



Submission Documentation

Define-XML 2.0 What can it do for you?

Sally Cassells



Next Step Clinical Systems LLC

Define-XML 2.0

- Provides an important new tool for communicating information about submission data to reviewers
- Major upgrade (not backward compatible)
 - Machine readable value level metadata
 - Improved support for traceability, links to documents, derivation documentation about
 - Improved support for CDISC Controlled Terminology



WHAT DOES DEFINE-XML DO?

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, ETC	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, TSPARMCD, 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 Demographics Reviewers Guide
SE	Subject Elements	SPECIAL PURPOSE	One record per actual Element per subject	Tabulation	STUDYID, USUBJID, SESTDTC, SEENDTC, TAETORD, ETC	se.xpt	
SV	Subject Visits	SPECIAL PURPOSE	One record per actual visit per subject	Tabulation	STUDYID, USUBJID, SVSTDTC, VISITNUM	sv.xpt	
CM	Concomitant Medications	INTERVENTIONS	One record per recorded medication occurrence or constant-dosing interval per subject	Tabulation	STUDYID, USUBJID, CMSTDTC, CMENDTC, CMCAT, CMTRT, CMDOSTXT,	cm.xpt	

Home - Clinical Metadata Re... Home - Clinical Metadata Re... Overview (Java Platform SE... Suggested Sites Web Slice Gallery							
2.0	NOTES FOR THE TABULATION DATASETS [Instructional, to be removed: Any domain specific information that may be helpful to the reviewer can be described within this section.]						
2.1	DEMOGRAPHICS (DM) Two subjects in the Demographics (DM) domain were randomized but did not receive study drug. These subjects were assigned a treatment arm, but do not have any records in the Exposure (EX) domain. This will cause a HIGH severity error. A list of the subjects and treatment arms are in Table 3–1. Table 3–1. Subjects Randomized and Did Not Receive Drug <table> <tr> <th>Subject</th><th>Treatment Arm</th></tr> <tr> <td>STUDY01-10001</td><td>XXXCD</td></tr> <tr> <td>STUDY01-10010</td><td>YYYCD</td></tr> </table>	Subject	Treatment Arm	STUDY01-10001	XXXCD	STUDY01-10010	YYYCD
Subject	Treatment Arm						
STUDY01-10001	XXXCD						
STUDY01-10010	YYYCD						

	STUDYID	DOMAIN	USUBJID	QSSEQ	QSTESTCD	QSTEST	QSCAT	QSSCAT	QSORRES	QSSTRESC	QSSTRESN
1	CDISC01	QS	CDISC01.1	3	CSDD01	Anxiety	CORNELL S	Mood - Re	Severe	2	2
2	CDISC01	QS	CDISC01.1	4	CSDD01	Anxiety	CORNELL S	Mood - Re	Severe	2	2
3	CDISC01	QS	CDISC01.1	5	CSDD01	Anxiety	CORNELL S	Mood - Re	Mild or I	1	1
4	CDISC01	QS	CDISC01.1	6	CSDD01	Anxiety	CORNELL S	Mood - Re	Mild or I	1	1
5	CDISC01	QS	CDISC01.1	7	CSDD02	Sadness	CORNELL S	Mood - Re	Severe	2	2
6	CDISC01	QS	CDISC01.1	8	CSDD02	Sadness	CORNELL S	Mood - Re	Mild or I	1	1
7	CDISC01	QS	CDISC01.1	9	CSDD02	Sadness	CORNELL S	Mood - Re	Absent	0	0
8	CDISC01	QS	CDISC01.1	10	CSDD02	Sadness	CORNELL S	Mood - Re	Absent	0	0
9	CDISC01	QS	CDISC01.1	11	CSDD03	Lack of R	CORNELL S	Mood - Re	Absent	0	0
10	CDISC01	QS	CDISC01.1	12	CSDD03	Lack of R	CORNELL S	Mood - Re	Absent	0	0
11	CDISC01	QS	CDISC01.1	13	CSDD03	Lack of R	CORNELL S	Mood - Re	Absent	0	0
12	CDISC01	QS	CDISC01.1	14	CSDD03	Lack of R	CORNELL S	Mood - Re	Absent	0	0
13	CDISC01	QS	CDISC01.1	15	CSDD04	Irritabil	CORNELL S	Mood - Re	Mild or I	1	1
14	CDISC01	QS	CDISC01.1	16	CSDD04	Irritabil	CORNELL S	Mood - Re	Mild or I	1	1
15	CDISC01	QS	CDISC01.1	17	CSDD04	Irritabil	CORNELL S	Mood - Re	Absent	0	0
16	CDISC01	QS	CDISC01.1	18	CSDD04	Irritabil	CORNELL S	Mood - Re	Mild or I	1	1
17	CDISC01	QS	CDISC01.1	19	CSDD05	Agitation	CORNELL S	Behaviora	Mild or I	1	1

Questionnaire-QSCS (QSCS) [Location: [qscs.xpt](#)]

Variable	Label	Key	Type	Length	Controlled Terms or Format	Origin	Derivation/Comment
STUDYID	Study Identifier	1	text	7		Protocol	
DOMAIN	Domain Abbreviation		text	2	["QS" = "Questionnaires"] < Domain Abbreviation (QS) >	Assigned	
USUBJID	Unique Subject Identifier	2	text	14		Derived	Concatenation of STUDYID and SUBJID
QSSEQ	Sequence Number		integer	3		Derived	Sequential number identifying records within each USUBJID in the domain.
QSTESTCD	Question Short Name	4	text	7	QS Test Codes	Assigned	
QSTEST	Question Name		text	37	QS Tests	CRF Page 17	
QSCAT	Category of Question	3	text	50		CRF Page 17	
QSSCAT	Subcategory for Question		text	40		CRF Pages 14 - 15	
QSORRES	Finding in Original Units		text	20		CRF Page 17	
QSSTRESC	Character Result/Finding in Std Format		text	2		Derived	If questionnaire result is a numerically relevant code then code stored here; else character result equal to QSORRES. Character results contain mixed case data.
QSSTRESN	Numeric Finding in Standard Units		integer	2		Derived	When QSORRES is numeric then QSSTRESN equal to numeric result

file:///C:/Users/Administrator/Desktop/Define-XML-2-0/sdtm/blankcrf.pdf#page=17 - Windows Internet Explorer

file:///C:/Users/Administrator/Desktop/Define-XML-2-0/sdtm/blankcrf.pdf#page=17

File Edit Go to Favorites Help

Home - Clinical MetaData Re... Home - Clinical MetaData Re... Overview (Java Platform SE... Suggested Sites Web Slice Gallery

Tools Sign Comment

14 / 22 122%

Bookmarks

- Domains
 - 12-Lead ECG
 - Adverse Events
 - Clinical Global Impression
 - Cornell Scale for Depression in Dementia**
 - Demography
 - Enrollment
 - Inclusion Criteria
 - Informed Consent
 - Exclusion Criteria
 - Laboratory
 - Medication Record
 - Medical and Surgical History
 - Mini-Mental State Examination
 - Screening
 - Week 24
 - Physical Examination
 - Prior/Concomitant

CDISC

Study CDISC01

Assessment Date **QSDTC**

CORNELL SCALE FOR DEPRESSION IN DEMENTIA (CSDD) (PAGE 1 OF 2)

QSCAT **RTRINIT in SUPPQS**

Instructions: For each item, select the "cue" which best characterizes the symptoms and signs occurring during the past week, and circle the appropriate number. **QSEVLINT**

Rater's Initials:

QSSCAT **A. MOOD - RELATED SIGNS**

	ABSENT	CUE MILD OR INTERMITTENT	SEVERE
1. ANXIETY QSTEST Anxious expression, ruminations, worrying	0 QSORRES when QSTESTCD = CSDD01	1	2
2. SADNESS Sad expression, sad voice, tearfulness	0 QSORRES when QSTESTCD = CSDD02	1	2
3. LACK OF REACTIVITY TO PLEASANT EVENTS	0 QSORRES when QSTESTCD = CSDD03	1	2
4. IRRITABILITY Easily annoyed, short tempered	0 QSORRES when QSTESTCD = CSDD04	1	2
B. BEHAVIORAL DISTURBANCE			
5. AGITATION Restlessness, handwringing, hairpulling	0 QSORRES when QSTESTCD = CSDD05	1	2
6. RETARDATION Slow movements, slow speech, slow reactions	0 QSORRES when QSTESTCD = CSDD06	1	2

ECG Test Results (EG) [Location: [eg.xpt](#)]

Variable	Label	Key	Type	Length	Controlled Terms or Format	Origin	Derivation/Comment
STUDYID	Study Identifier	1	text	7		Protocol	
DOMAIN	Domain Abbreviation		text	2	["EG" = "Electrocardiogram"] < Domain Abbreviation (EG) >	Assigned	
USUBJID	Unique Subject Identifier	2	text	14		Derived	Concatenation of STUDYID and SUBJID
EGSEQ	Sequence Number		integer	2		Derived	Sequential number identifying records within each USUBJID in the domain.
EGSPID	Sponsor-Defined Identifier		text	6		Assigned	ECG parameter ordering variable
EGTESTCD	ECG Test or Examination Short Name	3	text	6	ECG Test Code	Assigned	
EGTEST	ECG Test or Examination Name		text	40	ECG Test Name	CRF Page 12	
EGPOS	ECG Position of Subject		text	6		Protocol	Equal to SUPINE
EGORRES	Result or Finding in Original Units		text	15		CRF Page 12	
EGORRESU	Original Units		text	4	["BEATS/MIN" = "Beats per Minute", "msec" = "Millisecond"] < Unit (EGRESU) >	CRF Page 12	
EGSTRESC	Character Result/Finding in Std Format		text	15		Derived	Data collected in non-standard units is converted using standard conversion factors to standard units.
EGSTRESN	Numeric Result/Finding in Standard Units		float	5		Derived	EGSTRESN = numeric value of EGSTRESC, when EGSTRESC contains numeric data.
EGSTRESU	Standard Units		text	9	["BEATS/MIN" = "Beats per Minute", "msec" = "Millisecond"] < Unit (EGRESU) >	Assigned	

ECG Test Code [CL.EGTESTCD, C71153]

Permitted Value (Code)	Display Value (Decode)
INTP [C41255]	Interpretation
PRMEAN [C62086]	Summary (Mean) PR Duration
QRS DUR [C62087]	Summary (Mean) QRS Duration
QTcB [C62112]	QTcB - Bazett's Correction Formula
QTcF [C62113]	QTcF - Fridericia's Correction Formula
QTMEAN [C62089]	Summary (Mean) QT Duration
VRMEAN [C62152]	Summary (Mean) Ventricular Rate

Value Level Metadata - EG [EGSTRESC]

Variable	Where	Type	Length / Display Format	Controlled Terms or Format	Origin	Derivation/Comment
EGSTRESC	EGTESTCD EQ QTcB (QTcB - Bazett's Correction Formula)	float	5.1		Derived	QTcB = QT interval / square root of (60 / heart rate). For the complete algorithm see the referenced external document. Complex Algorithms (complexalgorithms.pdf)
EGSTRESC	EGTESTCD EQ QTcF (QTcF - Fridericia's Correction Formula)	float	5.1		Derived	QTcF = QT interval / cubic root of (60 / heart rate). For the complete algorithm see the referenced external document. Complex Algorithms (complexalgorithms.pdf)

Value Level Metadata - VS [VSORRES]

Variable	Where	Type	Length / Display Format	Controlled Terms or Format	Origin	Derivation/Comment
VSORRES	VSTESTCD EQ DIABP (Diastolic Blood Pressure)	integer	2		CRF Page 11	
VSORRES	VSTESTCD EQ FRMSIZE (Body Frame Size)	txt	6	["LARGE", "MEDIUM", "SMALL"] < Size >	CRF Page 11	
VSORRES	VSTESTCD EQ HEIGHT (Height)	float	5.1		CRF Page 11	
VSORRES	VSTESTCD EQ PULSE (Pulse Rate)	integer	2		CRF Page 11	
VSORRES	VSTESTCD EQ SYSBP (Systolic Blood Pressure)	integer	3		CRF Page 11	
VSORRES	VSTESTCD EQ WEIGHT (Weight)	float	5.1		CRF Page 11	

Value Level Metadata - SUPPDM [QVAL]

Variable	Where	Type	Length / Display Format	Controlled Terms or Format	Origin	Derivation/Comment
QVAL	QNAM EQ RACE1 (Race 1)	text	5	Race	CRF Page 6	
QVAL	QNAM EQ RACE2 (Race 2)	text	25	Race	CRF Page 6	
QVAL	QNAM EQ RACE3 (Race 3)	text	5	Race	CRF Page 6	
QVAL	QNAM EQ RACEOTH (Race, Other)	text	14		CRF Page 6	
QVAL	QNAM EQ RAND (Randomized Population Flag)	text	1	["N" = "No", "U" = "Unknown", "Y" = "Yes"] < No Yes Response Subset >	CRF Page 16	
QVAL	QNAM EQ RANDNO (Randomization Number)	text	4		CRF Page 16	
QVAL	QNAM EQ SAFETY (Safety Population Flag)	text	1	["N" = "No", "U" = "Unknown", "Y" = "Yes"] < No Yes Response Subset >	Derived	SAFETY = "Y" for randomized subjects who took at least one dose study medication. Null otherwise.

Analysis Datasets for Study CDISC-Sample (ADaM-IG 1.0)

Dataset	Description	Class	Structure	Purpose	Keys	Location	Documentation
ADSL	Subject-Level Analysis	SUBJECT LEVEL ANALYSIS DATASET	one record per subject	Analysis	USUBJID	adsl.xpt	Screen Failures are excluded since they are not needed for this study analysis
ADQSADAS	ADAS-Cog Analysis	BASIC DATA STRUCTURE	One record per subject per parameter per analysis visit per analysis date	Analysis	USUBJID, PARAMCD, AVISIT, ADT	adqsadas.xpt	See referenced dataset creation program and Analysis Data Reviewer's Guide, Section 2.1 adqsadas.sas Analysis Data Reviewer's Guide

	AVISITN	VISIT	VISITNUM	ADY	ADT	PARAM	PARAMCD	PARAMN	AVAL	BASE	CHG
1	0	BASELINE	3	1	05AUG2012	Word Reca	ACITM01	1	4	4	.
2	8	WEEK 4	5	29	02SEP2012	Word Reca	ACITM01	1	2	4	-2
3	24	RETRIEVAL	201	198	18FEB2013	Word Reca	ACITM01	1	3	4	-1
4	0	BASELINE	3	1	05AUG2012	Naming Ob	ACITM02	2	2	2	.
5	8	WEEK 4	5	29	02SEP2012	Naming Ob	ACITM02	2	2	2	0
6	24	RETRIEVAL	201	198	18FEB2013	Naming Ob	ACITM02	2	0	2	-2
7	0	BASELINE	3	1	05AUG2012	Delayed W	ACITM03	3	2	2	.
8	8	WEEK 4	5	29	02SEP2012	Delayed W	ACITM03	3	2	2	0
9	24	RETRIEVAL	201	198	18FEB2013	Delayed W	ACITM03	3	3	2	1
10	0	BASELINE	3	1	05AUG2012	Commands	ACITM04	4	0	0	.
11	8	WEEK 4	5	29	02SEP2012	Commands	ACITM04	4	0	0	0
12	24	RETRIEVAL	201	198	18FEB2013	Commands	ACITM04	4	0	0	0
13	0	BASELINE	3	1	05AUG2012	Construct	ACITM05	5	3	3	.
14	8	WEEK 4	5	29	02SEP2012	Construct	ACITM05	5	2	3	-1
15	24	RETRIEVAL	201	198	18FEB2013	Construct	ACITM05	5	4	3	1
16	0	BASELINE	3	1	05AUG2012	Ideationa	ACITM06	6	0	0	.
17	8	WEEK 4	5	29	02SEP2012	Ideationa	ACITM06	6	0	0	0
18	24	RETRIEVAL	201	198	18FEB2013	Ideationa	ACITM06	6	0	0	0

ADAS-Cog Analysis (ADQSADAS) [Location: [adqsadas.xpt](#)]

Variable	Label	Type	Length / Display Format	Controlled Terms or Format	Source/Derivation/Comment
VISIT	Visit Name	text	19	VISIT	Predecessor: QS.VISIT
VISITNUM	Visit Number	float	8	VISITNUM	Predecessor: QS.VISITNUM
ADY	Analysis Relative Day	integer	8		Derived: ADY = ADT - TRTSDT + 1, if ADT > TRTSDT. ADY = ADT - TRTSDT, if ADT < TRTSDT.
ADT	Analysis Date	integer	8		Derived: SAS date from QS.QSDTC
PARAM	Parameter	text	100	PARAM_ADQSADAS	Assigned:
PARAMCD	Parameter Code	text	8	PARAMCD_ADQSADAS	Assigned:
PARAMN	Parameter (N)	integer	8	PARAMN_ADQSADAS	Assigned: Assign a numeric code for each value of PARAMCD (see codelist PARAMN_ADQSADAS)
AVAL	Analysis Value	integer	8		Derivations are described per parameter in the parameter value level metadata
BASE	Baseline Value	integer	8		Derived: QS.QSSTRESN when QS.QSBLFL=Y (QS.VISITNUM=3)
CHG	Change from Baseline	integer	8		Derived: AVAL - BASE
PCHG	Percent Change from Baseline	integer	8		Derived: 100* (CI IG/BASE)
ABLFL	ABLFL	text	1	["Y" = "Yes"] < No Yes Response - Y subset >	Predecessor: QS.QSBLFL
ANL01FL	Analysis Record Flag 01	text	1	["Y" = "Yes"] < No Yes Response - Y subset >	Derived: If multiple visits fall into the same visit window, then the one closest to the target date is chosen for analysis. These are flagged with ANL01FL="Y".
DTYPE	Derivation Type	text	7	["LOCF" = "Last Observation Carried Forward"] < Derivation Type >	

PARAMCD_ADQSADAS [CL.PARAMCD_ADQSADAS]

Permitted Value (Code)	Display Value (Decode)
ACITM01	Word Recall Task
ACITM02	Naming Objects And Fingers (Refer To 5 C
ACITM03	Delayed Word Recall
ACITM04	Commands
ACITM05	Constructional Praxis
ACITM06	Ideational Praxis
ACITM07	Orientation
ACITM08	Word Recognition
ACITM09	Attention/Visual Search Task
ACITM10	Maze Solution
ACITM11	Spoken Language Ability
ACITM12	Comprehension Of Spoken Language
ACITM13	Word Finding Difficulty In Spontaneous Speech
ACITM14	Recall Of Test Instructions
ACTOT	Adas-Cog(11) Subscore

ADAS-Cog Analysis (ADQSADAS) [Location: [adqsadas.xpt](#)]

Variable	Label	Type	Length / Display Format	Controlled Terms or Format	Source/Derivation/Comment
VISIT	Visit Name	text	19	VISIT	Predecessor: QS.VISIT
VISITNUM	Visit Number	float	8	VISITNUM	Predecessor: QS.VISITNUM
ADY	Analysis Relative Day	integer	8		Derived: ADY = ADT - TRTSDT + 1, if ADT > TRTSDT. ADY = ADT - TRTSDT, if ADT < TRTSDT.
ADT	Analysis Date	integer	8		Derived: SAS date from QS.QSDTC
PARAM	Parameter	text	100	PARAM_ADQSADAS	Assigned:
PARAMCD	Parameter Code	text	8	PARAMCD_ADQSADAS	Assigned:
PARAMN	Parameter (N)	integer	8	PARAMN_ADQSADAS	Assigned: Assign a numeric code for each value of PARAMCD (see codelist PARAMN_ADQSADAS)
AVAL	Analysis Value	integer	8		Derivations are described per parameter in the parameter value level metadata
BASE	Baseline Value	integer	8		Derived: QS.QSSTRESN when QS.QSBLFL=Y (QS.VISITNUM=3)
CHG	Change from Baseline	integer	8		Derived: AVAL - BASE
PCHG	Percent Change from Baseline	integer	8		Derived: 100* (CIIG/BASE)
ABLFL	ABLFL	text	1	["Y" = "Yes"] < No Yes Response - Y subset >	Predecessor: QS.QSBLFL
ANL01FL	Analysis Record Flag 01	text	1	["Y" = "Yes"] < No Yes Response - Y subset >	Derived: If multiple visits fall into the same visit window, then the one closest to the target date is chosen for analysis. These are flagged with ANL01FL="Y".
DTYPE	Derivation Type	text	7	["LOCF" = "Last Observation Carried Forward"] < Derivation Type >	

Parameter Value List - ADQSADAS [AVAL]

Variable	Where	Type	Length / Display Format	Controlled Terms or Format	Origin	Derivation/Comment
AVAL	PARAMCD NOTIN ("ACTOT" (Adas-Cog(11) Subscore))	integer	8		Derived	QS.QSSTRESN where QSTESTCD=PARAMCD
AVAL	PARAMCD EQ ACTOT (Adas-Cog(11) Subscore)	integer	8		Derived	Sum of ADAS scores for items 1, 2, 4, 5, 6, 7, 8, 11, 12, 13, and 14, see Analysis Data Reviewers Guide (Page 3) for details on adjusting for missing values. Analysis Data Reviewer's Guide (analysis-data-reviewers-guide.pdf)

Parameter Value List - ADQSADAS [DTYPE]

Variable	Where	Type	Length / Display Format	Controlled Terms or Format	Origin	Derivation/Comment
DTYPE	PARAMCD NOTIN ("ACTOT" (Adas-Cog(11) Subscore))	text	10	["LOCF" = "Last Observation Carried Forward"] < Derivation Type >	Assigned	Value: null
DTYPE	PARAMCD EQ ACTOT (Adas-Cog(11) Subscore)	text	10	["LOCF" = "Last Observation Carried Forward"] < Derivation Type >	Assigned	Value: LOCF denotes that the LOCF imputation method was used to impute the value for the given parameter and analysis visit.

Summary

- Define-XML 2.0 provides new capabilities
 - Communicating more metadata details
 - Expressing SDTM and ADaM concepts
 - Providing smoother links with external documents
- Define-XML 2.0 is preferred by the FDA
 - No need to send Define-XML.PDF