PharmaSUG 2024 - Paper DS-276

Your Guide to Successfully Upversioning CDISC Standards

David Neubauer, Kapila Patel, Toral Patel, and Soumya Rajesh; IQVIA

ABSTRACT

As of early 2023, new versions of the CDISC standards (i.e., SDTM v1.7, SDTM IG v3.3, Define.xml v2.1) are either required or supported by the industry's regulatory agencies. This paper relays challenges and best practices the authors have experienced while up-versioning to these standards. Not all these practices are found in published standards. This paper will bring together the resources and lessons learned in one place, so that readers can skillfully navigate through the challenges of adopting these new standards. Highlights include strategies for dealing with value level metadata for variables with multiple codelist references, a new domain class, new domains, and domains referenced in TAUGs not seen in the IGs. We'll discuss best practices for data modeling: when to use new variables, supplemental qualifiers, and targeting the appropriate domains. We'll include experiences interpreting and dispositioning validation output from the applicable conformance rules.

INTRODUCTION

The purpose of this paper is to provide topics for consideration and discussion to those implementing new versions of standards. This paper will focus on SDTM standards and the Define-XML standards in support of SDTM. Strategies around upversioning CDASH, SEND, ADaM, and Controlled Terminology standards can be considered out of scope for this discussion.

When implementing CDISC standards for a regulatory submission, users must be aware of the agencies' published guidance for those submissions to be successful. The scope of this paper is limited to submissions to the U.S. Food & Drug Administration (FDA) with some clarifications for the Pharmaceutical and Medical Devices Agency (PMDA) in Japan. Both the FDA and PMDA have published Technical Conformance Guides (TCG), Data Standards Catalogs, and sets of Validation Rules. Per that guidance, the FDA requires SDTM v1.7 and SDTMIG v3.3 for clinical study datasets and Define-XML 2.1 for clinical study data definition files, both, for studies starting on or after March 15, 2023. The PMDA supports SDTM v1.7 and SDTMIG v3.3 for clinical study datasets for studies starting on or after April 1, 2023. As of September, 2023, the PMDA has not published support for Define-XML 2.1. The PMDA currently supports Define-XML 2.0 and 1.0 for clinical study data definition files for studies starting on or after October 1, 2016. The PMDA's support for Define-XML 1.0 ends March 31, 2025.

Topics covered include:

- Alignment of CDISC Standards
- When/Why of Upversioning
- Changes in SDTMIG v3.3 at a Glance
- Permissible Variables
- Subject Visits (SV)
- Physical Exams (PE)

- MO vs. Morphology/Physiology Domains
- Changes in Define-XML v2.1 at a Glance
- VLM for Multiple Codelists/Use Cases
- Conformance/Validation

ALIGNMENT OF CDISC STANDARDS

A large part of the strategy for upversioning a study or program of studies is knowing how standards align with each other. Per CDISC.org, SDTMIG v3.3 currently aligns with the following standards (as seen in Figure 1. SDTMIG v1.7 Alignment):

- SDTM (model) v1.7
- SDTM IG MD (Medical Devices) v1.1
- SDTM MSG (Metadata Submission Guidelines) v2.0
- SDTM and SDTM IG Conformance Rules v1.1

Note: While the SDTM Conformance Rules v1.1 release contains new rules for SDTMIG v3.3 and SDTM v1.7, the newer SDTM and SDTM IG Conformance Rules v2.0 contains updates to rules for SDTMIG v3.3 and SDTM v1.7 and can be considered for use in validation.

Additionally, and perhaps unofficially, SDTMIG v3.3 would also align with the following standards:

- SDTM IG AP (Associated Persons) v1.0
- SDTM IG PGx (Pharmacogenomics/Genetics) v1.0 (Provisional)

Note: The SDTM IG AP v1.0 initially aligned with SDTMIG v3.2 but updates to the APIG haven't been published since it's release as "Final" in 2013. The PGxIG was provisionally released in 2015, prior to SDTMIG v3.3, and has since been deprecated with the release of SDTMIG v3.4. Sponsors should consider the deprecation of PGxIG when working with v3.3 and genomics/pharmacogenomics data and decide whether or not to use PGxIG as written, use SDTM v1.7 domain templates with the domain abbreviations as appearing in SDTMIG v3.4, or create SDTM v1.7 based custom domains.

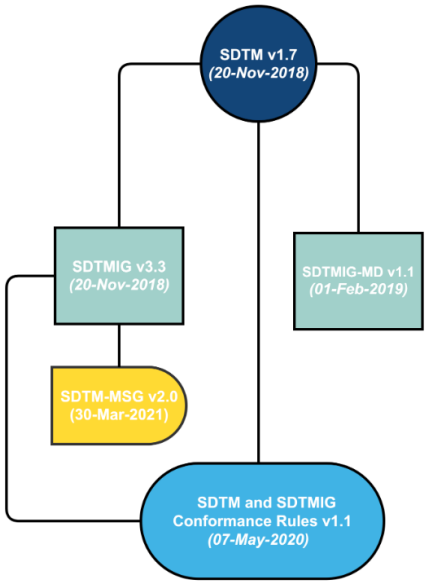

Figure 1. SDTMIG v1.7 Alignment

Define-XML 2.1, currently aligns with the following standards (as seen in Figure 2. Define-XML v2.1 Alignment):

- ODM v1.3.2
- Analysis Results Metadata (ARM) v1.0
- Dataset-XML
- Conformance Rules for Define-XML v2.1
- ADaM MSG v1.0
- SDTM MSG v2.0

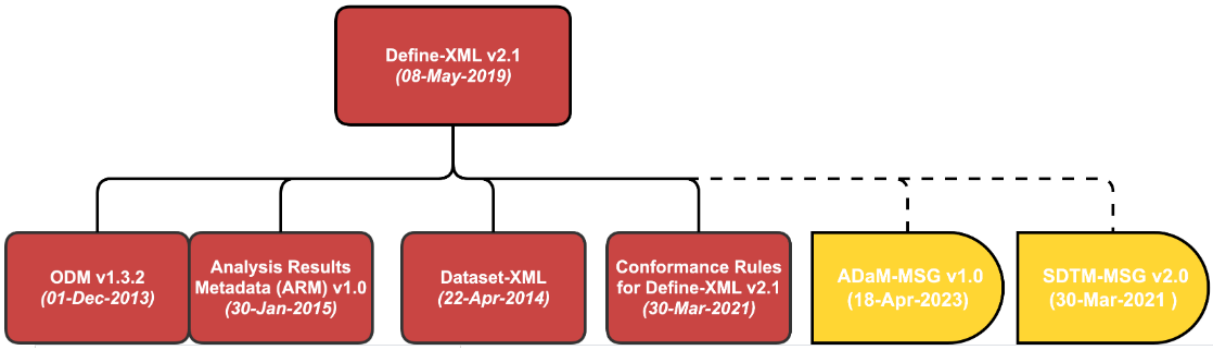

Figure 2. Define-XML v2.1 Alignment

With regard to Define-XML and SDTM Controlled Terminology, best practice would be to use a version by date following conventions described in the data standards catalogs by the regulatory agencies.

WHEN/WHY OF UPVERSIONING

Sponsors need to take into account the scope of a submission prior to up-versioning. They need to consider the reasons for doing so. Up-versioning shouldn't necessarily be just at a study level or decided upon by one person. Some points to consider while up-versioning are:

- Does the Sponsor have a Study Data Standardization Plan (SDSP)?
- Is there agreement between the Sponsor and the regulatory agency on which versions of standards to use?
- Does up-versioning align with company's processes?
- Is the study going to be part of pooled analyses (ISS/ISE)?
- What are the statuses of the studies to be included in the effort? Are some studies ongoing or are they legacy studies that need converting?
- Evaluate which domains, variables, codelists need to be updated/up-versioned or perhaps deprecated
- Documents to be updated as part of up-versioning like SDTM aCRF, SDRG, and define.xml
- Time, cost, and benefits of up-versioning

CHANGES IN SDTMIG V3.3 AT A GLANCE

There are a lot of changes in SDTMIG v3.3 and SDTM v1.7 from their previous versions. The Study Reference domain class is new with 3 domains (DI – Device Identifiers, OI – Non-host Organism Identifiers, and PB – Pharmacogenomic/Genetic Biomarker Identifiers) now described under that section. Newly modeled are 12 domains (7 findings, 2 interventions, and 1 each for relationship, special purpose, and trial design class domains):

CV – Cardiovascular System Findings.

FT – Functional Tests

MK – Musculoskeletal System Findings

NV – Nervous System Findings

OE – Ophthalmic Examinations

RE – Respiratory System Findings

UR – Urinary System Findings

AG – Procedure Agents

ML – Meal Data

RELSUB – Related Subjects

SM – Subject Disease Milestones

TM – Trial Disease Milestones

There are 168 standard variables newly modeled across the IG domains (highlights being ARMNRS, ACTARMUD, FOCID, --LOBXFL). There are also changes to handling permissible variables, expanded use cases for previously established domains (RS, MB, and MS), and a few deprecated items/components. A full revision history from prior versions for the IG and model can be found in the appendices of each document and in diff reports from the CDISC library.

PERMISSIBLE VARIABLES

A seemingly small but substantial change from prior versions to SDTMIG v3.3 are the expectations involving permissible variables.

As per SDTMIG v3.2, section 2.5 The SDTM Standard Domain Models:

“As long as no data was collected for Permissible variables, a sponsor is free to drop them and the corresponding descriptions from the Define-XML.”

As per SDTMIG v3.3, section 2.5 The SDTM Standard Domain Models:
“A Permissible variable should be used in an SDTM dataset wherever appropriate.

- If a study includes a data item that would be represented in a Permissible variable, then that variable must be included in the SDTM dataset, even if null. Indicate no data were available for that variable in the Define-XML document.
- If a study did not include a data item that would be represented in a Permissible variable, then that variable should not be included in the SDTM dataset and should not be declared in the Define- XML document.”

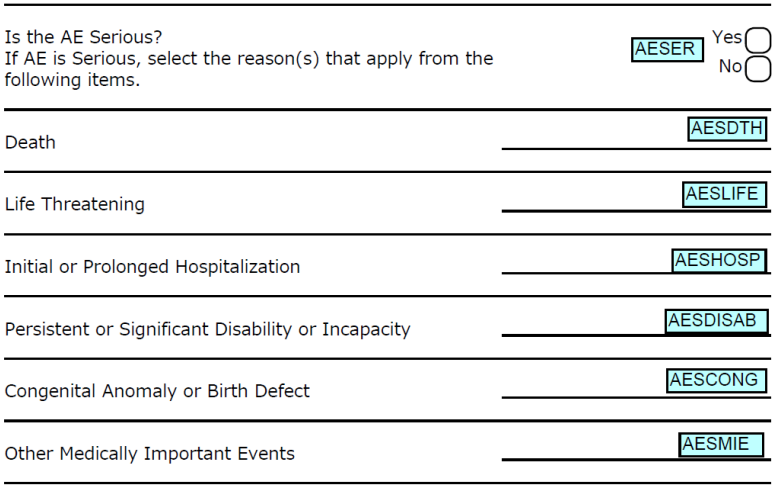

Image 1. Example Adverse Event (AE) SDTM aCRF

Categories of serious events are collected on the CRF but for this example study, no serious adverse events were reported. If the study implements SDTMIG v3.2 and no values were collected/reported for AESDTH, AESLIFE, AESHOSP, AESCONG, AESDISAB, and AESMIE, then they could be dropped from ae.xpt (as evident in Table 1). However, if using SDTMIG v3.3, in this scenario (no data reported) then the categories of serious events variables must be included in ae.xpt (as shown in Table 2) and indicated as having no data in the define.xml.

AESCAN (Involves Cancer) and AESOD (Occurred with Overdose) which are modeled permissible variables in SDTMIG v3.2 and v3.3 are not planned to have data collected on this example Adverse Events form and thus not included in ae.xpt (as evident in Table 1 and Table 2).

ROW	STUDYID	DOMAIN	USUBJID	AESEQ	AETERM	AESEV	AESER	AESTDTC
1	ABC123	AE	123101	1	HEADACHE	MODERATE	N	2023-01-21
2	ABC123	AE	123101	2	FEVER	MILD	N	2023-01-22
3	ABC123	AE	123101	3	BACK PAIN	MILD	N	2023-01-24

Table 1. SDTMIG v3.2 ae.xpt

ROW	STUDYID	DOMAIN	USUBJID	AESEQ	AETERM	AESEV	AESER	AESCONG	AESDISAB	AESDTH	AESHOSP	AESLIFE	AESMIE	AESTDTC
1	ABC123	AE	123101	1	HEADACHE	MODERATE	N							2023-01-21
2	ABC123	AE	123101	2	FEVER	MILD	N							2023-01-22
3	ABC123	AE	123101	3	BACK PAIN	MILD	N							2023-01-24

Table 2. SDTMIG v3.3 ae.xpt

AESCONG [No Data]		Congenital Anomaly or Birth Defect	text	Record Qualifier	1	No Yes Response, subset for variables with only "Y" or null values • "Y" = "Yes"	Collected (Source: Investigator) Annotated CRF [22]
AESDISAB [No Data]		Persist or Signif Disability/Incapacity	text	Record Qualifier	1	No Yes Response, subset for variables with only "Y" or null values • "Y" = "Yes"	Collected (Source: Investigator) Annotated CRF [22]
AESDTH [No Data]		Results in Death	text	Record Qualifier	1	No Yes Response, subset for variables with only "Y" or null values • "Y" = "Yes"	Collected (Source: Investigator) Annotated CRF [22]
AESHOSP [No Data]		Requires or Prolongs Hospitalization	text	Record Qualifier	1	No Yes Response, subset for variables with only "Y" or null values • "Y" = "Yes"	Collected (Source: Investigator) Annotated CRF [22]
AESLIFE [No Data]		Is Life Threatening	text	Record Qualifier	1	No Yes Response, subset for variables with only "Y" or null values • "Y" = "Yes"	Collected (Source: Investigator) Annotated CRF [22]
AESMIE [No Data]		Other Medically Important Serious Event	text	Record Qualifier	1	No Yes Response, subset for variables with only "Y" or null values • "Y" = "Yes"	Collected (Source: Investigator) Annotated CRF [22]

Image 2. Define-XML v2.1 ae.xpt HasNoData

SUBJECT VISITS (SV)

The Visit Events (VE) domain was an interim suggestion that was rolled out in the “Guidance for Ongoing Studies Disrupted by COVID-19” to capture missed visits or visits that did not occur, due to COVID. This was because of the limitations of SV domain in SDTMIG v3.3. There was no supplemental qualifier or additional variables in SV to capture missed visits. New variables have since been introduced in SV domain in SDTMIG v3.4, in response to FDA feedback, however the FDA is not yet accepting this version of the standards. So, it is advisable to continue using SV domain to submit to SDTMIG v3.3, by introducing the new variables from SDTM v3.4 as non-standard variables in SV.

From the FDA TCG June 2023: “It is the current preference of the Agency that for all clinical studies, not limited to those impacted by COVID-19, subject visit data for scheduled (whether or not they occurred), and unscheduled visits be submitted in one single dataset structured as the current CDISC Subject Visits (SV) domain. It is also Agency preference that three non-standard variables (NSVs) for missed visits, --REASOC (Reason for Occur Value), --EPCHGI (Epi/Pandemic Related Change Indicator), and --CNTMOD (Contact Mode), outlined in the CDISC document “Guidance for Ongoing Studies Disrupted by COVID-19 Pandemic” be included within the SV domain and not within the supplemental SUPPSV domain or in other SDTM datasets.”.

In the define.xml for SV in 3.3 SDTMIG implementation have those variables (SVPRESP, SVOCCUR, SVREASOC, SVCNTMOD, SVEPCHGI) flagged w/ def:IsNonStandard = "Yes". In the SDRG explain any warnings/errors about using these non-standard variables as they meet expectations set forth by the FDA TCG.

ROW	STUDYID	DOMAIN	USUBJID	VISITNUM	VISIT	SVPRESP	SVOCCUR	SVREASOC	SVCNTMOD	SVEPCHGI	SVSTDTC	SVENDTC	SVUPDES
1	123456	SV	101	1	SCREEN	Y	Y		IN PERSON		2020-02-13	2020-02-18	
2	123456	SV	101	2	DAY 1	Y	Y		IN PERSON		2020-02-19	2020-02-19	
3	123456	SV	101	3	WEEK 1	Y	Y		IN PERSON		2020-02-25	2020-02-25	
4	123456	SV	101	4	WEEK 2	Y	N	CLINIC CLOSED DUE TO HOT WEATHER					
5	123456	SV	101	4.1	WEEK 2: UNSCHEDULED 1				REMOTE AUDIO VIDEO	Y	2020-03-07	2020-03-07	EVALUATION OF AE
6	123456	SV	101	8	FOLLOW-UP	Y	Y		TELEPHONE CALL	Y	2020-03-16	2020-03-16	

Table 3. SDTMIG v3.3 sv.xpt

PHYSICAL EXAMS (PE)

PE maps to PR: It has been observed that sponsors tend to map trigger questions on Physical Exam (or Neurological Exam) pages to PEORRES (or NVORRES) as "Performed" when the response to the lead question is "Yes". This is not the way the PE domain is intended to be used, and it creates false or redundant records in SDTM without an actual result, because it goes against the purpose of PE Per CDISC (and CDASH). Per SDTMIG v3.3 PEORRES should be the “Text description of any abnormal findings” per body system defined in PETEST/PETESTCD. When you don’t have actual results, but rather just a prompt for whether the Physical Exam was done, it should map to PR as the occurrence of a procedure, based on the guidance provided in the Knowledge Base Article (KBA): “How should I represent whether a physical exam was performed in SDTM?”. According to the KBA, SDTMIG v3.4 describes this CDASH best practice as a section: "PE - Physical Examination", that describes a best practice for collecting physical examination data. Basically, any abnormalities would be recorded as medical history or adverse events, and depending on timing whether an exam was performed, it would be recorded by treating the exam as a procedure. There should be no PE dataset in SDTM at all. The below example annotations show how not to map physical exams in SDTMIG v3.3 (Image 3) and how to correctly map them (Image 4).

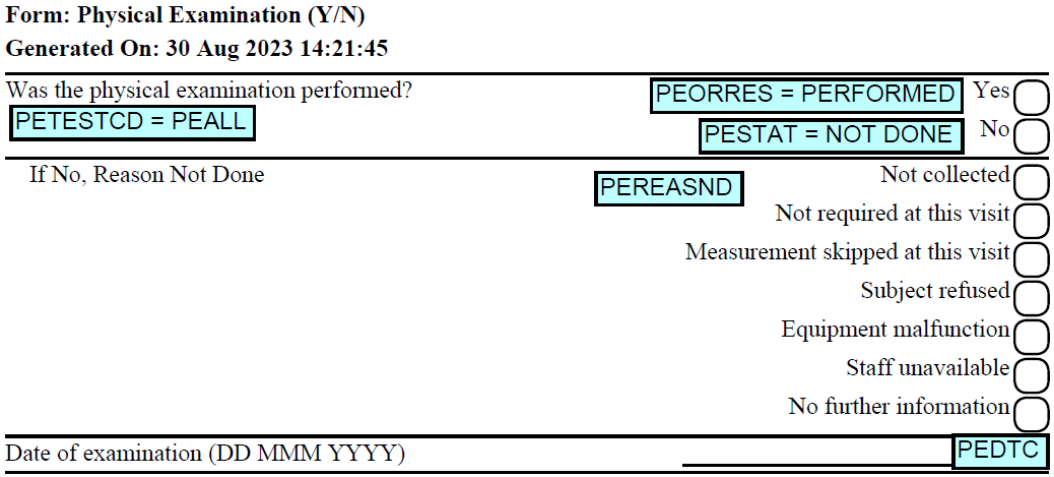

Image 3. How Not to Map PE in SDTMIG v3.3

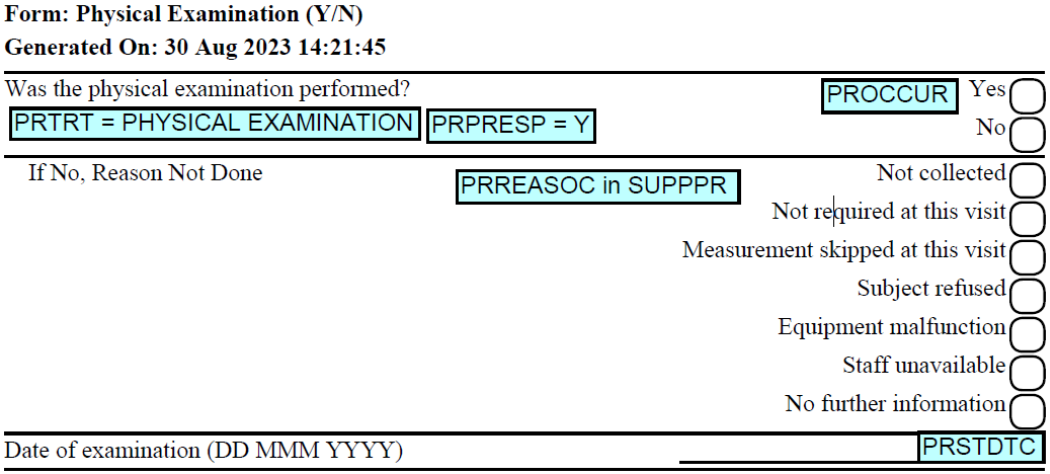

Image 4. How to Map Prompt Questions to PR in SDTMIG v3.3

MO VS. MORPHOLOGY/PHYSIOLOGY DOMAINS

Morphology (MO) versus morphology/physiology (or body system) domains (CV, NV, MK, OE, RE, RP, and UR): When the Morphology domain was introduced in SDTMIG v3.2, the CDISC planned to represent morphology and physiology findings in

separate domains. Morphology findings would be modeled in the MO domain and physiology findings in separate domains by body systems. Since then, the team found that separating morphology and physiology findings was more difficult than anticipated and provided little added value. This led to the decision create body system-based domains that cover both morphology and physiology findings, and to deprecate the single MO domain in a future version of the SDTMIG.

- The SDTMIG v3.3 includes several domains for physiology and morphology findings for different body systems. These differ only in body system, in domain code, and in informative content.
- All body system-based physiology/morphology domains share the same structure. Although time point is not in the structure, it can be included in the structure of a particular domain if time point variables were included in the data represented.
- CDISC controlled terminology includes codelists for TEST and TESTCD values for each body-system based domain. Codelist for MOTEST and MOTESTCD are removed from CT based on the decision to decommission MO from future versions.
- Although SDTMIG v3.3 has the MO domain, it has been deprecated in SDTMIG v3.4. Therefore, it is recommended to use body-system based domains and to not use MO if possible, when implementing SDTMIG v3.3.

ROW	STUDYID	DOMAIN	USUBJID	MOSEQ	MOTESTCD	MOTEST	MOORRES	MOORRESU	MOSTRESC	MOSTRESN	MOSTRESU	MOLOC
1	STUDY01	MO	232-P01	1	VOLUME	Volume	50	mL	50	50	mL	KIDNEY
2	STUDY01	MO	232-P01	2	VOLUME	Volume	100	mL	100	100	mL	LIVER
3	STUDY01	MO	232-P01	3	MASS	Mass	225	g	225	225	g	HEART,LEFT VENTRICLE

Table 4. SDTMIG v3.2 mo.xpt

ROW	STUDYID	DOMAIN	USUBJID	URSEQ	URTESTCD	URTEST	URORRES	URORRESU	URSTRESC	URSTRESN	URSTRESU	URLOC
1	STUDY01	UR	232-P01	1	VOLUME	Volume	50	mL	50	50	mL	KIDNEY
2	STUDY01	UR	232-P01	2	VOLUME	Volume	100	mL	100	100	mL	LIVER

Table 5. SDTMIG v3.3 ur.xpt

ROW	STUDYID	DOMAIN	USUBJID	CVSEQ	CVTESTCD	CVTEST	CVORRES	CVORRESU	CVSTRESC	CVSTRESN	CVSTRESU	CVLOC
1	STUDY01	CV	232-P01	1	MASS	Mass	225	g	225	225	g	HEART, LEFT VENTRICLE

Table 6. SDTMIG v3.3 cv.xpt

CHANGES IN DEFINE-XML V2.1 AT A GLANCE

Not to be outdone by the SDTM model and IG there are also a lot of changes in Define-XML v2.1 from v2.0. A variable's origin metadata has been expanded upon from v2.0 to contain greater detail. The def:Origin element now includes the Type and Source attributes in v2.1. Type

indicates how the data for the variable originated while Source identifies the party responsible for the data's origin. Type and Source are required attributes for SDTM datasets except when Type is Predecessor. In such cases, Source is not used. Controlled terminology for values in Origin (Type) and Source have been established (as seen in Table 7. Origin Type and Source Codelists). As with previous versions, Origin can be specified at the variable or value level as appropriate, the cardinality of def:Origin is one or more (meaning multiple values can be specified for a variable or value in a value list), and there are other business rules (mostly unchanged) for its use in the Define-XML specification. Below is a textual and rendered representation of Origin Type and Source (Image 5. Define-XML v2.1 LBORRES Origin/Source where LBTESTCD = 'HCT').

ItemDef Definitions

```
    <ItemDef OID="IT.LB.LBORRES.HCT.LBSPEC.BLOOD.VENDOR" Name="HCT" DataType="float" Length="4" SignificantDigits="2"
SASFieldName="HCT">
      <Description>
        <TranslatedText xml:lang="en">Hematocrit</TranslatedText>
      </Description>
      <def:Origin Type="Collected" Source="Vendor">
        <Description>
          <TranslatedText xml:lang="en">From Central lab (LB.LBNAM NE "LOCAL LAB")</TranslatedText>
        </Description>
      </def:Origin>
    </ItemDef>
    <ItemDef OID="IT.LB.LBORRES.HCT.LBSPEC.BLOOD.CRF" Name="HCT" DataType="float" Length="4" SignificantDigits="2"
SASFieldName="HCT">
      <Description>
        <TranslatedText xml:lang="en">Hematocrit</TranslatedText>
      </Description>
      <def:Origin Type="Collected" Source="Investigator">
        <Description>
          <TranslatedText xml:lang="en">From Local lab (LB.LBNAM="LOCAL LAB"). Note that the CRF page reference is given only for illustration
purposes. The sample acrf.pdf does not include the local lab CRF page.</TranslatedText>
        </Description>
        <def:DocumentRef leafID="LF.acrf">
          <def:PDFPageRef PageRefs="1" Type="PhysicalRef"/>
        </def:DocumentRef>
      </def:Origin>
    </ItemDef>
```

Variable	Where Condition	Label / Description	Type	Length or Display Format	Controlled Terms or ISO Format	Origin / Source / Method / Comment
LBORRES VLM		Result or Finding in Original Units	text	8		Origin specified at Value Level Metadata
	LBTESTCD = "HCT" (Hematocrit) and LBSPEC = "BLOOD" and LBNAM ≠ "LOCAL LAB"	Hematocrit	float	4		Collected (Source: Vendor) From Central lab (LB.LBNAM NE "LOCAL LAB")
	LBTESTCD = "HCT" (Hematocrit) and LBSPEC = "BLOOD" and LBNAM = "LOCAL LAB"	Hematocrit	float	4		Collected (Source: Investigator) From Local lab (LB.LBNAM="LOCAL LAB"). Note that the CRF page reference is given only for illustration purposes. The sample acrf.pdf does not include the local lab CRF page. Annotated CRF [1]

Image 5. Define-XML v2.1 LBORRES Origin/Source where LBTESTCD = 'HCT'

Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	CDISC Submission Value
C170450		No	Origin Source	ORIGINS
C25936	C170450		Origin Source	Investigator
C70793	C170450		Origin Source	Sponsor
C41189	C170450		Origin Source	Subject
C68608	C170450		Origin Source	Vendor
C170449		No	Origin Type	ORIGINT
C170547	C170449		Origin Type	Assigned
C170548	C170449		Origin Type	Collected
C170549	C170449		Origin Type	Derived
C126101	C170449		Origin Type	Not Available
C17649	C170449		Origin Type	Other

C170550	C170449		Origin Type	Predecessor
C170551	C170449		Origin Type	Protocol

Table 7. Origin Type and Source Codelists

Refer to the Define-XML v2,1 specification for business rules involving the use of Origin as well as section 4.3.2.1 “Origin for SDTM Datasets” which includes a table showing valid uses of the controlled terminology combined in Origin Type and Source.

Two other impactful changes in Define-XML v2.1 are the def:HasNoData and the def:IsNonStandard attributes. The def:HasNoData attribute can be used conditionally for both the ItemRef and ItemGroupDef elements when either have no data (the former for an all null variable and the latter for an empty/missing dataset). The previous example for all null permissible Adverse Events variables (intended to be collected) shows the use of def:HasNoData (as rendered in Image 2. Define-XML v2.1 ae.xpt HasNoData). Note that a comment is required for each use of the def:HasNoData = “Yes” in the define.xml.

The def:IsNonStandard attribute, for ItemRef, ItemGroupDef, and CodeList elements, is used to identify contents (variables, datasets, and/or codelists) that are non-standard. A dataset is considered non-standard if it is either a sponsor-defined custom domain or a domain based on an unpublished draft of a CDISC dataset. If a dataset variable is not part of the def:Standard referenced by the dataset, it is also to be considered non-standard for regulatory submission purposes; even in the case where the variable is defined in a different version of the same standard. Sponsor-defined codelists referenced in the define.xml should make use of def:IsNonStandard. An example of such is shown below (Image 6. Define-XML v2.1 DSSCAT Non-Standard Codelist).

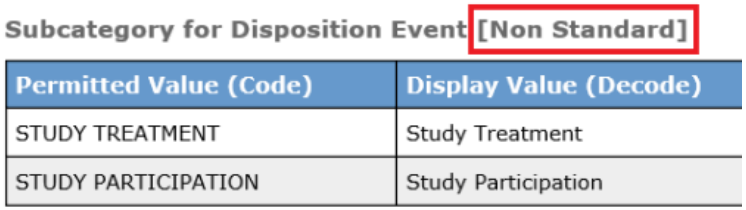

Image 6. Define-XML v2.1 DSSCAT Non-Standard Codelist

Other changes include the dataset Class attribute being re-implemented as a child element to an ItemGroup with support for the definition of Subclass underneath Class (for future consideration/utility in SDTM). A full description of the changes in Define-XML v2.1 from Define-XML v2.0 can be found in Define-XML v2.1 Section 1.1.3 “Relationship to Prior Define-XML Specifications”.

VLM FOR MULTIPLE CODELISTS/USE CASES

With SDTMIG v3.3, there are now representations of multiple codelists for single variables. It makes sense to have separate codelists for different use cases (usually different CRF forms) to control responses/display discrete value lists depending on the use. DSDECOD is a great example of such cases with NCOMPLT (Completion/Reason for Non-Completion) and PROTMLST (Protocol Milestone) codelists used for different disposition events for different forms (as seen in Images 7 through 9 below). Perhaps the only issue with representing the data this way in a submission is that the Define-XML standard’s CodeListRef Element has a usage cardinality of one (as seen in Image 10). This means we simply cannot specify two codelists as child elements for 1 ItemDef for a variable. Instead, we need to create 1 ValueListDef with 2 ItemRefs, 2 WhereClauseDefs, and

3 ItemDefs (2 to relate the 2 codelists to the 1 variable as VLM, and 1 for the remaining variable metadata) . The example text and rendered define.xml (Image 11) below show the conventions used for DSDECOD. This has practical application for other variables and domains as well (e.g., Oncology versus Clinical Classifications use cases in RS with variables RSCAT, RSTESTCD, RSTEST, RSSTRESC having multiple codelists for each variable).

DSDECOD	Standardized Disposition Term	Char	(NCOMPLT)(PROTMLST)	Synonym Qualifier	Controlled terminology for the name of disposition event or protocol milestone. Examples of protocol milestones: "INFORMED CONSENT OBTAINED", "RANDOMIZED". There are separate codelists used for DSDECOD where the choice depends on the value of DSCAT. Codelist "NCOMPLT" is used for disposition events and codelist "PROTMLST" is used for protocol milestones. The variable may be	Req

Image 7. SDTMIG v3.3 DSDECOD Specification

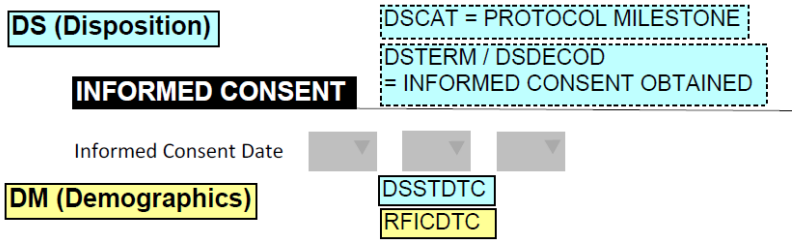

Image 8. Example Informed Consent (DS) Form

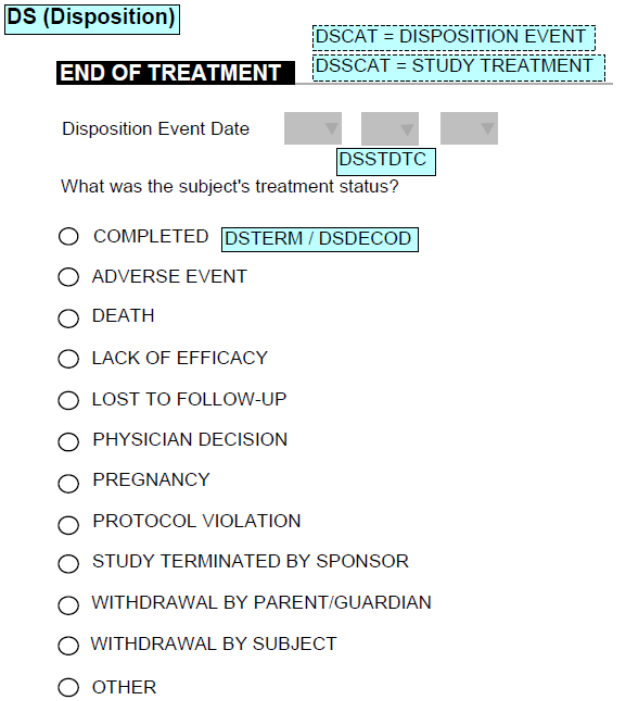

Image 9. Example End of Treatment (DS) Form

5.3.12.1 CodeListRef Element

Provides a reference to the Codelist element defining controlled terminology if applicable.

Element Name	CodeListRef
Element XPath(s)	/ODM/Study/MetaDataVersion/ItemDef/CodeListRef
Element Textual Value	None
Usage	• Requirement: Optional • Cardinality: One

Image 10. Define-XML v2.1 CodeListRef Cardinality

DSDECOD VLM		Standardized Disposition Term	text	Synonym Qualifier	29		
	DSSCAT ≠ ""	Standardized Disposition Term	text		29	Completion/Reason for Non-Completion [13 Terms]	Collected (Source: Investigator) Annotated CRF [27 28]
	DSSCAT = ""	Standardized Disposition Term	text		29	Protocol Milestone • "INFORMED CONSENT OBTAINED" = "Informed Consent"	Assigned (Source: Sponsor) Annotated CRF [5]

Image 11. Example Define.xml Render for DSDECOD VLM

Excerpts of the XML used to describe the VLM, where clauses, and variable metadata for DSDECOD:

ValueList Definitions

```
<def:ValueListDef OID="VL.DSDECOD">
  <ItemRef ItemOID="IT.DS.DSDECOD.3" OrderNumber="1" Mandatory="No">
    <def:WhereClauseRef WhereClauseOID="WC.DSDECOD1"/>
  </ItemRef>
  <ItemRef ItemOID="IT.DS.DSDECOD.4" OrderNumber="2" Mandatory="No">
    <def:WhereClauseRef WhereClauseOID="WC.DSDECOD2"/>
  </ItemRef>
</def:ValueListDef>
```

WhereClause Definitions

```
<def:WhereClauseDef OID="WC.DSDECOD1">
  <RangeCheck Comparator="NE" SoftHard="Soft" def:ItemOID="IT.DS.DSSCAT">
    <CheckValue></CheckValue>
  </RangeCheck>
</def:WhereClauseDef>
<def:WhereClauseDef OID="WC.DSDECOD2">
  <RangeCheck Comparator="EQ" SoftHard="Soft" def:ItemOID="IT.DS.DSSCAT">
    <CheckValue></CheckValue>
  </RangeCheck>
</def:WhereClauseDef>
```

ItemGroup Definitions

```
<!-- Dataset Definition (DS) -->
<ItemGroupDef OID="IG.DS" Name="DS" Domain="DS"
```

```
      Repeating="Yes" IsReferenceData="No" SASDatasetName="DS"
      def:Structure="One record per disposition status or protocol milestone per subject"
      Purpose="Tabulation" def:StandardOID="STD.1"
 def:ArchiveLocationID="LF.DS">
      <Description>
        <TranslatedText xml:lang="en">Disposition</TranslatedText>
      </Description>
      <ItemRef ItemOID="IT.DS.STUDYID" Mandatory="Yes" OrderNumber="1" KeySequence="1" Role="Identifier"/>
      <ItemRef ItemOID="IT.DS.DOMAIN" Mandatory="Yes" OrderNumber="2" Role="Identifier"/>
      <ItemRef ItemOID="IT.DS.USUBJID" Mandatory="Yes" OrderNumber="3" KeySequence="2" Role="Identifier"/>
      <ItemRef ItemOID="IT.DS.DSSEQ" Mandatory="Yes" OrderNumber="4" MethodOID="MT.SEQ" Role="Identifier"/>
      <ItemRef ItemOID="IT.DS.DSLNKID" Mandatory="No" OrderNumber="5" Role="Identifier"/>
      <ItemRef ItemOID="IT.DS.DSTERM" Mandatory="Yes" OrderNumber="6" Role="Topic"/>
      <ItemRef ItemOID="IT.DS.DSDECOD" Mandatory="Yes" OrderNumber="7" KeySequence="6" Role="Synonym Qualifier"/>
      <ItemRef ItemOID="IT.DS.DSCAT" Mandatory="No" OrderNumber="8" KeySequence="4" Role="Grouping Qualifier"/>
      <ItemRef ItemOID="IT.DS.DSSCAT" Mandatory="No" OrderNumber="9" KeySequence="5" Role="Grouping Qualifier"/>
      <ItemRef ItemOID="IT.DS.EPOCH" Mandatory="No" OrderNumber="10" MethodOID="MT.EPOCH" Role="Timing"/>
      <ItemRef ItemOID="IT.DS.DSSTDTC" Mandatory="No" OrderNumber="11" KeySequence="3" Role="Timing"/>
      <ItemRef ItemOID="IT.DS.DSSTDY" Mandatory="No" OrderNumber="12" MethodOID="MT.DAYCALC" Role="Timing"/>
      <def:Class Name="EVENTS"/>
      <def:leaf ID="LF.DS" xlink:href="ds.xpt">
        <def:title>ds.xpt</def:title>
      </def:leaf>
    </ItemGroupDef>

ItemDef Definitions
 <ItemDef OID="IT.DS.DSDECOD" Name="DSDECOD" DataType="text" Length="29" SASFieldName="DSDECOD">
      <Description>
        <TranslatedText xml:lang="en">Standardized Disposition Term</TranslatedText>
      </Description>
      <def:ValueListRef ValueListOID="VL.DSDECOD"/>
    </ItemDef>
 <ItemDef OID="IT.DS.DSDECOD.3" Name="DSDECOD" DataType="text" Length="29" SASFieldName="DSDECOD">
      <Description>
        <TranslatedText xml:lang="en">Standardized Disposition Term</TranslatedText>
      </Description>
      <CodeListRef CodeListOID="CL.NCOMPLT"/>
      <def:Origin Type="Collected" Source="Investigator">
        <def:DocumentRef leafID="LF.acrf">
          <def:PDFPageRef PageRefs="27 28" Type="PhysicalRef"/>
```

```
    </def:DocumentRef>
  </def:Origin>
</ItemDef>
<ItemDef OID="IT.DS.DSDECOD.4" Name="DSDECOD" DataType="text" Length="29" SASFieldName="DSDECOD">
  <Description>
    <TranslatedText xml:lang="en">Standardized Disposition Term</TranslatedText>
  </Description>
  <CodeListRef CodeListOID="CL.PROTMLST"/>
  <def:Origin Type="Assigned" Source="Sponsor">
    <def:DocumentRef leafID="LF.acrf">
      <def:PDFPageRef PageRefs="5" Type="PhysicalRef"/>
    </def:DocumentRef>
  </def:Origin>
</ItemDef>
```

CONFORMANCE/VALIDATION

When performing validation of SDTM v3.3 datasets, we advise you use the SDTM and SDTMIG Conformance Rules v2.0 Rules for. There have been rules added, modified, and deleted for SDTMIG v3.3 in the changes from Conformance Rules v1.1 to v2.0. Conformance Rules for Define-XML 2.1 were released in March of 2021 and should be used to validate the define.xml alone and with the data for conformance. Use the rules engine/regulatory agency set of validation/business rules as applicable for your submission.

CONCLUSION

We have shared some of our practical experiences and insights working with SDTMIG v3.3 and Define-XML v2.1. We hope these best practices and resources can be used by others to successfully upversion their SDTM and Define-XML standards. Provided in this paper are effective strategies for dealing with a lot of the larger and more subtle changes within these versions. There are additional changes throughout SDTMIG v3.3 and Define-XML v2.1 that we could not cover and we encourage users of those guidances to see the references provided below. Please feel free to contact the authors with any questions related to this paper.

REFERENCES

SDTMIG v3.3

SDTM v1.7

SDTMIG for Medical Devices v1.1

SDTMIG-AP v1.0

SDTM Metadata Submission Guidelines v2.0

Conformance Rules v1.1 for SDTMIG v3.2 and v3.3

SDTM and SDTMIG Conformance Rules v2.0

Knowledge Base Article - Subject Visits and COVID-19

Guidance for Ongoing Studies Disrupted by COVID-19

Knowledge Base Article - How Should I Represent Whether Physical Exam Was Performed In SDTM

Define-XML v2.1

Conformance Rules for Define-XML v2.1

CDISC Controlled Terminology

CDISC Library

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

David Neubauer
IQVIA
david.neubauer@iqvia.com

Kapila Patel
IQVIA
kapila.patel@iqvia.com

Toral Patel
IQVIA
toral.patel@iqvia.com

Soumya Rajesh
IQVIA
soumya.rajesh@iqvia.com