



PRAHEALTHSCIENCES

THE FUTURE OF CLINICAL DEVELOPMENT

ANNUAL CONFERENCE



Edinburgh 2017

**Digital Innovation
in Healthcare**
8th–11th October

Edinburgh International
Conference Centre



PRAHEALTHSCIENCES

Paper DS07 – PhUSE 2017

CDISC Transport Standards - A Glance

Giri Balasubramanian, PRA Health Sciences

Edwin Ponraj Thangarajan, PRA Health Sciences



- Paper Abstract
- CDISC Standards – Types
- Why Transport Standards?
- What are the CDISC Transport Standards?
- Types & Purpose of Transport Standards
- Conclusion

Paper Abstract

CDISC Transport Standards - A Glance



CDISC has published various standards over a period of years as part of transport standards to enable exchange of data conformant with CDISC foundation standards and their therapeutic area extensions. Such transport standards have gained immense usage across clinical trial life cycle, enabling regulatory to review them quickly and facilitate exchange of information across. This presentation would touch upon such standards such as,

- ❖ Clinical Trial Registry XML (CTR-XML)
- ❖ ODM-XML
- ❖ SDM-XML
- ❖ Define-XML
- ❖ Dataset-XML
- ❖ RDF
- ❖ LAB

giving details pertaining to the content it carries, structure, value it serves as a transport standard, regulatory implications, and roadmap of using such transport standards in clinical trial life cycle.



CDISC Standards

CDISC Standards



CDISC has two basic type of standards,

- one that holds the content, data, metadata and terminology
- the other which is called transport standard to move the data using XML technology.

XML is preferred over other technologies for data transfer because of:



Simplicity



Extensibility



Interoperability



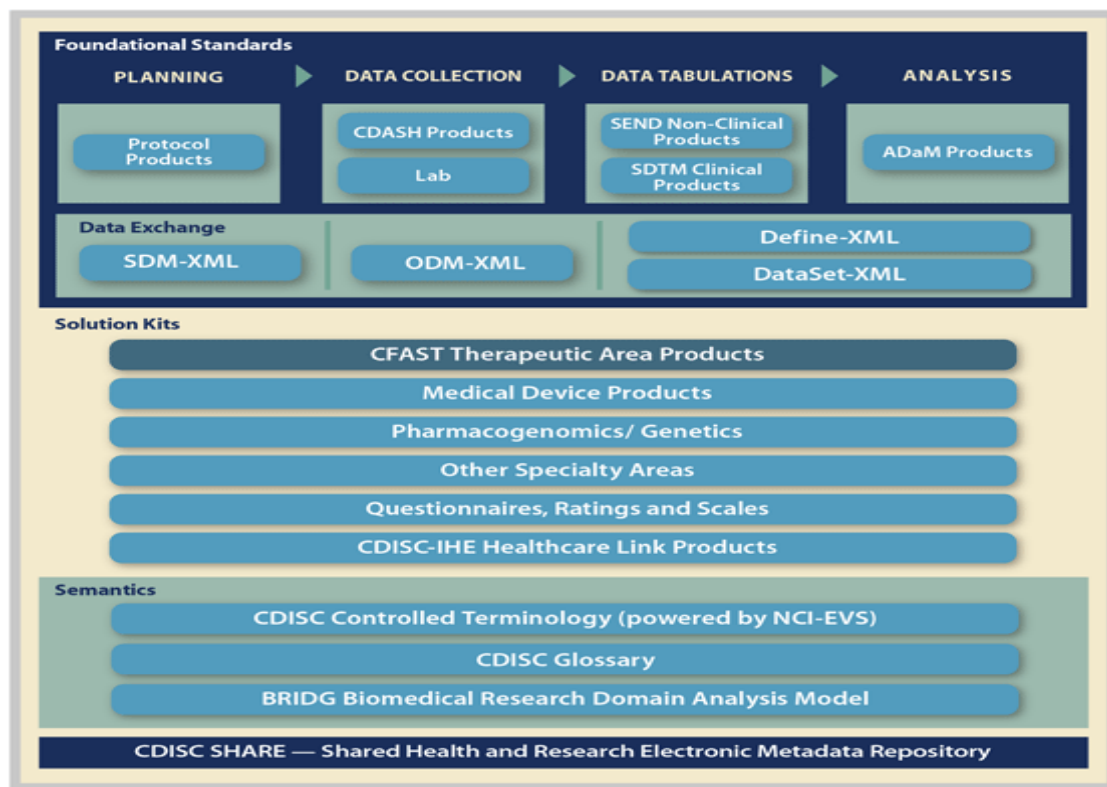
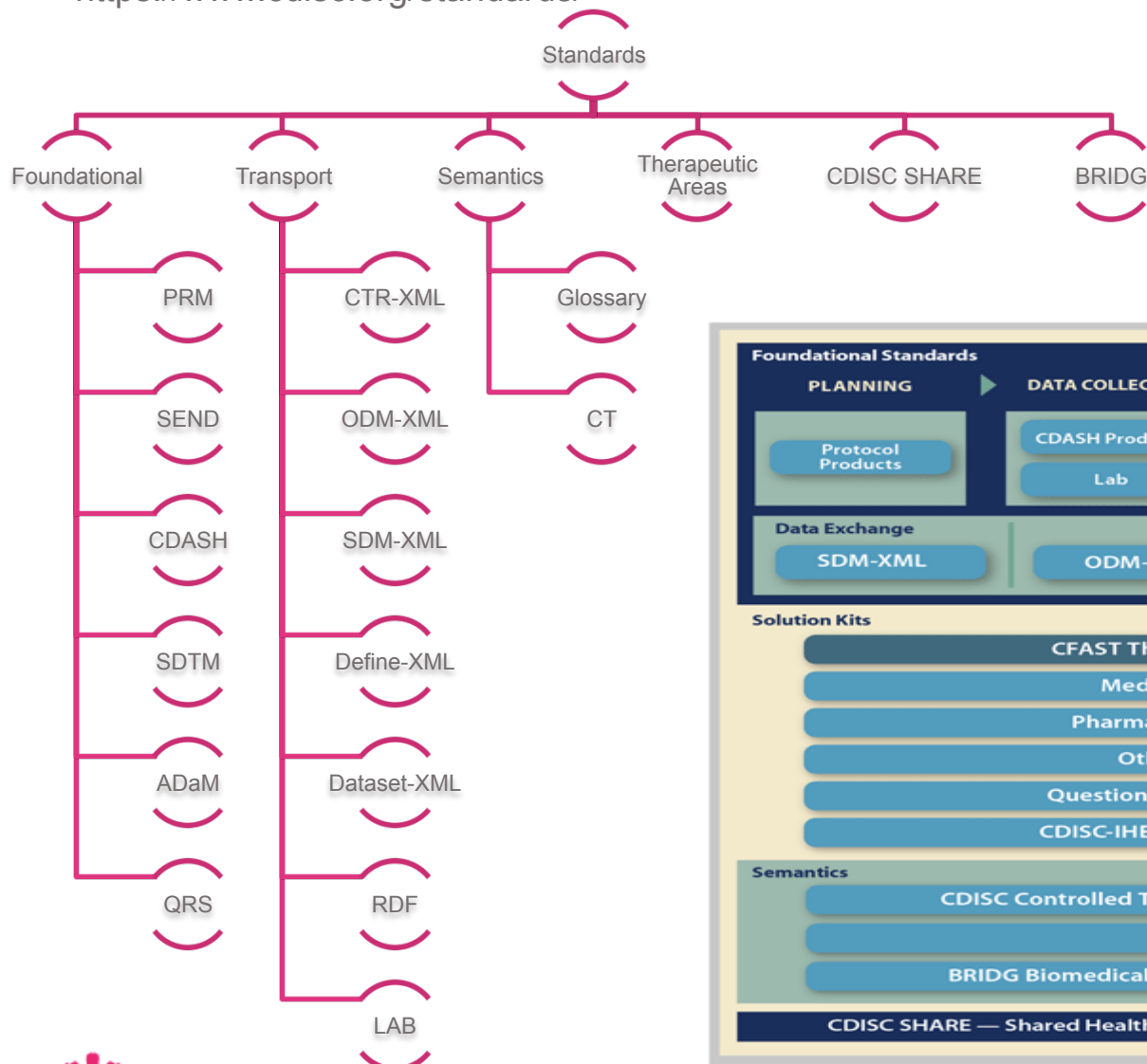
Openness



Experience

CDISC Standards

<https://www.cdisc.org/standards/>



THE SOLUTION: FORMAT DATA TO ENABLE SHARING

CDISC Transport Standards



PRAHEALTHSCIENCES



- Seamless movement of trial data across systems in a Clinical Trial Life Cycle
- Maintains traceability of information
- Replacement of age old submission of trial data to support eSubmissions
- Improves Operational Efficiency & Effectiveness
- Encourages automation of processes and even moving towards Robotic Process Automation of entire life cycle

CDISC Transport Standards

Why Transport Standards?



The FDA has stated that, “**improving the efficiency and effectiveness of medical product development is a national priority**”. Regulatory electronic submissions have grown more complex with the average submission now a staggering 3.4 million pages, an increase of 1,423% since 2005. With this scale, inefficiencies in the clinical research data lifecycle add considerable time and expense to new medical product development. Increasing efficiency requires that the networked organizations participating in clinical development exchange data seamlessly. The 2014 CDISC business case claims that using CDISC standards from the beginning of the process can save approximately \$180 million per submission.

The very question of improving efficiency & effectiveness led to development and adoption of transport standards widely by Pharmaceutical companies and Software Vendors. Highlights of why new set of transport standards were discussed early on in November 2012 is listed below:

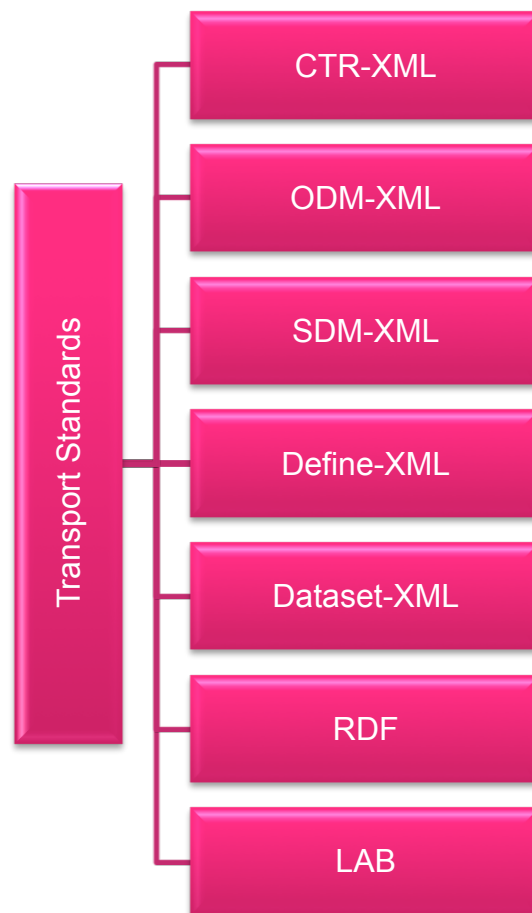
The SAS XPORT transport format should be replaced by a more modern data exchange standard for electronic regulatory submissions to FDA based on current prevailing XML technology

The choice of transport standards for study data should capitalize on existing knowledge and investment within the global bio-pharmaceutical industry

The choice of transport standards should ensure that commonly used data structures, specifically domain datasets and analysis files and their associated metadata, can be accurately exchanged, utilized and reproduced.

CDISC Transport Standards

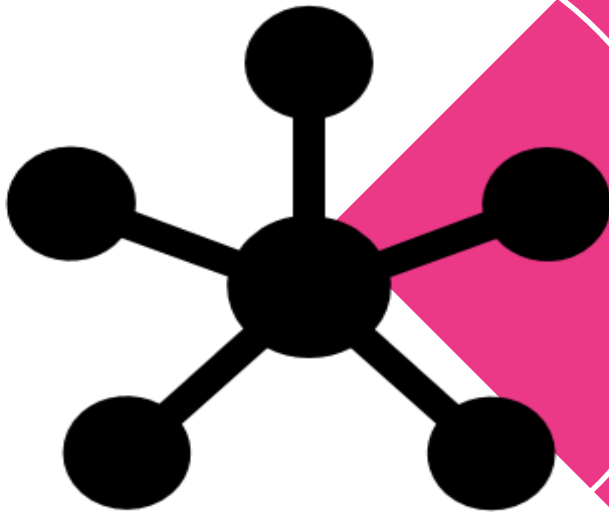
<https://www.cdisc.org/standards/transport>



CDISC Transport Standards enable the exchange of data conformant with CDISC Foundational Standards and their Therapeutic Area extensions.

CLINICAL TRIAL REGISTRY XML (CTR-XML)

<https://www.cdisc.org/standards/transport/ctr-xml>



- Standard that facilitate submission of clinical research for multiple, global clinical trial registries; specifically the World Health Organization (WHO), European Medicines Agency (EMA) and to ClinicalTrials.gov from a single file
- Built over reuse of existing transport standard by extending ODM-XML and including SDM-XML
- Generalized to populate any trial registry with structured content from a clinical trial sponsor's systems

Future roadmap:

- Extend the standard, creating an underlying structured protocol standard, results summaries and IDMP compliance to further increase the richness, reliability and traceability of registry information
- Project Kicked Off as CTR2 in Sept '17 - The CTR2 (Clinical Trial Registration) Project will extend and enrich CTR-XML by including summary results, extending Protocol elements, and harmonizing the CDISC Foundational Standards with the ISO IDMP Standards [<https://www.cdisc.org/ctr2-project>]

CLINICAL TRIAL REGISTRY XML (CTR-XML)



The minimum amount of trial information that must appear in a register in order for a given trial to be considered fully registered. There are currently 20 items in the WHO Trial Registration Data Set. It is sometimes referred to as the TRDS.

- A clinical trials registry is an official platform and catalog for registering a clinical trial. Clinical trials are conducted to allow safety and efficacy data to be collected for health interventions (e.g., drugs, diagnostics, devices, therapy protocols).
- A clinical trials register is the formal record of an internationally agreed minimum amount of information about a clinical trial. This record is usually stored in and managed using a database. A clinical trials registry is the entity that houses the register, and is responsible for ensuring the completeness and accuracy of the information it contains, and that the registered information is used to inform health care decision making. A clinical trials registry is more than its database.
- At its core, the CTR standard is inspired by the International Committee of Medical Journal Editors (ICMJE), and is based upon the 20-item WHO Trial Registration Data Set as well as EudraCT specific extensions. The CTR standard maximizes the re-use of existing CDISC transport standards by extending the Operational Data Model (ODM-XML) and including Study/Trial Design Model (SDM-XML) content. This makes the standard a more general solution that any trial registry could potentially use as the means to populate their registry with structured content from a clinical trial sponsor's system.

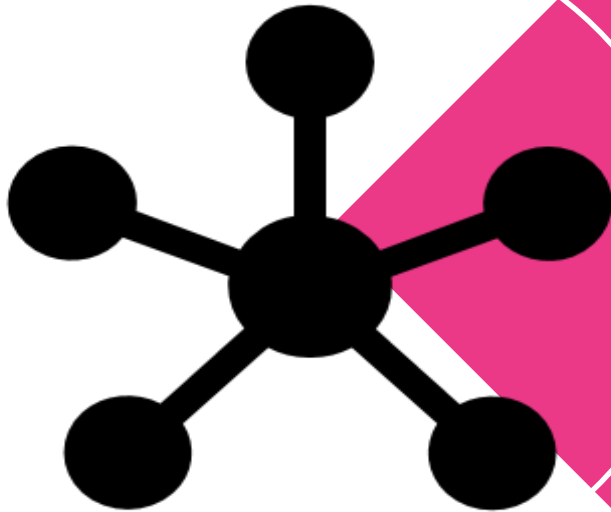
#	WHO Trial Registration Data Set
1	Primary Registry and Trial Identifying Number
2	Date of Registration in Primary Registry
3	Secondary Identifying Numbers
4	Source(s) of Monetary or Material Support
5	Primary Sponsor
6	Secondary Sponsor(s)
7	Contact for Public Queries
8	Contact for Scientific Queries
9	Public Title
10	Scientific Title
11	Countries of Recruitment
12	Health Condition(s) or Problem(s) Studied
13	Intervention(s)
14	Key Inclusion and Exclusion Criteria
15	Study Type
16	Date of First Enrollment
17	Target Sample Size
18	Recruitment Status
19	Primary Outcome(s)
20	Key Secondary Outcomes

CTR-XML uses elements and attributes from four namespaces. The following principles were applied in deciding which namespaces to use for each piece of information to be submitted:

- CTR information elements that are part of the ODM are used as is.
- CTR information elements that are part of SDM-XML are used as is. In a few cases where a SDM-XML element is conceptually the same as a CTR element but is missing a sub-component, the SDM-XML element has been extended.
- CTR information elements that are not part of the ODM or SDM-XML, but are defined in the EudraCT XML schemas, are used without extensions.
- CTR information elements that are not part of the ODM, SDM-XML or the EudraCT XML schemas are defined as part of the CTR-XML extension.

OPERATIONAL DATA MODEL (ODM)-XML

<https://www.cdisc.org/standards/transport/odm>



- ODM-XML is a vendor-neutral, platform-independent format for exchanging and archiving clinical and translational research data, along with their associated metadata, administrative data, reference data, and audit information.

- ODM-XML facilitates the regulatory-compliant acquisition, archival and exchange of metadata and data.

- It has become the language of choice for representing case report form content in many electronic data capture (EDC) tools.

- ODM was not originally developed based on an existing clinical research or healthcare data model, but instead was designed using a bottom-up approach to meet the established
 - data interchange,
 - archival, and
 - audit trail requirements.
- The first production version of ODM was published in October 2000 and was demonstrated in two connections on events in 2001. The current ODM version, v1.3.2, was published in December of 2013. ODM, now based on XML schema, remains under active development by the CDISC XML Technologies Team, and while the original ODM requirements remain highly relevant, use of the standard has extended well beyond the original design.

OPERATIONAL DATA MODEL (ODM)-XML



Characteristics of ODM

- Efficient communication of a clinical study protocol and case report forms during all stages of a human clinical study is important for many stakeholders.
- An electronic and structured study representation format that can be used throughout the whole study life-span can improve such communication and potentially lower total study costs.
- ODM's initial objective was exchange of case report forms data but it is increasingly utilized in other contexts.
- The initial focus was on a general, vendor neutral structure and syntax; industry level data models and semantics were given little consideration.
- The ODM standard plays a key role in clinical research informatics, including areas such as data exchange, archival, U.S. Food and Drug Administration (FDA) submission, and interoperability with healthcare data. Within the highly data-centric domain of clinical research, the XML-based ODM is the standard exchange format for case report form (CRF) data and metadata.

OPERATIONAL DATA MODEL (ODM)-XML

Characteristics of ODM



- ODM-XML v1.3.2 is the most current version of the standard.
- Many CDISC standards have been developed by extending ODM-XML including: Define-XML, SDM-XML, Dataset-XML, CTR-XML and CT-XML.
- ODM-XML provides a common base structure for these standard extensions easing the learning curve and implementation complexity



OPERATIONAL DATA MODEL (ODM)-XML

<https://www.cdisc.org/standards/transport/odm>



The CDISC foundational standards covered by ODM, and standardized extensions such as Clinical Data Acquisition Standards Harmonization (CDASH) that describes the basic data collection fields for domains, the Study Data Tabulation Model (SDTM) that describes a standard structure for study data tabulations, and the Analysis Data Model (ADaM) that describes metadata models and examples for analysis datasets.

Foundational Standards

PLANNING

Protocol
Products

DATA COLLECTION

CDASH Products

Lab

DATA TABULATIONS

SEND Non-Clinical
Products

SDTM Clinical
Products

ANALYSIS

ADaM Products

Data Exchange

SDM-XML

ODM-XML

Define-XML

DataSet-XML



OPERATIONAL DATA MODEL (ODM)-XML

<https://www.cdisc.org/standards/transport/odm>



CDISC's Operational Data Model (ODM) is the most versatile standard of CDISC's suite. It is primarily an XML-based transport format defined via an XML Schema. Additionally, ODM has a powerful underlying data model that represents,

all clinical facts about subjects that were acquired during the study plus audit log entries

administrative information like user accounts and electronic signatures

reference data that help to interpret the clinical data

support for versioning

is extensible by custom vendor extensions

OPERATIONAL DATA MODEL (ODM)-XML

Reasons for usage



The CDISC Operational Data Model (ODM), which has been in production use for more than ten years, is an ideal choice as a new study data exchange standard for the following reasons:

ODM can streamline the clinical development process by supporting metadata-driven data transport end-to-end across the entire clinical research lifecycle, with traceability from protocol through analysis.

ODM is fully compliant with regulatory guidance and 21 CFR Part 11, including audit trail and electronic signatures.

ODM is fully compatible with current metadata submission standards, and is the basis for the CDISC define.xml standard already accepted by FDA

ODM is supported by NCI EVS as an exchange format for CDISC controlled terminology.

ODM has been successfully used in conjunction with HL7 CDA formatted data from Electronic Healthcare Record systems to support research under an HHS sponsored interoperability specification.

ODM is fully compliant with regulatory guidance and 21 CFR Part 11, including audit trail and electronic signatures.

ODM is already widely understood and used extensively for global clinical research, and can be deployed for submissions without significant added financial burden on industry.

ODM accurately represents and easily reproduces tabular dataset structures, including those structured according to the CDISC Study Design Model, CDASH, SDTM, SEND and ADaM standards that are already widely used in industry and at the FDA.

ODM is already supported by major technology providers of clinical data information systems used for regulated clinical research.

ODM can represent more complex relationships between data events recorded per the research protocol.

ODM can be easily and rapidly extended through the CDISC standards development process to address emerging new requirements as they arise.

OPERATIONAL DATA MODEL (ODM)-XML

Entities and Elements



The ODM model assumes that a study's clinical data will consist of several kinds of entities. These include subjects, study events, forms, item groups, items, and annotations

Entities and Elements	Description
item	An item is an individual clinical data item, such as a single systolic blood pressure reading. Items are collected together into item groups
item group	An item group is a closely related set of items that are generally analyzed together. (Item groups are sometimes referred to as "records" and are associated with "panels" or "tables".) Item groups are aggregated into forms.
form	A form is analogous to a page in a paper CRF book or electronic CRF screen. A form generally collects a set of logically and temporally related information. A series of forms is collected as part of a study event.
study event	A study event is a reusable package of forms usually corresponding to a study data-collection event.
subject	A subject is a patient participating in the study.
annotation	An annotation is a comment applied to a subject, study event, form, item group, or item. Annotations can also be applied to pairs of entities.
StudyEventDef	A StudyEventDef describes a particular type of study event (mostly by listing the types of forms it can contain).
FormDef	A FormDef describes a particular type of form.
ItemGroupDef	An ItemGroupDef describes a particular type of item group.
ItemDef	An ItemDef describes a particular type of item.

Example of prolog and top level ODM element:

```
<?xml version="1.0" encoding="UTF-8"?>
<ODM xmlns="http://www.cdisc.org/ns/odm/v1.3"
      xmlns:ds="http://www.w3.org/2000/09/xmldsig#"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

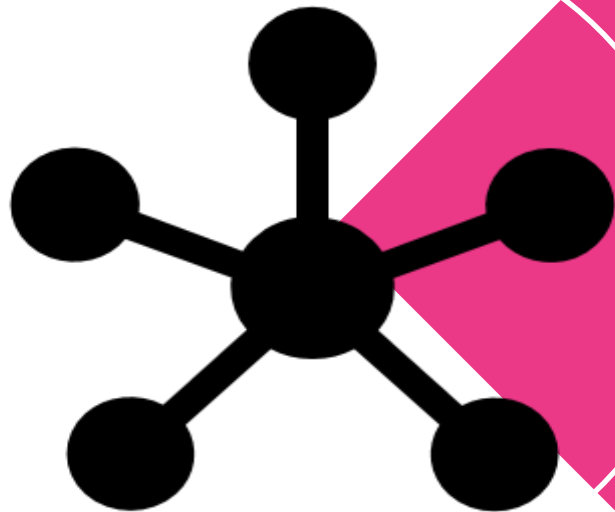
      xsi:schemaLocation="http://www.cdisc.org/ns/odm/v1.3 ODM1-3-2.xsd"

      ODMVersion="1.3.2"
      FileOID="000-00-0000"
      FileType="Transactional"
      Description="Sample ItemData Extension"

      AsOfDateTime="2013-02-04T07:57:00"
      CreationDateTime="2013-02-06T10:30:00">
```

STUDY/TRIAL DESIGN MODEL-XML

<https://www.cdisc.org/standards/transport/sdm-xml>



- Study/Trial Design Model in XML (SDM-XML) is an extension of ODM-XML and allows organizations to provide rigorous, machine-readable, interchangeable descriptions of the designs of their clinical studies, including treatment plans, eligibility and times and events.
- SDM-XML defines three key sub-modules - Structure, Workflow, and Timing
- Permitting various levels of detail in any representation of a clinical study's design.

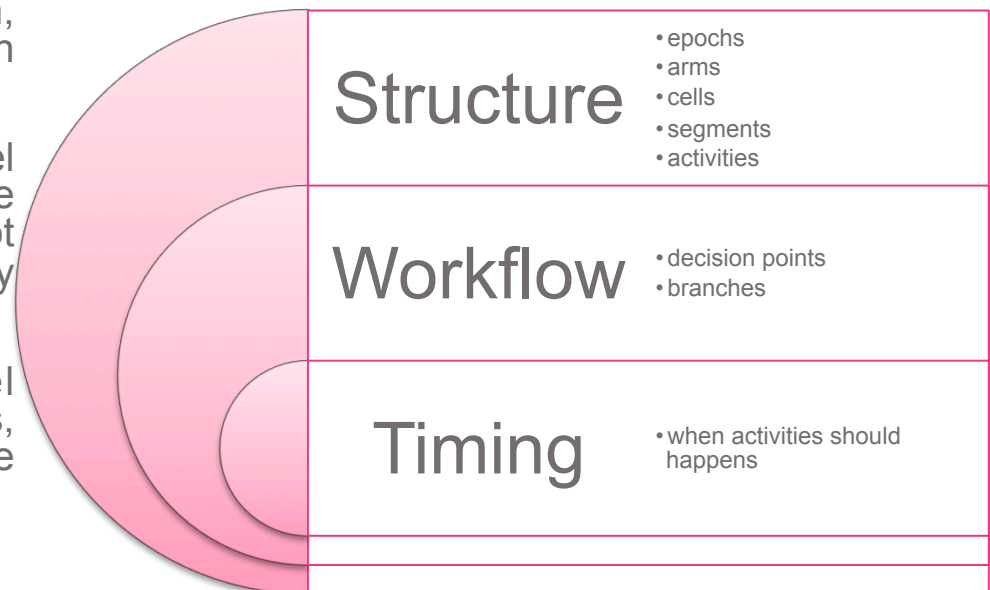
STUDY/TRIAL DESIGN MODEL-XML

<https://www.cdisc.org/standards/transport/sdm-xml>



- The clinical research study protocol is the plan that describes the study's objectives, methodology, statistical considerations, and the organization of the study.
- This plan includes the design of the study, which includes the arm descriptions, the schedule of activities, the eligibility criteria and summary information. Several CDISC standards represent aspects of the study design, but do not specify the study design completely.
- For instance, the Operational Data Model (ODM) represents the metadata for the data collected in the study, but does not describe the planned timing of the study events.
- The Study Data Tabulation Model (SDTM) includes trial design datasets, but only pertains to the visits, which are only part of the activity schedule.

- As for the Protocol Representation Model (PRM), it is a conceptual model that includes the study design, but has no specification details.
- The CDISC Study Design Model (SDM) has been developed to specify the study design. It extends the core ODM and consists of the following sub-components that model the design of the study, not its execution.
- The SDM is modelled in XML.



STUDY/TRIAL DESIGN MODEL-XML

Structural Elements



Structural elements are comprised of the “building blocks” of a study design: objects such as Epochs, Cells, Arms and Segments, as well as Activities. These are the objects that can act as nodes in a study design workflow, or as objects between which timing constraints may be applied.

Structure Overview

	Screening epoch	Treatment epoch	Follow-up epoch
Placebo arm	Screening Cell	Treatment Cell (Blinded)	Follow-up Cell (Blinded)
Low dose arm			
High dose arm			

Epochs

Name	Description	Entry Criteria	Exit Criteria
Screening epoch	Screening Epoch		
Treatment epoch	Treatment Epoch	- Entry criterion for treatment epoch (Entry condition for treatment epoch) The inclusion/exclusion criteria for the study are applied here.	Exit criterion for treatment epoch (Exit condition for treatment epoch)
Follow-up epoch	Follow-up Epoch		

Arms

Name	Description
Placebo arm	Placebo arm: a placebo is being administered
Low dose arm	Low-dose arm: 50 cm2 TTS Formulation E, 54 mg xanomeline
High dose arm	High-dose arm: 75 cm2 TTS Formulation E, 81 mg xanomeline

Cells

Name	Description	Blinded/Unblinded	Epoch	Arm(s)	Segment(s)	Entry Criteria	Exit Criteria
Screening Cell		Unblinded	Screening epoch	Placebo arm Low dose arm High dose arm	Screening segment		
Treatment Cell		Blinded	Treatment epoch	Placebo arm Low dose arm High dose arm	Treatment segment		
Follow-up Cell		Blinded	Follow-up epoch	Placebo arm Low dose arm High dose arm	Follow-up segment		

STUDY/TRIAL DESIGN MODEL-XML

Workflow



Study workflows are defined using a set of constructs that make it possible for a study designer to specify possible study participant paths through a study.

Workflow is specified in a section of XML distinct from that of the structural elements. However, workflow objects commonly reference objects defined in the Structure section of the document. This separation of concerns allows the potential for different workflow representations to be applied to the same set of structural elements.

Workflow Transitions

Name	Description	Source Activity	Target Activity	Condition	Transition Name	Transition Description	Timing Constraint
Trial start transition		trial start activity	informed consent activity	Default	(default) transition to informed consent		
Transition from informed consent activity		informed consent activity	patient number assignment activity	Informed Consent condition	conditional transition from informed consent to patient number assigned		• Transition timing constraint 1 name
			trial finish activity	Default	default transition from informed consent to study finish		
transition to activity taking demographics, education and habits		patient number assignment activity	Demographics, Education and Habits data collection activity	Default	default transition to transition to activity taking demographics, education and habits		
Transition to mini-mental state test activity		Demographics, Education and Habits data collection activity	Mini-mental state activity	Default	Default transition to mini-mental state test activity		
transition to Hachinski test		Mini-mental state activity	Hachinski test Activity	Condition that subject has MMSE score of 10 to 23	conditional transition to Hachinski test		
			trial finish activity	Default	Transition to Trial Finish as MMS leads to exclusion		
Transition between Hachinski test and Medical History		Hachinski test Activity	medical history activity	Condition Hachinski Ischemic Scale score of <= 4	Conditional Transition between Hachinski test and Medical History		
			trial finish activity	Default	Default transition to Trial Finish due to Hachinski score exclusion		
Transitions from Medical History Activity		medical history activity	vital signs activity	Default	Transition to Vital Signs Activity		
Trial finish due to high fever		assess high fever activity	trial finish activity	Default	Trial finish due to high fever		• at high fever, go to trial finish within 1 day

Triggers

Name	Description	Structural Element	Structural Element Type	Target Activity	Condition	Transition Name	Transition Description	Timing Constraint
Fever trigger		Treatment epoch	Epoch	assess high fever activity	Condition that subject has high fever (>38.5C / 101.3F)	conditional transition when subject has high fever		
				assess fever activity	Default	default transition from when subject has fever		• Transition timing constraint 2
Headache trigger		Treatment epoch	Epoch	Assess headache	Default	Default transition upon headache		

STUDY/TRIAL DESIGN MODEL-XML



Timing

Timing constraints, like structural definitions and workflow definitions, are declared in their own sub-section of an SDM-XML document, within an element named Timing. Timing constraints may apply either to activities or to workflow transitions. Note, however, that this relationship is one-way – elements declared in the structural or workflow areas of the document never reference timing elements.

Workflow Transition Timing Constraints

Name	Description	Transition	Type	Target	Granularity	PreWindow	PostWindow	Base Subsequent Schedule On
Transition timing constraint 1 name	english description	conditional transition from informed consent to patient number assigned	StartToStart	PT24H				Planned
Transition timing constraint 2	Upon fever, react within one hour	default transition from when subject has fever	FinishToStart	PT1H				Planned
at high fever, go to trial finish within 1 day		Trial finish due to high fever	StartToStart	P1D				Planned

Absolute Timing Constraints

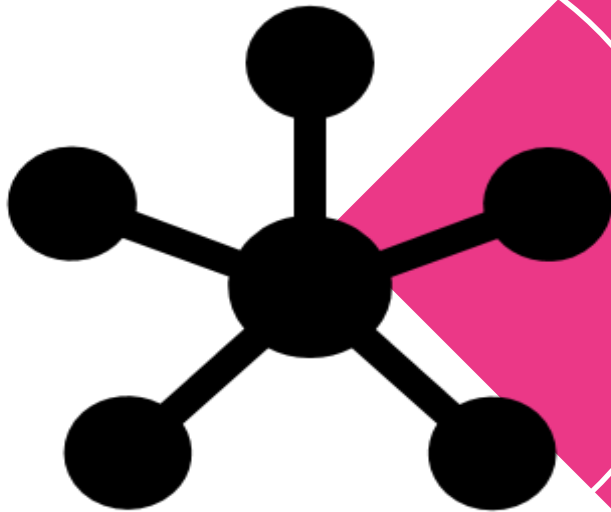
Name	Description	Target	Granularity	PreWindow	PostWindow
Absolute timing constraint on informed consent activity		2009-09-20T11:07:23-05:00			

Relative Timing Constraints

Name	Description	Predecessor Activity	Successor Activity	Type	Target	Granularity	PreWindow	PostWindow	Base Subsequent Schedule On
relative timing constraint between activity informed consent and activity patient number assignment		informed consent activity	patient number assignment activity	FinishToStart	P7D				Planned
Timing constraint between ambulatory ECG placement and removal	All patients will undergo a 24-hour Ambulatory ECG at Visit 2 (prior to the initiation of study medication). Although every effort will be made to obtain the entire 24-hour ambulatory ECG recording, this may not always be feasible because of patient behavior or technical difficulties. The minimal recording period for an ambulatory ECG to be considered interpretable will be 8 hours, of which at least 3 hours must be sleep.	ambulatory ECG placement activity	ambulatory ECG removal activity	StartToStart	PT24H		PT16H		Planned
Timing constraint for treatment visit 1 (visit 4)	For the first 8 weeks of treatment, patients will be assessed at clinic visits every 2 weeks.	patient randomization activity	Start of treatment visit 1 (visit 4)	StartToStart	P14D				Planned

DEFINE-XML

<https://www.cdisc.org/standards/transport/define-xml>



- Define-XML transmits metadata that describes any tabular dataset structure.
- When used with the CDISC content 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 is required by the United States Food and Drug Administration (FDA) and the Japanese Pharmaceuticals and Medical Devices Agency (PMDA) for every study in each electronic submission to inform the regulators which datasets, variables, controlled terms, and other specified metadata were used.

Future roadmap:

- Define-XML v2.1 updates the v2.0 standard originally published in 2013 with a number of improvements and new features.

DEFINE-XML

<https://www.cdisc.org/standards/transport/define-xml>



- FDA CDER and CBER have indicated support for v1.0 and v2.0 of Define-XML in the FDA Data Standards Catalog. The FDA has also announced the end of support for Define-XML v1.0 for studies that start 12 months after March 15, 2017.
- Define-XML v2.0 represents a significant update to Define-XML v1.0 in response to implementation experience with v1.0, the evolution of the SDTM, SEND and ADaM standards and best practices by SDTM and ADaM metadata experts. Key enhancements include:

Support for CDISC Controlled Terminology

Flexible definition of Value Level Metadata

Enhanced documentation of data origin or source

Improved support for ADaM metadata

Improved handling of comments

DEFINE-XML

Document Structure



The below shows the XML that would comprise the minimal structure of any ODM 1.3.2 document that contains a Define-XML document. It illustrates a valid Define-XML document header and the gray box illustrates the set of elements that comprise this standard in the order in which they should appear in a valid Define-XML file.

```
<?xml version="1.0" encoding="UTF-8"?>
<ODM xmlns="http://www.cdisc.org/ns/odm/v1.3"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:def="http://www.cdisc.org/ns/def/v2.0"
  FileType="Snapshot" ODMVersion="1.3.2"
  FileOID="Studydisc01-Define-XML_2.0.0"
  CreationDateTime="2012-06-21T11:07:23-05:00"
  Originator="CDISC XML Technologies Team">
  <Study OID="cdisc01">
    <GlobalVariables>
      <StudyName>CDISC01</StudyName>
      <StudyDescription>CDISC Test Study</StudyDescription>
      <ProtocolName>CDISC 01</ProtocolName>
    </GlobalVariables>
    <MetaDataVersion OID="MDV.CDISC01.SDTMIG.3.1.2.SDTM.1.2">
      Name="Study CDISC01, Data Definitions"
      Description="Study CDISC01, Data Definitions"
      def:DefineVersion="2.0.0"
      def:StandardName="CDISC SDTM"
      def:StandardVersion="3.1.2">
```

```
< Annotated Case Report Forms (def:AnnotatedCRF) >
< Supplemental Data Definitions (def:SupplementalDoc) >
< Value Level Metadata (def:ValueListDef) >
< Where Clause Definitions (def:WhereClauseDef) >
< Domain Level Metadata (ItemGroupDef) >
< Variable Level Metadata (ItemDef) >
< Controlled Terminology Metadata (CodeList) >
< Computational Algorithms (MethodDef) >
< Comments (def:CommentDef) >
< Referenced Documents (def:leaf) >
```

```
</MetaDataVersion>
</Study>
</ODM>
```

SDTM-IG 3.1.2

Annotated Case Report Form
Reviewers Guide
Complex Algorithms
Tabulation Datasets
Value Level Metadata
Controlled Terminology
Computational Algorithms
Comments

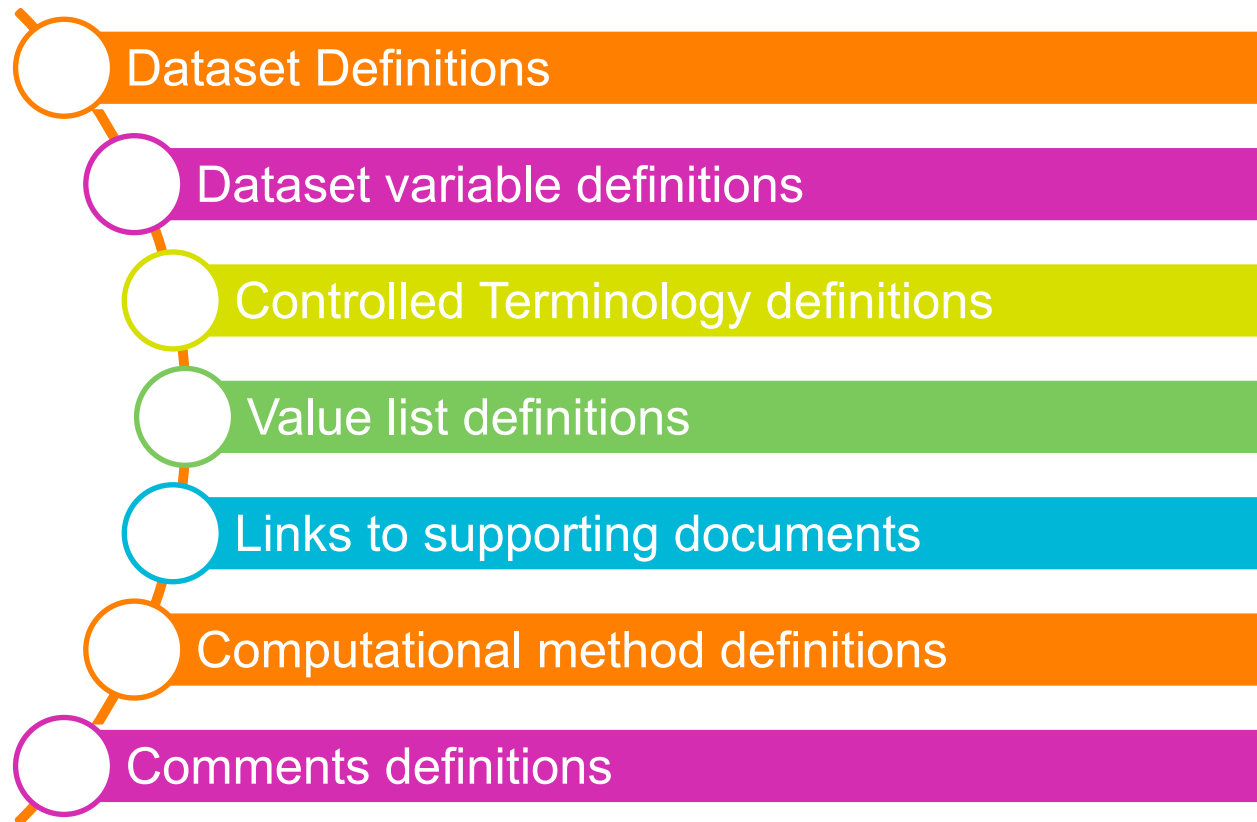
Date of document generation: 2013-03-03T17:04:44

Stylesheet version: 2013-04-24

Tabulation Datasets for Study CDISC01 (SDTM-IG 3.1.2)

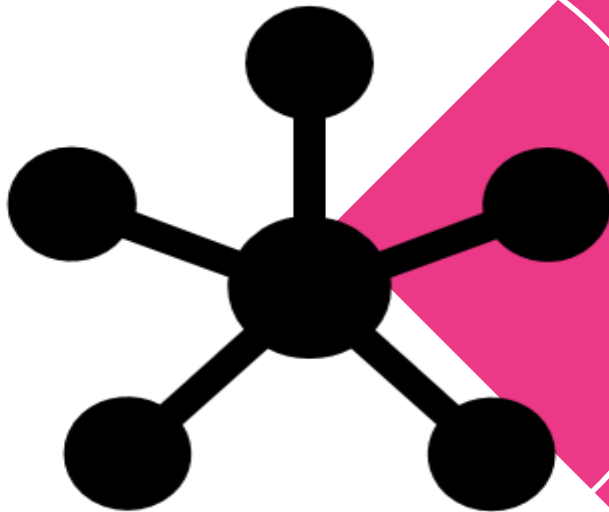
Dataset	Description	Class	Structure	Purpose	Keys	Location	Documentation
TA	Trial Arms	TRIAL DESIGN	One record per planned Element per Arm	Tabulation	STUDYID, ARMCD, TAETORD	ta.xpt	
TE	Trial Elements	TRIAL DESIGN	One record per planned Element	Tabulation	STUDYID, ETCD	te.xpt	
TI	Trial Inclusion/Exclusion Criteria	TRIAL DESIGN	One record per I/E criterion	Tabulation	STUDYID, IETESTCD	ti.xpt	
TS	Trial Summary	TRIAL DESIGN	One record per trial summary parameter value	Tabulation	STUDYID, TSARMCD, TSSEQ	ts.xpt	
TV	Trial Visits	TRIAL DESIGN	One record per planned Visit per Arm	Tabulation	STUDYID, VISITNUM, ARMCD	tv.xpt	
DM	Demographics	SPECIAL PURPOSE	One record per subject	Tabulation	STUDYID, USUBJID	dm.xpt	See Reviewer's Guide, Section 2.1.1 Demographics Reviewers Guide
SE	Subject Elements	SPECIAL PURPOSE	One record per actual Element per subject	Tabulation	STUDYID, USUBJID, SESTOTC, SEENDTC, TAETORD, ETCD	se.xpt	
SV	Subject Visits	SPECIAL PURPOSE	One record per actual visit per subject	Tabulation	STUDYID, USUBJID, SVSTOTC, VISITNUM	sv.xpt	
CM	Concomitant Medications	INTERVENTIONS	One record per recorded medication occurrence or constant-dosing interval per subject	Tabulation	STUDYID, USUBJID, CMSTOTC, CMENDTC, CMCAT, CMTRT, CMDOSTXT, CMDOSE, CMINDC, CMDOSEFRQ	cm.xpt	
EX	Exposure	INTERVENTIONS	One record per constant dosing	Tabulation	STUDYID, USUBJID	ex.xpt	

The key metadata components to support submissions are:



DATASET-XML

<https://www.cdisc.org/standards/transport/dataset-xml>



- Dataset-XML supports the exchange of dataset data based on Define-XML metadata.
- Dataset-XML complements Define-XML and provides an alternative to the SAS V5 Transport format for the exchange of study datasets for CDISC's Foundational standards.
- Dataset-XML is a truly non-proprietary, global standard, removing many SAS V5 Transport file restrictions (the current file format required by the FDA and PMDA), such as 8-character variable names and 200-character text fields

Future roadmap:

- Dataset-XML v1.0 has been shown to work as a SAS V5 XPORT replacement, the CDISC XML Technologies Team will add additional features in the next versions, including improved relationships and traceability.

DATASET-XML

<https://www.cdisc.org/standards/transport/dataset-xml>



- CDISC developed Dataset-XML v1.0 as a drop-in replacement for SAS V5 XPORT to enable testing using existing processes.
- New Dataset-XML features were intentionally not implemented to simplify the comparison.
- Dataset-XML supports exchanging tabular data in clinical research applications using ODM-based XML technologies, enabling communication of study results and regulatory submissions.
- Dataset-XML is a truly non-proprietary, global standard, removing many SAS V5 Transport file restrictions, such as 8-character variable names and 200-character text fields.

DATASET-XML

<https://www.cdisc.org/standards/transport/dataset-xml>



Dataset-XML can represent any tabular dataset including SDTM, ADaM, SEND, or non-standard legacy datasets. Noteworthy items relating to Dataset-XML v1.0 include:

Alternative to SAS Version 5 Transport (XPORT) format for datasets

ODM-based model for representation of SEND, SDTM, ADaM or legacy datasets

Capable of supporting CDISC regulatory data submissions

Based on Define-XML v2 or v1 metadata, easy to reference

Dataset-XML supports all language encodings supported by XML

In November 2013, the Food and Drug Administration (FDA) issued a Federal Register (FR) Notice of a Pilot Project called "Transport Format for the Submission of Regulatory Study Data." The purpose of the pilot was to conduct an initial analysis of CDISC's DS-XML as an alternative solution to the challenges of SAS XPORT V5 transport. Additional testing will be needed to evaluate cost versus effectiveness as an alternate transport format. FDA envisions conducting several pilots to evaluate new transport formats before a decision is made to support a new format. The initial pilot ended with challenges encoding and file sizes.

DATASET-XML

Document Structure



The below show the basic structure of any ODM v1.3.2 document that contains Dataset-XML content.

Dataset-XML document structure using the ClinicalData element

```
<?xml version="1.0" encoding="UTF-8"?>
<ODM
  xmlns="http://www.cdisc.org/ns/odm/v1.3"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:data="http://www.cdisc.org/ns/Dataset-XML/v1.0"
  FileType="Snapshot"
  ODMVersion="1.3.2"
  data:DatasetXMLVersion="1.0"
  FileOID="www.cdisc.org.Studydisc01-Define-XML_2.0.0(IG.CM)"
  PriorFileOID="www.cdisc.org.Studydisc01-Define-XML_2.0.0"
  Originator="CDISC Dataset-XML Team"
  CreationDateTime="2014-03-20T21:45:33">
  <ClinicalData
    StudyOID="cdisc01"
    MetaDataVersionOID="MDV.CDISC01.SDTMIG.3.1.2.SDTM.1.2">
    <!-- Dataset (CM) -->
    <ItemGroupData ItemGroupOID="IG.CM" data:ItemGroupDataSeq="1">
      <ItemData ItemOID="IT.STUDYID" Value="CDISC01"/>
      <ItemData ItemOID="IT.CM.DOMAIN" Value="CM"/>
      <ItemData ItemOID="IT.USUBJID" Value="CDISC01.100008"/>
      <ItemData ItemOID="IT.CM.CMSEQ" Value="1"/>
      <ItemData ItemOID="IT.CM.CMTRT" Value="PROCARDIA XL"/>
    </ItemGroupData>
  </ClinicalData>
</ODM>
```

Dataset-XML document structure using ReferenceData

```
<?xml version="1.0" encoding="UTF-8"?>
<ODM
  xmlns="http://www.cdisc.org/ns/odm/v1.3"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:data="http://www.cdisc.org/ns/Dataset-XML/v1.0"
  FileType="Snapshot"
  ODMVersion="1.3.2"
  data:DatasetXMLVersion="1.0"
  FileOID="www.cdisc.org.Studydisc01-Define-XML_2.0.0(IG.TA)"
  PriorFileOID="www.cdisc.org.Studydisc01-Define-XML_2.0.0"
  Originator="CDISC Dataset-XML Team"
  CreationDateTime="2014-03-20T21:45:33" >
  <ReferenceData
    StudyOID="cdisc01"
    MetaDataVersionOID="MDV.CDISC01.SDTMIG.3.1.2.SDTM.1.2">
    <!-- Dataset (TA) -->
    <ItemGroupData ItemGroupOID="IG.TA" data:ItemGroupDataSeq="1">
      <ItemData ItemOID="IT.STUDYID" Value="CDISC01"/>
      <ItemData ItemOID="IT.TA.DOMAIN" Value="TA"/>
      <ItemData ItemOID="IT.TA.ARMCD" Value="PLACEBO"/>
      <ItemData ItemOID="IT.TA.ARM" Value="Placebo"/>
      <ItemData ItemOID="IT.TA.TAETORD" Value="1"/>
      <ItemData ItemOID="IT.TA.ETCD" Value="SCREEN"/>
      <ItemData ItemOID="IT.TA.ELEMENT" Value="Screening"/>
      <ItemData ItemOID="IT.TA.EPOCH" Value="SCREEN"/>
    </ItemGroupData>
  </ReferenceData>
</ODM>
```

DATASET-XML AND DEFINE-XML

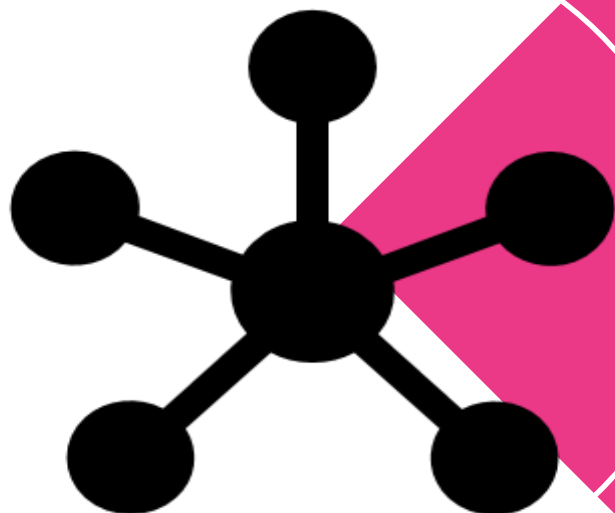
- Dataset-XML defines a standard format for transporting tabular dataset data in XML.
- The metadata for a dataset contained within an Dataset-XML document must be specified using the Define-XML standard.
- The Define-XML must be contained within the same folder as the dataset document files.
- Each Dataset-XML file contains data for a single dataset but a single define.xml file describes all the datasets included in the folder.
- The Dataset-XML file containing the data may be linked to the define.xml file containing the metadata by the PriorFileOID attribute on the root ODM node.

BENEFITS OF DATASET-XML

- Open, non-proprietary standard without the field width or data set and variable naming restrictions of SAS V5 Transport files
- Supports representation of data relationships, metadata versions and audit trails
- Harmonized with BRIDG, CDISC Controlled Terminology
- Data elements include references to metadata in Define-XML
- Straightforward implementation starting from SDTM data in SAS
- Supports FDA goal of encouraging open source reviewer tool development
- Facilitates Validation since both data and metadata share underlying technology
- Enables re-thinking some of the length restrictions in standards

RESOURCE DESCRIPTION FRAMEWORK (RDF)

<https://www.cdisc.org/standards/transport/rdf>



- CDISC Standards in RDF provides a representation of the CDISC Foundational standards in a model based on the Resource Description Framework (RDF).
- RDF provides executable, machine-readable CDISC standards from CDISC SHARE. This file format is a “linked data” view of the standards as an ontology.

RESOURCE DESCRIPTION FRAMEWORK (RDF)

Characteristics of RDF



- The Resource Description Framework (RDF) provides a universal, mathematically precise, and computable language that can express a wide range of information.
- RDF can express information about meta-models, models, and data in the same universal language.
- Once expressed in RDF, information can be represented, accessed, computed, integrated, and exchanged without the need for any translations.
- This representation in RDF avoids information mismatches that often happen when systems interface with each other. RDF provides a consistent language and modeling framework, identical at design and run-time. As a W3C Recommendation, it can fulfill the role of a standard language to express and exchange standards.
- Version 1.0 of the CDISC Standards in RDF, prepared by the PhUSE CS Semantic Technology Working Group, consists of two documents:
 - CDISC Standards in RDF Reference Guide v1 Final - provides a reference to the representation of the existing foundational CDISC standards in a model based on the Resource Description Framework (RDF).
 - CDISC Standards in RDF User Guide v1 Final – describes how to access and use the RDF files and provides background on their creation.
- The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. It is a collaborative effort led by W3C with participation from many researchers and industrial partners. RDF, which is one of the fundamental building blocks of the Semantic Web, gives a formal definition for that interchange.

RESOURCE DESCRIPTION FRAMEWORK (RDF)

Characteristics of RDF



- RDF- the Resource Description Framework is a standard model for data interchange on the Web.
- RDF has features that facilitate data merging even if the underlying schemas differ, and it specifically supports the evolution of schemas over time without requiring all the data consumers to be changed.
- The fundamental model of RDF is independent of XML.
- RDF is a model describing qualified (or named) relationships between two (Web) resources, or between a Web resource and a literal.
- At that fundamental level, the only commonality between RDF and the XML World is the usage of the XML Schema datatypes to characterize literals in RDF.
- The key elements of the RDF meta-model are based on the ISO 11179 standard for Metadata Registries (MDR), a standard that also provides the background model for the CDISC metadata registry called SHARE.

RESOURCE DESCRIPTION FRAMEWORK (RDF)

RDF SCHEMAS FOR CDISC FOUNDATIONAL STANDARDS



- Three RDF schemas have been defined and layered that enable the description of the CDISC foundational standards for CDASH, SDTM, SEND, and ADaM, together with their controlled terminologies.

The meta-model schema (namespace prefix mms) is a generic RDF schema with elements from ISO 11179 that allows the specification of data oriented models in a unified way.

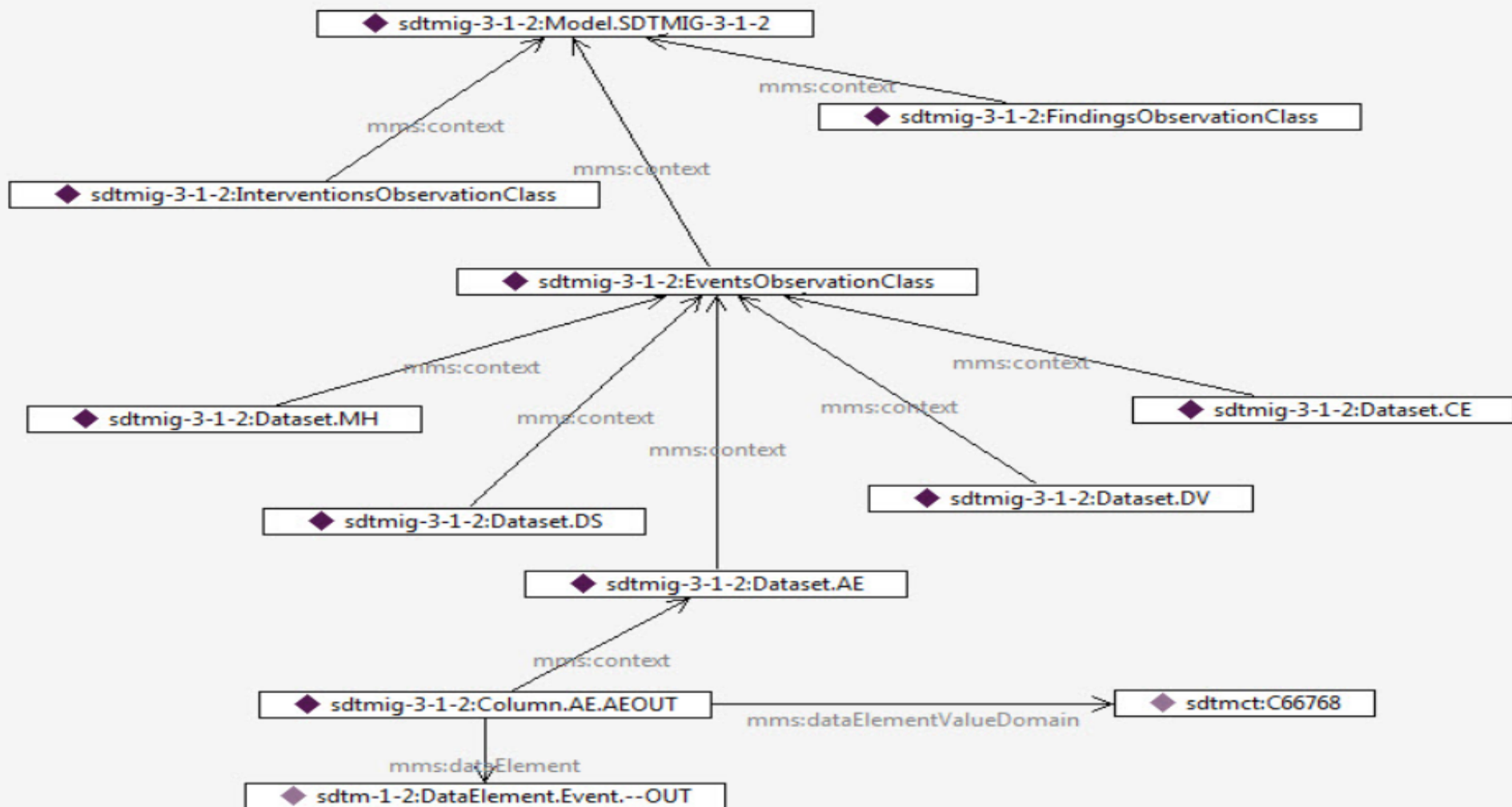
The CT schema (namespace prefix cts) defines additional predicates used by the NCI EVS to publish the CDISC controlled terminology in RDF.

The CDISC schema (namespace prefix cdiscs) introduces additional classes and predicates to capture CDISC specific model information

RESOURCE DESCRIPTION FRAMEWORK (RDF)

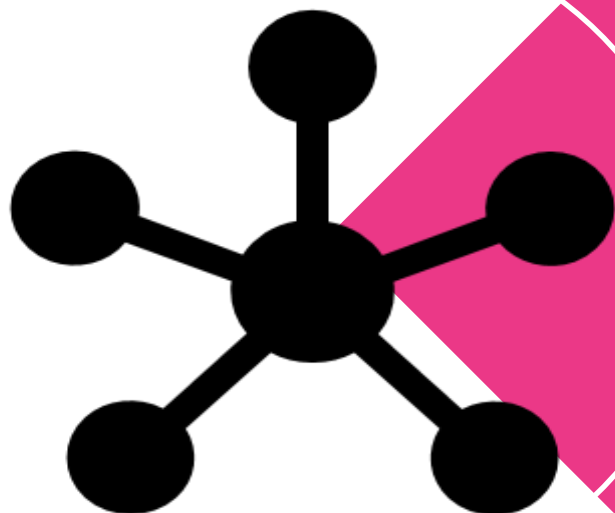
EXAMPLE RDF GRAPH FOR SDTM IG 3.1.2

Representing subjects and objects of triples as nodes, and predicates of triples as directed edges between nodes, one can see that the information expressed by a set of triples is in fact a directed graph, which expresses the intrinsic data model of an RDF data set.



LABORATORY DATA MODEL (LAB)

<https://www.cdisc.org/standards/transport/lab>



- LAB provides a standard model for the acquisition and exchange of laboratory data, primarily between labs and sponsors or CROs.
- The LAB standard was specifically designed for the interchange of lab data acquired in clinical trials.

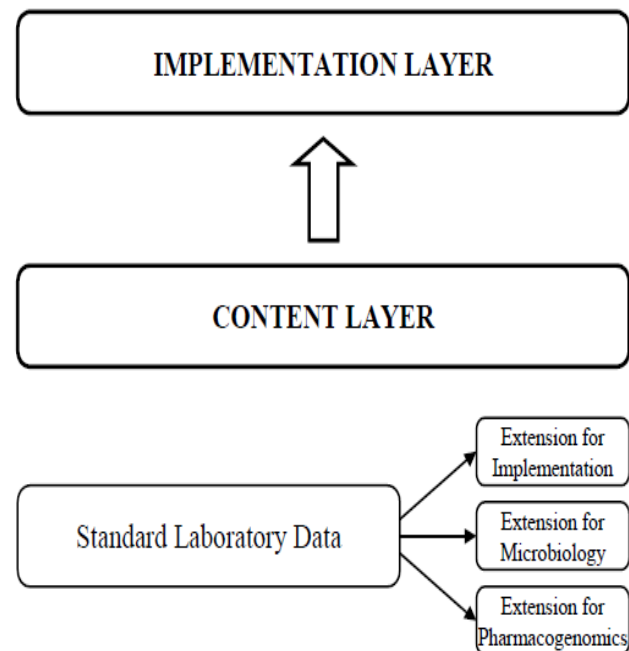
LABORATORY DATA MODEL (LAB)

<https://www.cdisc.org/standards/transport/lab>



Standard models for the interchange of laboratory data do exist already but they are very seldom used within the biopharmaceutical industry. Examples of such standards are ACDM, ASTM, HL7 and X12. The main reason standards such as these have not been more accepted by the industry is that they have limited applicability to clinical trial data and hence have limited use to central laboratories, CROs or biopharmaceutical companies.

The design of the model is thus as follows:



The first layer would be the content layer and above that would be an implementation layer, the idea being that the content would not change but the implementation could. The advantage of this approach is that it offers flexibility but retains control: it doesn't make the use of the model dependent upon any one implementation and if different implementations are used the content remains the same so the standard still applies.

Standards Development

Drivers



CDISC Team & Volunteers



SHARE Ecosystem



CDISC Standards are required for regulatory submissions to FDA (U.S.) and PMDA (Japan), endorsed by China FDA, and requested for use by the European Innovative Medicines Initiative (IMI).

BENEFITS OF IMPLEMENTING CDISC STANDARDS

- Fostered efficiency
- Enhanced innovation
- Facilitated data sharing
- Increased predictability
- Complete traceability
- Improved data quality
- Reduced costs
- Streamlined processes



Conclusion



CDISC and PhUSE partner to further the mission of each organization, with CDISC focusing on developing global, platform-independent data standards, and PhUSE focusing on implementing CDISC standards. The two organizations combine efforts on key initiatives around end-to-end standards and semantics, strengthening an interdependent process.

CDISC Transport Standards enable the exchange of data conformant with CDISC Foundational Standards and their Therapeutic Area extensions.

Each of the standards described above plays its roles in transportation of data during the entire phase of clinical trials from Study Design and all the way to regulatory submission.

These standards are used as references for transportation of trial data from trial sites, central laboratories, and regulatory filings.

With advent of RDF and Meta Data Repositories, Transport standards especially ODM would play a vital role in integrating with Fast Healthcare Interoperability Resources (FHIR) to support Real World Evidence (RWE).

WE ARE DEDICATED TO
THE FUTURE OF CLINICAL DEVELOPMENT
AND TO EVERY LIFE IT SAVES.



PRAHEALTHSCIENCES