsor is using the model for retrospective data ingestion

Retrospective has a lower risk as a first point of entry into using USDM

Study Build Application

Study Build & EDC Configuration

Retrospective Studies

Prospective Studies

The "footnote conundrum"

- *Retrospective study re-creation brings a few challenges*
- *We are not constrained by the "2D" paper world. USDM enables an "improved" reconstruction*
- *Sponsors need to consider their "philosophy", their approach to "reconstruction" of protocols*

Study Build Application

Population of USDM from old protocols

Open Study Builder
Study Build Application

Academic institution using the model as a component of a data import application

cdisc

Phase Three: USDM Meets M11

Next Steps – Phase Three

Slide from May 2023

USDM

SDR

1

- Baseline model for specifying a study in digital format
- Model supports use of a CRF link to specify which forms to use in EDC
- Handles simple study designs

- Consume digitized study specification from an upstream source e.g., study builder)
- Store, view and search study concepts
- Downstream EDC systems may pull study specification to aid in set-up

2

- Improved support for complex study designs with a fully specified digitized Schedule of Activities (SoA)
- Model supports the identification of the appropriate CRFs for data collection to enable automated, faster configuration via use of Biomedical Concepts
- Improved CPT alignment
- Initial 'T' Domain support

- Downstream vendors can readily consume the SoA from the SDR
- Sponsor system admins can perform a visual check that SoA data received from an upstream system displays an accurate, human-readable SoA table
- Opportunity to aggregate robust historical protocol information to support analytics to drive smart design and assess risk

3

Focus for Phase 3 is currently being determined. Current expectations are:

- Expand ability to handle increasingly complex studies
- ICH M11 & CPT alignment

cdisc

M11 Is …

ICH CLINICAL ELECTRONIC STRUCTURED HARMONISED PROTOCOL (CeSHarP)

https://www.ich.org/page/multidisciplinary-guidelines

ICH harmonisation for better health

INTERNATIONAL COUNCIL FOR HARMONISATION OF TECHNICAL REQUIREMENTS FOR PHARMACEUTICALS FOR HUMAN USE

ICH HARMONISED GUIDELINE

CLINICAL ELECTRONIC STRUCTURED HARMONISED PROTOCOL

(CESHARP)

M11

Draft version

Endorsed on 27 September 2022

Currently under public consultation

At Step 2 of the ICH Process, a consensus draft text or guideline, agreed by the appropriate ICH Expert Working Group, is transmitted by the ICH Assembly to the regulatory authorities of the ICH regions for internal and external consultation, according to national or regional procedures.

Provides background, purpose, and scope as a guideline

ICH harmonisation for better health

INTERNATIONAL COUNCIL FOR HARMONISATION OF TECHNICAL REQUIREMENTS FOR PHARMACEUTICALS FOR HUMAN USE

ICH HARMONISED GUIDELINE

CLINICAL ELECTRONIC STRUCTURED HARMONISED PROTOCOL

(CESHARP)

M11 TEMPLATE

Draft version

Endorsed on 27 September 2022

Currently under public consultation

At Step 2 of the ICH Process, a consensus draft text or guideline, agreed by the appropriate ICH Expert Working Group, is transmitted by the ICH Assembly to the regulatory authorities of the ICH regions for internal and external consultation, according to national or regional procedures.

Provides the written format for the 'Interventional Clinical Trial Protocol Template'

ICH harmonisation for better health

INTERNATIONAL COUNCIL FOR HARMONISATION OF TECHNICAL REQUIREMENTS FOR PHARMACEUTICALS FOR HUMAN USE

ICH HARMONISED GUIDELINE

CLINICAL ELECTRONIC STRUCTURED HARMONISED PROTOCOL

(CESHARP)

M11 TECHNICAL SPECIFICATION

Draft version

Endorsed on 27 September 2022

Currently under public consultation

At Step 2 of the ICH Process, a consensus draft text or guideline, agreed by the appropriate ICH Expert Working Group, is transmitted by the ICH Assembly to the regulatory authorities of the ICH regions for internal and external consultation, according to national or regional procedures.

Provides the technical representation aligned with the guideline and protocol template

cdisc

Idea for slide layout shamelessly stolen from Rebecca Baker

M11 Simple Example

Template Specification

Protocol Full Title:	[Protocol Full Title] The protocol should have a descriptive title that identifies the scientific aspects of the trial sufficiently to ensure it is immediately evident what the trial is investigating and on whom, and to allow retrieval from literature or internet searches.
Sponsor Confidentiality Statement:	[Sponsor Confidentiality Statement] Insert the Sponsor's confidentiality statement, if applicable, otherwise delete.
Protocol Number:	[Protocol Number] A unique alphanumeric identifier for the trial, designated by the Sponsor, is a standard part of trial data, and should be included for most trials.
Version:	[Version] An optional field for use by the Sponsor at their discretion.
Amendment Number:	[Amendment Number] Enter the amendment number. If this is the original instance of
Trial Phase:	[Trial Phase] [Description of Trial Phase Other] Acceptable entries are: "Early Phase 1", "Phase 1", "Phase 1/Phase 2", "Phase 2", "Phase 2/Phase 3", "Phase 3", "Phase 4",
Compound Number(s):	[Compound Number] Enter the Sponsor's unique identifier for investigational compound(s) in the trial. Add or delete additional fields as needed.
Compound Name(s):	[Nonproprietary Name]. [Proprietary Name]. [Additional Proprietary Name] Delete this line from the table if a nonproprietary name has not yet been assigned. Omit proprietary name fields if not yet established.
Trial Phase:	[Trial Phase] [Description of Trial Phase Other] Acceptable entries are: "Early Phase 1", "Phase 1", "Phase 1/Phase 2", "Phase 2", "Phase 2/Phase 3", "Phase 3", "Phase 4",

Technical Specification

Term (Variable)	Trial Phase
Data Type	Pick list
Topic, Value or Header	D
Definition	
User Guidance	For trials combining investigational drugs or vaccines with devices, classify according to the phase of drug development.
Conformance	Required
Cardinality	
Relationship content from ToC representing the protocol hierarchy	Title Page
Relationship (reference to high level conceptual model)	
Value	Early Phase 1 Phase 1 Phase 1/Phase 2 Phase 2 Phase 2/Phase 3 Phase 3 Phase 4 Other
Business rules	**Value Allowed:** yes **Relationship:** n/a **Concept:** Protocol short title
Duplicate field in other sections	

Controlled Terms

Template Specification

Protocol Full Title:	[Protocol Full Title] The protocol should have a descriptive title that identifies the scientific aspects of the trial sufficiently to ensure it is immediately evident what the trial is investigating and on whom, and to allow retrieval from literature or internet searches.
Sponsor Confidentiality Statement:	[Sponsor Confidentiality Statement] Insert the Sponsor's confidentiality statement, if applicable, otherwise delete.
Protocol Number:	[Protocol Number] A unique alphanumeric identifier for the trial, designated by the Sponsor, is a standard part of trial data, and should be included for most trials.
Version:	[Version] An optional field for use by the Sponsor at their discretion.
Amendment Number:	[Amendment Number] Enter the amendment number. If this is the original instance of
Trial Phase:	[Trial Phase] [Description of Trial Phase Other] Acceptable entries are: "Early Phase 1", "Phase 1", "Phase 1/Phase 2", "Phase 2", "Phase 2/Phase 3", "Phase 3", "Phase 4",
Compound Number(s):	[Compound Number] Enter the Sponsor's unique identifier for investigational compound(s) in the trial. Add or delete additional fields as needed.
Compound Name(s):	[Nonproprietary Name]. [Proprietary Name]. [Additional Proprietary Name] Delete this line from the table if a nonproprietary name has not yet been assigned. Omit proprietary name fields if not yet established.
Trial Phase:	[Trial Phase] [Description of Trial Phase Other] Acceptable entries are: "Early Phase 1", "Phase 1", "Phase 1/Phase 2", "Phase 2", "Phase 2/Phase 3", "Phase 3", "Phase 4",

Technical Specification

Term (Variable)	Trial Phase
Data Type	Pick list
Topic, Value or Header	D
Definition	
User Guidance	For trials combining investigational drugs or vaccines with devices, classify according to the phase of drug development.
Conformance	Required
Cardinality	
Relationship content from ToC representing the protocol hierarchy	Title Page
Relationship (reference to high level conceptual model)	
Value	Early Phase 1 Phase 1 Phase 1/Phase 2 Phase 2 Phase 2/Phase 3 Phase 3 Phase 4 Other
Business rules	**Value Allowed:** yes **Relationship:** n/a **Concept:** Protocol short title
Duplicate field in other sections	

CDISC CT

Trial Phase Response (C66737)

NOT APPLICABLE
PHASE 0 TRIAL
PHASE I TRIAL
PHASE I/II TRIAL
PHASE II TRIAL
PHASE II/III TRIAL
PHASE IIA TRIAL
PHASE IIB TRIAL
PHASE III TRIAL
PHASE IIIA TRIAL
PHASE IIIB TRIAL
PHASE IV TRIAL
PHASE V TRIAL

cdisc

Put Into Context … the USDM

Walkthrough

- *A StudyVersion has a studyPhase*
- *Linked to the AliasCode class via the relationship studyPhase*
- *AliasCode can store:*
  - *a single standard code (CDISC Trial Phase Response, C66737) using the Code class*
  - *And zero or more other alias codes (e.g. a sponsor phase code) using the Code class to provide flexibility*

Breadth versus Depth

Depth driven by individual use cases.

Some capability exists today, can be expanded incrementally or in one phase.

Shift of Focus

- Phases One & Two
  - Focused on the structured elements of the protocol, e.g. the Schedule of Activities (SoA)
  - The protocol document was an external entity into which the structured content could be exported
- Phase Three
  - Now contains structured and unstructured elements
  - The entire protocol document is held within the USDM
  - Allows for the protocol document to be generated from the model

M11 Template Example Document

- First attempt to create a protocol document from the USDM, both structured [non-narrative] and unstructured [narrative text] content
- Functionality has been added to the Excel test data tool
- More work is needed, this is very much a first draft

VERY DRAFT

Document doesn't look right? We'll help you out!

5 TRIAL POPULATION

5.1 Selection of Trial Population

5.2 Rationale for Trial Population

5.3 Inclusion Criteria

Patients may be included in the study only if they meet **all** the following criteria:

[1] Males and postmenopausal females at least 50 years of age.
[2] Diagnosis of probable AD as defined by National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA) guidelines (Attachment LZZT.7).
[3] MMSE score of 10 to 23.
[4] Hachinski Ischemic Scale score of ≤4 (Attachment LZZT.8).
[5] CNS imaging (CT scan or MRI of brain) compatible with AD within past 1 year. The following findings are incompatible with AD:

a. Large vessel strokes
   1. Any definite area of encephalomalacia consistent with ischemic necrosis in any cerebral artery territory.
   2. Large, confluent areas of encephalomalacia in parieto-occipital or frontal regions consistent with watershed infarcts. The above are exclusionary. Exceptions are made for small areas of cortical asymmetry which may represent a small cortical stroke or a focal area of atrophy provided there is no abnormal signal intensity in the immediately underlying parenchyma. Only one such questionable area allowed per scan, and size is restricted to ≤1cm in frontal/parietal/temporal cortices and ≤2 cm in occipital cortex.

b. Small vessel ischemia
   1. Lacunar infarct is defined as an area of abnormal intensity seen on CT scan or on both T1 and T2 weighted MRI images in the basal ganglia, thalamus or deep white matter which is ≤1 cm in maximal diameter. A maximum of one lacune is allowed per scan.
   2. Leukoariosis or leukoencephalopathy is regarded as an abnormality seen on T2 but not T1 weighted MRIs, or on CT. This is accepted if mild or moderate in extent, meaning involvement of less than 25% of cortical white matter.

USDM Summary

- Structured content along with the ability to hold unstructured content

Summary

ICH & The CDISC Project

Phase Three Timeline

January 2024
Phase 3 development sprints complete

February 2024
Phase 3 public review

April 2024
Version 3 USDM published

Dates may be adjusted to align with ICH M11 publication dates.

Slides taken from CDISC US Interchange 2023 Presentation

"ICH M11 Clinical Electronic Structured Harmonized Protocol"

Ron Fitzmartin, PhD, MBA
Center for Biologics Evaluation and Research
Food and Drug Administration

Summary

- We are “understanding” the complexity
- We can start to remove the silos and join the dots
- Digital Data Flow / Unified Study Definitions Model (DDF/USDM) fills an important gap
- USDM is but one building block, but an important one
- USDM alignment with ICH M11 will be an important step forward
- Can support various use cases, the prospective versus the retrospective
- We are only limited by our imagination

Thank You

Contacts:

Dave IH: diberson-hurst.external@cdisc.org
John Owen: jowen@cdisc.org

Link:

Github: https://github.com/cdisc-org/DDF-RA

CDSIC Team:

- Gerry Campion
- Drew Mills
- Erin Muhlbradt
- John Owen
- Berber Snoeijer
- Craig Zwickl