cdisc

2023
EUROPE
INTERCHANGE
COPENHAGEN | 26-27 APRIL

CORE

Presented by: Amy Palmer, CDISC, Head of Standards Development Operations

Meet the Speaker

Amy Palmer

Title: Head of Standards Development Operations

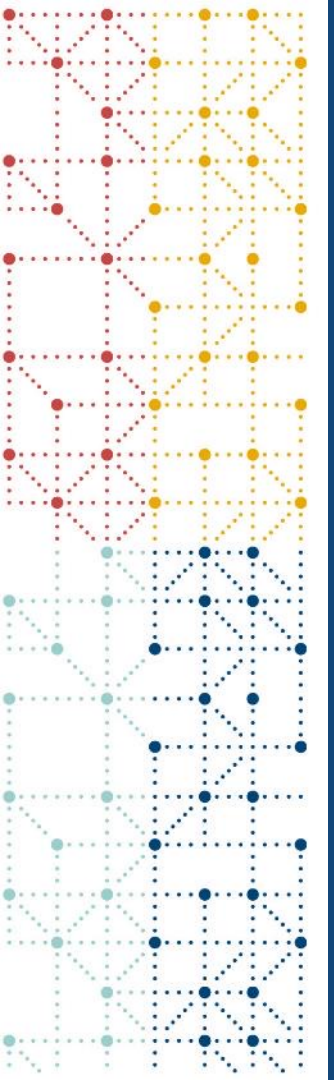
Organization: CDISC

Amy Palmer is the Head of Standards Development Operations at CDISC. Amy has been with CDISC since 2013. She is a member of the CDISC Technical Leadership Team and leads the Global Governance Group. Amy has over 28 years' experience working in clinical research. She has been involved in the development of multiple therapeutic area user guides as well the foundational standards and has been working with CDISC standards since 2010.

Amy has a BS from Mary Washington College and an MPH from the University of Montana.

Agenda

1. CDISC Conformance Rules – Where we started
2. Source of Rules – Where do they come from
3. Impacts to Standards – What we learned
4. Conformance Rules Best Practices – What we are doing
5. Moving Forward
6. How to Become Involved



CDISC Conformance Rules

Where we started…

About CORE

- The CORE Project objectives are to:
- Ensure each standard has a set of unambiguous, executable Conformance Rules
- Ensure consistency across Conformance Rule implementations
- Expedite the availability of executable Conformance Rules for new Foundational Standards
- Create executable Conformance Rules vetted by the CDISC standards development teams
- Create a Reference Implementation of an open-source engine that executes the Rules
- Release the open-source engine under the CDISC Open-Source Alliance (COSA)

Rules Specifications and Executable Rules Development

Conformance Rule Specification Development

Human-readable Specification

Rule ID	SDTMIG Version	Rule Version	Class	Domain	Variable	Condition	Rule
CG0225	3.4	1	ALL	ALL	VISITDY	VISITNUM is NOT in TV.VISITNUM	VISITDY = null
Document	**Section**	**Item**	**Cited Guidance**				
IG v3.4	4.4.5		VISITDY must not be populated for unplanned visits, since VISITDY is, by definition, the planned study day of visit, and since the actual study day of an unplanned visit belongs in a --DY variable.				

Authoring Sources:

- CDISC Standards
- FDA Validation Rules
- Community proposals

Centralized

Specifications Catalog

CORE Rule Development

Executable Rule (YAML) in CORE Rule Editor

```yaml
Core:
  Id: CDISC.SDTMIG.CG0225
  Version: '1'
Authority:
  Organization: CDISC
Description: Trigger error if VISITYDY is populated when VISITNUM is not in TV.
References:
    Origin: SDTM and SDTMIG Conformance Rules
    Rule Identifier:
      Id: CG0225
      Version: '1'
    Version: '2.0'
Sensitivity: Record
Severity: Warning
Rule Type: Value Presence
Scopes:
  Classes:
    Include:
      All
  Domains:
    Include:
      All
  Standards:
    Name: SDTMIG
    Version: '3.4'
Operations:
  domain: TV
  id: $tv_visitnum
  name: VISITNUM
  operator: distinct
```

Rule developed and tested in CORE Rule Editor and CORE Engine

Publish

CDISC Library

Catalog of Conformance Rules

SDTMIG Rules

SENDIG Rules

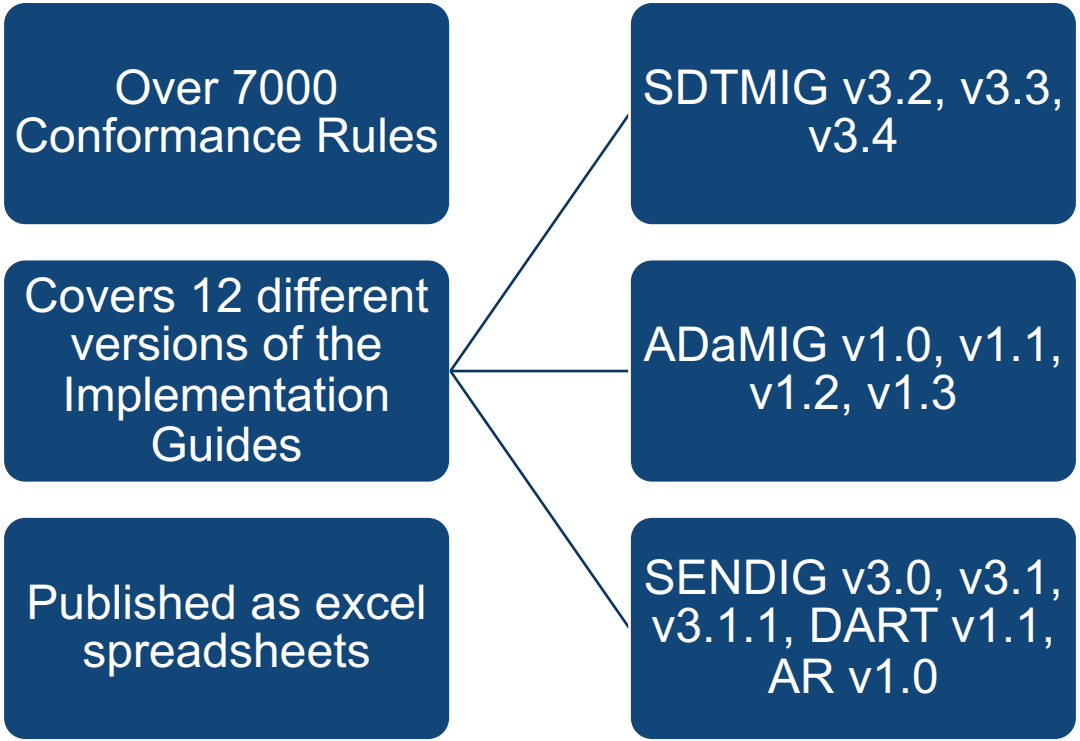
ADaMIG Rules

Define.xml Cross-check rule

- Some part of IG rules, some are needed to ensure consistency between Define.xml file and SDTM or ADaM datasets

Regulatory Validator Checks

Volume of Conformance Rules



Developing CORE Rules

Use published Conformance Rules as the source for CORE Rules

Data Checks – Unit Testing

- Positive result – data is conformant to the standard
- Negative result – data is not conformant to the standard

Rules are entered as failure criteria in the Rule Editor

cdisc

Rule Executability

Not Executable

- The rule has been identified as not being able to be programmatically checked

Partially Executable

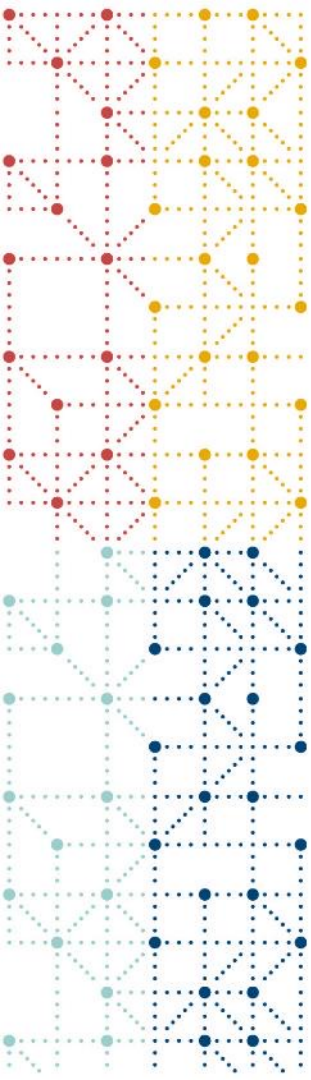
- Only a portion of the rule is able to be programmatically checked

Partially Executable – possibly overreporting

Partially Executable – possibly underreporting

Fully Executable

- The rule is able to be programmatically checked



Source of Rules

Where do they come from

Where do these rules come from?

- An **Expected** variable is any variable necessary to make a record useful in the context of a specific domain. Expected variables may contain some null values, but in most cases will not contain null values for every record. When the study does not include the data item for an expected variable, however, a null column must still be included in the dataset, and a comment must be included in the Define-XML document to state that the study does not include the data item.

	Rule ID	SDTMIG Version	Rule Version	Class	Domain	Variable	Condition	Rule
38	CG0016	3.2	1	ALL	ALL	GEN	Variable Core Status = Expected	Variable present in dataset
39	CG0016	3.3	1	ALL	ALL	GEN	Variable Core Status = Expected	Variable present in dataset
40	CG0016	3.4	1	ALL	ALL	GEN	Variable Core Status = Expected	Variable present in dataset

Regulatory Validator Rules

FDA Validator Rule ID	Publisher	Publisher ID	FDA Validator Rule Message	FDA Validator Rule Description
SD1137	CDISC	CG0348	Non-missing value for --ORRES, when --DRVFL='Y'	Character Result/Finding in Original Units (--ORRES) value should be NULL, when Derived Flag (--DRVFL) value is 'Y'.
SD1235	CDISC	CG0554	Neither SPDEVID nor USUBJID values are populated	Value for Sponsor Device Identifier (SPDEVID) or Unique Subject Identifier (USUBJID) variables should be populated for all records in Divice In-Use (DU) domain
SD9999	CDISC	CG0320, CG0321, CG0463, CG0464, 266, 266.1, 267, 267.1, 268, 268.1, 269, 269.1	Dataset class not recognized	The structure for custom dataset should be based on one of the general observation classes (EVENTS, FINDINGS, INTERVENTIONS) defined by the SDTM model.
SD1075	CDISC	CG0467, 78	Variable not recommended for use	Variables described in IG as not recommended for usage should be not included in the dataset.
SD1072	CDISC	CG0204, CG0371, CG0203, 286	Missing IDVAR value, when RDOMAIN value is provided	Value of Identifying Variable (IDVAR) variable must be populated, when Related Domain Abbreviation (RDOMAIN) variable value is provided, with the only exception of 'DM' value for RDOMAIN.
SD2239	CDISC	CG0240	Inconsistent value for --TPT	Planned Time Point Name (--TPT) value must be consistent for all records with same Subject (USUBJID) and Assessment Date/Time (--DTC).

Now let's do this for an actual rule

13	DTHDTC	Date/Time of Death	Char	ISO 8601 datetime or interval	Record Qualifier		Not in nonclinical trials	C117450	The date or date and time of death, represented in a standardized character format.	Should represent the date/time that is captured in the clinical-trial database.

14	DTHFL	Subject Death Flag	Char		Record Qualifier		Not in nonclinical trials	C117451	An indication that the subject died.	A value of "Y" indicates the subject died. Should be "Y" or null. Should be populated even when the death date is unknown.

	1	2	3	4	5	6	7	8
1	Rule ID	SDTMIG Version	Rule Version	Class	Domain	Variable	Condition	Rule
1043	CG0435	3.2	1	SPC	DM	DTHFL	DTHDTC ^= null	DTHFL = 'Y'
1044	CG0435	3.3	1	SPC	DM	DTHFL	DTHDTC ^= null	DTHFL = 'Y'
1045	CG0435	3.4	1	SPC	DM	DTHFL	DTHDTC ^= null	DTHFL = 'Y'

```yaml
Check:
  all:
    - name: DTHDTC
      operator: non_empty
    - name: DTHFL
      operator: not_equal_to
      value: Y
Core:
  Id: CORE-000007
  Version: '1'
  Status: Published
Description: Raise an error when DTHDTC is not empty and DTHFL not equal to "Y"
Outcome:
  Message: DTHFL is not "Y", when DTHDTC is populated
  Output Variables:
    - DTHDTC
    - DTHFL
Rule Type: Record Data
Sensitivity: Record
Authorities:
  - Organization: CDISC
    Standards:
      - Name: SDTMIG
        Version: '3.4'
        References:
          - Origin: SDTM and SDTMIG Conformance Rules
            Rule Identifier:
              Id: CG0435
              Version: '1'
            Version: '2.0'
            Citations:
              - Cited Guidance: DTHDTC[The date or date and time of death, represented
                  in a standardized character format.] DTHFL[An indication that the
                  subject died.]
                Document: Model v2.0
                Item: DTHDTC|DTHFL
                Section: Demographics
Scope:
  Classes:
    Include:
```

Same Rule in the Rule Editor

Impacts to Standards

What we learned….

Conformance Rules Intent and Meaning

- Some rules are ambiguous?
- What does the standard say?
- Did the authors really mean to say this?

Deciphering Rule Language

Condition	Rule
MIDSTYPE = TM.MIDSTYPE and TM.TMRPT = 'Y'	MIDS is suffixed with a sequence number in ***consistent*** chronological order

Cited Guidance

For types of Disease Milestones that can occur multiple times, MIDS will usually be an abbreviated version of MIDSTYPE and will always end with a sequence number. Sequence numbers should start with one and indicate the chronological order of the instances of this type of Disease Milestone.

How will these learnings impact the standards going forward?

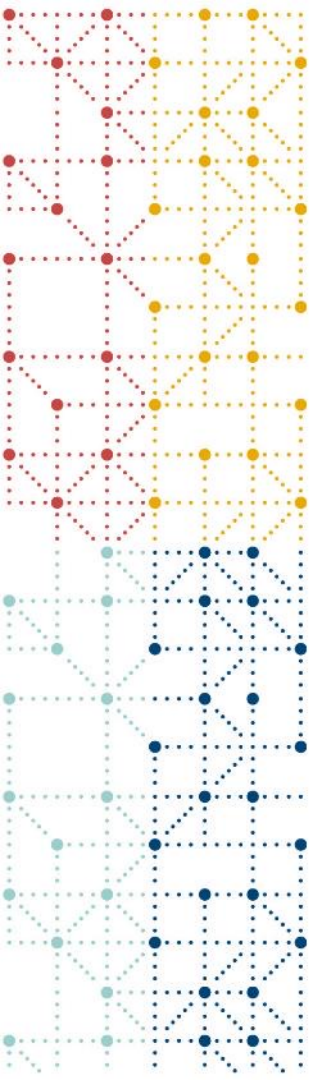
Write	Write rules with the Rule Editor in mind
Logic	Use logic to draft rules
Clear	Clear, concise, unambiguous language in the Implementation Guides and Models
Limit	Limit words like “should” and “may” in cited guided intended for conformance rules

Add clarifying language within the IGs to solidify the intent and meaning of the conformance rules

Currently the IG reads:
8. In RELREC, if a dataset-level relationship is defined for a split Findings About domain, then RDOMAIN may contain the 4-character dataset name, rather than the domain name "FA".

Comments from the CORE Team:

This whole assumption should be removed as rules cannot be developed to check this and FASEQ could still be used for RELREC as needed for record level and FALNKID/FALNKGRP used for dataset level relationships. RDOMAIN being a dataset is only mentioned here and muddies the definition of RDOMAIN.



Conformance Rules Best Practices

What we are doing…

Conformance Rule Specifications & CORE Rules

Rule specifications

- Human-readable
- Sources:
  - Typically, a part of CDISC foundational standards (developed per COP-001)
  - Alternately, Regulatory-provided rules
  - In the future, community-provided data quality rules or traceability rules

CORE Rules

- Machine-executable
- Re-expression of conformance rule specifications in machine-executable form
- Developed by the CDISC community
- Includes both an explanation or a description of the rule as well as a human-readable output message
- Includes all relevant citations and sources for a rule

Rules based on CDISC Standards are part of the CDISC Standards

- Draft rules within the CORE Rules Editor
- Use Rule Description and Outcome Message to review rule during Internal and Public Review
- Cited Guidance is part of the Rule
- Rule logic is transparent

Moving Forward

How do the different standards development teams consider conformance when updating and developing new standards?

Consider wording – “should” versus “must”

- Conformance = “must”, what about the “shoulds”?

Write rules in executable language – move away from pseudo code

- Rules developed using YAML schema

- Moving away from excel spreadsheets

Data Quality Checks

- Incorporation of data quality checks
- Community submissions
- Curated by CORE Rules Governance Team

Data Quality Check Examples

- Assessment dates prior to Informed Consent
- Inconsistent Coding
- Missing data
- Out of range data

CORE Key Contributors

Recognize individuals and their organizations for ongoing volunteer support for CORE project on the CDISC Website

CORE Badge available for email signatures and social media accounts

How to Become Involved

Rules Developer Skill Set

Core Skills

- Data savvy with science background; e.g., statistics, biometrics, data science
- A CDISC standards practitioner. Solid implementation experience with SEND, SDTM, and/or ADaM
- Experience in data specifications & associated verification & validation tasks

"Plus" Skills

- Some familiarity with the associated conformance rules
- Knowledgeable in structured data, such as XML, JSON, YAML
- A member of an organizational standards council or governance body

cdisc

Rules Developer Onboarding Process and Support

Training tools on the wiki, GitHub will further refine

Training webinars

Rule Workshops at F2F meetings

Weekly 1-hour CORE Rules Developer Meeting

Weekly 2-hour CORE Rules Developer "Office Hours"

Dedicated volunteers needed!

Currently limited number of active volunteers from industry

cdisc

How to Volunteer

- https://www.cdisc.org/volunteer/form
  - Select CORE Rules Team

Expected Engagement

- Time Period: 3 - 6 months, or longer, if able
- Hours per week: 2 - 4 hours, in addition to meeting attendance
- Weekly Meetings and Workshops: Recorded and available for review
  - Rules Development Workshop Tuesdays 9am - 11am ET
  - Weekly Team Meeting – Thursdays 11am - 12pm ET

Thank You!

apalmer@cdisc.org