Summing up - SDTM Trial Summary Domain

Yi Liu Celerion Inc., Lincoln, NE (USA)
Jenny Erskine Celerion Inc., Belfast, Northern Ireland
Stephen Read Celerion Inc., Belfast, Northern Ireland

ABSTRACT

Trial Design Model (TDM) datasets, specifically the Trial Summary Information (TS) domain, can provide a challenge for SAS programmers and SDTM reviewers based on the complexity of the study design and classification of study related characteristics as detailed in protocols, Statistical Analysis Plans (SAP), study related data or other supporting reference dictionaries and specifications. This paper gives a brief background and overview of key TS domain requirements and discusses some practical guidance in support of the TS domain build, its population and subsequent review. The paper also includes some recommendations around clearer definition of TS parameters in line with more recent SDTM and OpenCDISC validation enhancements and considerations around support of configurable data files that can be used to help drive the appropriate definition and documentation of TS data across a variety of protocols.

INTRODUCTION

The Trial Design Model describes the planned events of a clinical trial. As listed in Table 1, the Model includes Trial Arms, Trial Elements, Trial Visit, Trial Inclusion/Exclusion Criteria, Trial Summary, and Trial Disease Assessment, a new addition to SDTM-IG 3.2, In a standardized, clear, and concise manner, these datasets enable the reviewer to compare a subject or patient's progress in a clinical trial to the planned schedule of events.

Category	Description	Purpose	Domain
Trial Design	Describe the study characteristics and lists key facts (parameters) about the trial that are likely to appear in a registry of clinical Trials so that the reviewers have better understanding about the study design.	1. Clearly and quickly grasp the design of a clinical trial 2. Compare the designs of different trials 3. Search a data warehouse for clinical trials with certain features 4. Compare planned and actual treatments and visits for subjects in a clinical trial.	TA (Trial Arms), TE (Trial Elements), TV (Trial Visits), TI (Trial Inclusion/Exclusion Criteria), TS (Trial Summary), TD (Trial Disease Assessment)

Table 1. SDTM Trial Design Model Brief Description[1]

TRIAL SUMMARY DOMAIN STRUCTURE

As part of the Trial Design Model, the Trial Summary domain provides a high-level overview of key facts associated with a trial, recoding basic information such as study title, study design, study interventions, assessments, trial objectives etc. The domain is structured as follows: Study Identifier (STUDYID), Domain Abbreviation (DOMAIN), Sequence Number (TSSEQ), Group ID (TSGRPID), Trial Summary Parameter Short Name (TSPARMCD), Trial Summary Parameter (TSPARM), Parameter Value (TSVAL), Parameter Null Flavor (TSVALNF), Parameter Value Code (TSVALCD), Name of the Reference Terminology (TSVCDREF), and Version of the Reference Terminology (TSVCDVER) (Display 1).

TSSEQ is a required variable used to distinguish the multiple records for the same TSPARMCD. For example, the study might have two or more values for trial type, trial primary objective, and trial secondary objective. Thus TSSEQ might present as 1, 2 or 3 depending on the number of records for each given parameter. Take dosing frequency and dosing units as another example; If the study has two or more treatments, each treatment will have its own dose unit and frequency, therefore, we might have TSPARMCD = DOSFRQ, with one TSVAL = "ONCE", TSSEQ ="1" and another TSVAL = "QD", TSSEQ ="2". In such instances, TSGRPID can be used to clearly show which frequency and units are associated with which treatment.

---

[1] CDSIC SDTM Implementation Guide 3.2

While the code lists for Trial Summary Parameter Test Code and Trial Summary Parameter Test are extensible, most values for TSVAL for each parameter must match controlled terminology. There are some parameters however that allow free text from the Protocol, such as;: TITLE, OBJPRIM, OBJSEC, OUTMSPRI, OUTMSSEC, OUTMSEXP, and STOPRULE. There are other parameters that allow numbers, such as PLANSUB, ACTSUB, NARMS, and DOSE.

For parameter values that should be presented with Controlled Terminology, TSVALCD, TSVCDREF, and TSVCDVER must be populated. TSVALCD is to be populated with the appropriate Dictionary code for the given parameter with TSVCDREF and TSVCDVER referencing the name of the Dictionary and Dictionary version respectively. All dates and durations should be ISO 8601 compliant and in such instances, TSVCDREF should be populated with ISO 8601. In the event that TSVAL is null, TSVALNF should be populated based on the ISO 21090 standard, details of which will be discussed in the last section of the paper.

STUDYID	DOMAIN	TSSEQ	TSPARMCD	TSPARM	TSVAL	TSVALNF	TSVALCD	TSVCDREF	TSVCDVER
0000-000	TS	1	ACTSUB	Actual Number of Subjects					
0000-000	TS	1	ADAPT	Adaptive Design	N		C49487	CDISC	2015-03-27
0000-000	TS	1	ADDON	Added on to Existing Treatments	NA		C48660	CDISC	2015-03-27
0000-000	TS	1	AGEMAX	Planned Maximum Age of Subjects	P65Y			ISO 8601	
0000-000	TS	1	AGEMIN	Planned Minimum Age of Subjects	P19Y			ISO 8601	

Display 1. Trial Summary Domain Structure Example

TS PARAMETERS CLASSIFICATION

The Trial Summary dataset is a valuable tool for both internal review and external oversight. It is recommended to develop the TS domain early in the data specification process or even at the beginning of SDTM process. To understand the TS parameters and quickly create this dataset, all the TS parameters have been sub-classified and are listed in Tables 2 - 4 below.

TS PARAMETERS FROM THE PROTOCOLS

The TS Parameters listed in table 2 can be directly retrieved from study protocols and SAPs. It is recommended to add these parameters first on the TS template at the start of processing the datasets. However, given the increased amount of content and scope of the TS domain SAS programmers and SDTM specialists should not be afraid to reach out to colleagues in clinical study teams to review and confirm all TS parameter definitions are appropriate. For example, TTYPE (Trial Type) and INTMODEL (Intervention Model) might not be immediately clear to a SDTM programmer if it is not explicitly detailed in a SAP or protocol. Therefore programmers should not be afraid to consult their statistician or appropriate peers to clarify potential definitions and ensure the accuracy of the SDTM transfer.

TSPARMCD	TSPARM	TSPARMCD	TSPARM
ADDON	Added on to Existing Treatments	OBJSEC	Trial Secondary Objective
AGEMAX	Planned Maximum Age of Subjects	PLANSUB	Planned Number of Subjects
AGEMIN	Planned Minimum Age of Subjects	RANDOM	Trial is Randomized
CURTRT	Current Therapy or Treatment	ROUTE	Route of Administration
DOSE	Dose per Administration	SEXPOP	Sex of Participants
DOSFRQ	Dosing Frequency	STOPRULE	Study Stop Rules
DOSU	Dose Units	TBLIND	Trial Blinding Schema
FCNTRY	Planned Country of Investigational Sites	TCNTRL	Control Type
HLTSUBJI	Healthy Subject Indicator	TINDTP	Trial Indication Type
LENGTH	Trial Length	TITLE	Trial Title
NARMS	Planned Number of Arms	TPHASE	Trial Phase Classification
OBJPRIM	Trial Primary Objective	TTYPE	Trial Type
ADAPT	Adaptive Design	INTMODEL	Intervention Model

Table 2. TS Parameters from the Protocols

TS PARAMETERS FROM REAL DATA

Parameters in table 3 may not be located across the protocol as they require access to real data. The parameters defining the actual number of subjects, study start and end date might only be obtainable after study ends. Therefore, these parameters are most likely created after the completion of the study.

TSPARMCD	TSPARM	TSPARMCD	TSPARM
ACTSUB	Actual Number of Subjects	DCUTDTC	Data Cutoff Date
SENDTC	Study End Date	DCUTDESC	Data Cutoff Description
SSTDTC	Study Start Date		

Table 3. TS Parameters Require Real Data

TS PARAMETERS FROM DICTIONARIES

Table 4 lists parameters which need to be defined via additional supporting dictionaries such as Registry Identifier, Pharmacological Class of Invest. Therapy, and Diagnosis Group.

TSPARMCD	TSPARM	SOURCE LINK
REGID	Registry Identifier	website http://www.clinicaltrials.gov/
PCLAS	Pharmacological Class of Invest. Therapy	Website http://www.fda.gov/ForIndustry/DataStandards/StructuredProduct Labeling
TDIGRP	Diagnosis Group	Conditionally Required. IF healthy volunteers = N then TDIGRP must be present and not null. SNOMED CT http://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html

Table 4. TS Parameters from Additional Supporting Dictionaries

Registry Identifier for New Drug Application (NDA) submissions for all studies can be found on ClinicalTrial.gov website by providing protocol number. Figure 1 shows an example registry identifier: TSVAL="NCT017XX859"; TSVALCD="NCT017XX859"; TSVCDREF="CT.GOV" (Figure 1).

STUDYID	DOMAIN	TSSEQ	TSPARMCD	TSPARM	TSVAL	TSVALNF	TSVALCD	TSVCDREF	TSVCDVER
0000-000	TS	1	REGID	Registry Identifier	NCT017XX859		NCT017XX859	CT.GOV	

Figure 1. Registry Identifier from ClinicalTrial.gov

Identifying the parameter Pharmacological Class of Invest. Therapy can be found through the FDA Pharmacologic Classes file which can be downloaded from the website as shown on Table 5. Figure 2 displays NDF-RT NUI as N0000175430 for Clozapine, and the generated parameters TSVAL="Atypical Antipsychotic"; TSVALCD=" N0000175430"; TSVCDREF="NDF-RT" for TS domain.

Active Moiety Name	Active Moiety UNII	FDA Text Phrase	NDF-RT Concept	NDF-RT NUI
clotrimazole	G07GZ97H65		Azoles [Chemical/Ingredient]	N0000008217
clozapine	J60AR2IKIC	atypical antipsychotic	Atypical Antipsychotic [EPC]	N0000175430
codeine	Q830PW7520	opioid agonist	Opioid Agonist [EPC]	N0000175690

STUDYID	DOMAIN	TSSEQ	TSPARMCD	TSPARM	TSVAL	TSVALNF	TSVALCD	TSVCDREF	TSVCDVER
0000-000	TS	1	PCLAS	Pharmacological Class of Invest. Therapy	Atypical Antipsychotic		N0000175430	NDF-RT	

Figure 2. Pharmacological Class of Invest. Therapy from NDF-RT

OTHER USEFUL RESOURCES AND LINKS

Table 6 below contains useful sources to assist with understanding and implementing newly required TS parameters while Figures 3 - 4 demonstrate how to get this information from the source.

TSPARMCD	TSPARM	SOURCE LINK AND MORE DETAILS
SPONSOR	Clinical Study Sponsor	Website https://www.upik.de/en/upik_suche.cgi.
OUTMSPRI	Primary Outcome Measure	Website http://www.clinicaltrials.gov/
OUTMSSEC	Secondary Outcome Measures	Website http://www.clinicaltrials.gov/
TRT	Investigational Therapy or Treatment	Website http://fdasis.nlm.nih.gov/srs/

Table 6. Additional TS Parameters Source Link

Where previously it was sufficient to populate TSVAL with the name of the Sponsor where TSPARMCD=SPONSOR, SDTM-IG now requires the SDTM programmer to include the Data Universal Numbering System (DUNS) number found on the UPIK directory website. Taking Celerion, Inc as an example, entering the company name, address and the country into the website will return the DUNS number 962170390 which can be used to populate TS as shown in Figure 3.
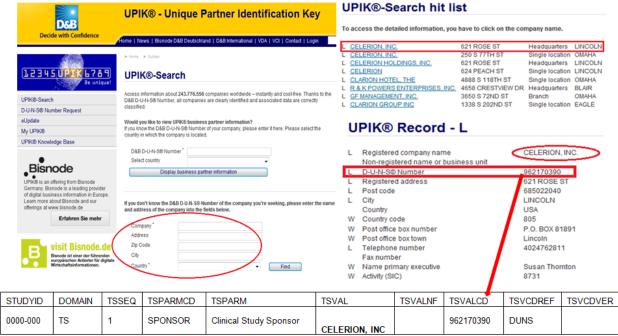.

STUDYID	DOMAIN	TSSEQ	TSPARMCD	TSPARM	TSVAL	TSVALNF	TSVALCD	TSVCDREF	TSVCDVER
0000-000	TS	1	SPONSOR	Clinical Study Sponsor	CELERION, INC		962170390	DUNS	

Figure 3. DUNS Number for SPONSOR from D&B

Information regarding Primary and Secondary Outcome Measures can easily be retrieved along with the Registry Identifier on the clinicaltrials, gov website. If a trial has not yet been registered, this information can generally be found in the study protocol or SAP.

The parameter for Investigational Therapy or Treatment is conditionally required when the Study Type parameter is "INTERVENTIONAL". Programmers must now not only list any drug(s), listed as the actual substance, but also provide the Unique Ingredient Identifier (UNII) registered for this substance as shown in Figure 4 below:
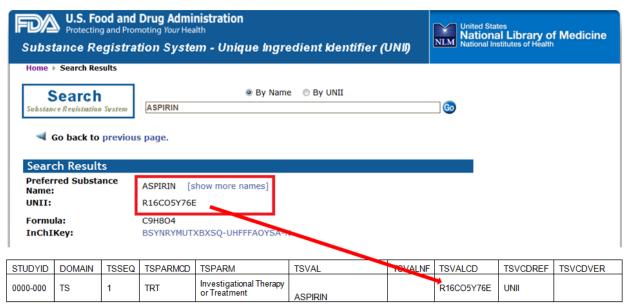
Figure 4. Investigational Therapy or Treatment from U.S. Food and Drug Administration

CONTROL TERMINOLOGY SOURCE LINK

Downloading latest controlled Terminology from the NCI FTP portal seems to play nicely with OpenCDISC.
http://evs.nci.nih.gov/ftp1/CDISC/

Or we can still downloaded directly from the NCI home page here:
http://www.cancer.gov/cancertopics/cancerlibrary/terminologyresources/cdisc

EXCEL METHODOLOGY FOR BUILDING THE SDTM TRIAL SUMMERY DOMAIN

To assist in the build, population, and review of the TS database, and to keep the process simple and efficient, programing teams are likely to find it easier to develop and maintain configurable data files to control and define an ever expanding database. A configurable excel based TS build spreadsheet that automatically links to drop down lists containing controlled terminology and auto-populating corresponding CT code list numbers via look up table links is well worth investigating as shown in Display 2.

TSPARMCD	TSPARM	PROTSPEC	TSVAL	TSVALNF	TSVALCD	TSVCDREF	TSVCDVER	TSVAL1	TSVAL2
ACTSUB	Actual Number of Subjects		20						
ADAPT	Adaptive Design		N		C49487	CDISC	2015-03-27		
ADDON	Added on to Existing Treatments		NA		C48660	CDISC	2015-03-27		
AGEMAX	Planned Maximum Age of Subjects		P65Y	.		ISO 8601			
AGEMIN	Planned Minimum Age of Subjects		P19Y			ISO 8601			
CURTRT	Current Therapy or Treatment		ASPIRIN		R16CO5Y76E	UNII			
DCUTDESC	Data Cutoff Description		DATABASE LOCK						
DCUTDTC	Data Cutoff Date		2015-03-26			ISO 8601			
DOSE	Dose per Administration		200						
DOSFRQ	Dosing Frequency		BID		C64496	CDISC	2015-03-27		
DOSU	Dose Units		mg		C28253	CDISC	2015-03-27		
FCNTRY	Planned Country of Investigational Sites		USA		USA	ISO 3166			
HLTSUBJI	Healthy Subject Indicator		Y		C49488	CDISC	2015-03-27		
INTMODEL	Intervention Model		SINGLE GROUP		C82640	CDISC	2015-03-27		
INTTYPE	Intervention Type		DRUG		C1909	CDISC	2015-03-27		
LENGTH	Trial Length		P8D			ISO 8601			
NARMS	Planned Number of Arms		1						
OBJPRIM	Trial Primary Objective	To compare	To compare the plasma pharmacokinetic profiles (						
OBJSEC	Trial Secondary Objective	To assess t	To assess the safety and tolerability of 200 mg Asp						
OUTMSPRI	Primary Outcome Measure	Pharmacoki	Pharmacokinetics						
PCLAS	Pharmacological Class of Invest. Therapy		Nonsteroidal Anti-inflammatory Dru	.	N0000175722	NDF-RT			
PLANSUB	Planned Number of Subjects		20						
RANDOM	Trial is Randomized		N		C49487	CDISC	2015-03-27		
REGID	Registry Identifier		NCT022866XX	.	NCT022866XX	CT.GOV			
ROUTE	Route of Administration		ORAL		C38288	CDISC	2015-03-27		
SENDTC	Study End Date		2015-02-28			ISO 8601			
SEXPOP	Sex of Participants		BOTH		C49636	CDISC	2015-03-27		
SPONSOR	Clinical Study Sponsor		XXXXX		XXX718596	DUNS			
SSTDTC	Study Start Date		2015-02-01			ISO 8601			
STOPRULE	Study Stop Rules			NA		ISO 21090			
STYPE	Study Type		INTERVENTIONAL		C98388	CDISC	2015-03-27		
TBLIND	Trial Blinding Schema		OPEN LABEL		C49659	CDISC	2015-03-27		
TCNTRL	Control Type		NONE		C41132	CDISC	2015-03-27		
TDIGRP	Diagnosis Group			NA		SNOMED			
TINDTP	Trial Indication Type		TREATMENT		C49656	CDISC	2015-03-27		
TITLE	Trial Title	A 3-Period,	A 3-Period, Fixed Sequence Study to Assess the Effe						
TPHASE	Trial Phase Classification		Phase I Trial		C15600	CDISC	2015-03-27		
TRT	Investigational Therapy or Treatment		ASPIRIN		R16CO5Y76E	UNII			
TTYPE	Trial Type		SAFETY		C49667	CDISC	2015-03-27		

Display 2. Example of a Controlled Drop down List in a Configurable TS Excel File

OPENCDISC UPDATES FOR SDTM IG V3.1.3 & V3.2

The TS domain saw significant changes in SDTM-IG v3.1.3 and v3.2 with increased demand for controlled terminology compared to the previous version SDTM-IG v3.1.2. The 2013 OpenCDISC releases were the first check for several of the newer TS classification parameters which were not expected for SDTM IG v3.1.2 (since SDTM 3.1.2 has no concept of "Required" for trial summary parameters) but which are now required or conditionally required in current SDTM-IG 3.1.3 and 3.2. There are forty-three parameters listed on SDTM-IG 3.1.3 and 3.2 as shown in table 7.

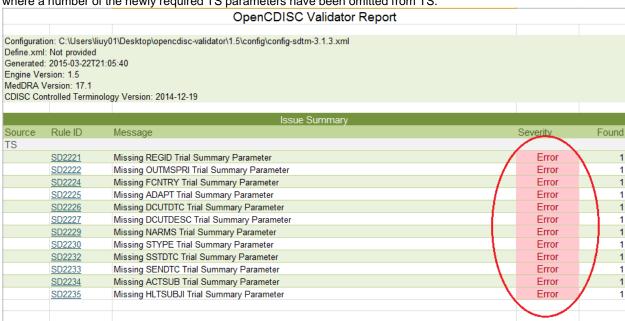| SDTM IG VERSION	TS PARAMETERS
SDTM IG 3.1.3 **&** **SDTM IG 3.2**	**REQUIRED:** ➢ ADDON, AGEMAX, AGEMIN, LENGTH, PLANSUB, RANDOM, SEXPOP, STOPRULE, TBLIND, TCNTRL, TITLE, TPHASE, TTYPE, OBJPRIM, SPONSOR, REGID, OUTMSPRI, FCNTRY, ADAPT, DCUTDTC, DCUTDESC, NARMS, STYPE, SSTDTC, SENDTC, ACTSUB, HLTSUBJI
	CONDITIONALLY REQUIRED: ➢ TDIGRP, TINDTP, CURTRT, TRT, RANDQT, PCLAS, INTMODEL, INTTYPE
	IF APPLICABLE: ➢ OBJSEC, COMPTRT, INDIC, STRATFCT, OUTMSSEC, OUTMSEXP, SDMDUR, CRMDUR

Table 7. SDTM IG 3.1.3[2] TS Parameters

---

[2] SDTM IG 3.1.3 section 7.6 & SDTM IG 3.2 Appendix C1

There are two main updates from V3.1.2 to V3.1.3, one of which is the addition of the parameter value code variables TSVALCD, TSVDREF and TSVCDVER, the other being the inclusion of TSVALNF, all of which have been described in the Trial Summary Domain Structure section above.

OPENCDISC V1.4.1/V1.5 WITH TS DOMAIN

OpenCDISC v1.4.1, superseded by OpenCDISC v1.5, was the first version of the open source validator to verify compliance of these new parameters. All the required parameters must be included into the TS domain otherwise OpenCDISC will return an error message. Display 3 shows an example of the type of OpenCDISC error message where a number of the newly required TS parameters have been omitted from TS.

OpenCDISC Validator Report

Configuration: C:\Users\liuy01\Desktop\opencdisc-validator\1.5\config\config-sdtm-3.1.3.xml
Define.xml: Not provided
Generated: 2015-03-22T21:05:40
Engine Version: 1.5
MedDRA Version: 17.1
CDISC Controlled Terminology Version: 2014-12-19

Issue Summary

Source	Rule ID	Message	Severity	Found
TS				
	SD2221	Missing REGID Trial Summary Parameter	Error	1
	SD2222	Missing OUTMSPRI Trial Summary Parameter	Error	1
	SD2224	Missing FCNTRY Trial Summary Parameter	Error	1
	SD2225	Missing ADAPT Trial Summary Parameter	Error	1
	SD2226	Missing DCUTDTC Trial Summary Parameter	Error	1
	SD2227	Missing DCUTDESC Trial Summary Parameter	Error	1
	SD2229	Missing NARMS Trial Summary Parameter	Error	1
	SD2230	Missing STYPE Trial Summary Parameter	Error	1
	SD2232	Missing SSTDTC Trial Summary Parameter	Error	1
	SD2233	Missing SENDTC Trial Summary Parameter	Error	1
	SD2234	Missing ACTSUB Trial Summary Parameter	Error	1
	SD2235	Missing HLTSUBJI Trial Summary Parameter	Error	1

Display 3. Example on TS Error Message for Some Missing Required Parameters.

OPENCDISC COMMUNITY 2.0 WITH TS DOMAIN

OpenCDISC Community 2.0 with its auto-update functionality for uploading new versions of Controlled Terminology greatly benefits the programmer. OpenCIDSC Community 2.0 added rule SD2241 which verifies that TSVCDREF is populated appropriately for parameters such as CURTRT or TRT where TSVCDREF value must be 'UNII', (Table 8). Previous versions of the validation tool did not go this far.

Rule ID	Message	Description	Category	Severity
SD2241	Invalid TSVCDREF value for CURTRT	TSVCDREF variable value must be 'UNII', when TSPARMCD='CURTRT'	Consistency	Error
SD2251	Invalid TSVALCD value for CURTRT	TSVALCD for CURTRT record must be a valid unique ingredient identifier from FDA Substance Registration System (SRS)	Consistency	Error

Table 8. New Rule for the OpenCDISC Community 2.0

NULL FLAVOR

Null favor is populated only in the event that TSVAL is null. It provides supplementary information about the parameter and the reason as to why TSVAL is null. For example, if TSPARMCD = STOPRULE and protocol does not specify any details regarding Study Stop Rules, the null flavor is "NA" meaning Not Applicable. Other possibilities for Null Flavor value are "NI" (No information), "PINF" (Positive infinity), "UNK" (Unknown), etc. based on different TSVAL variable with no value record. The description on Null flavor logic in the TS domain in the latest SDTM IG is worth a read for anyone involved with set up or reviews of this domain. As SDTMIIG 3.2 states the proposal to include a null flavor variable in the trial summary model needs a better way to illustrate these variables with no protocol-defined value; For example no maximum age of study subjects. Therefore, when the no upper limit described in the protocol, the Null Flavor value for parameter of AGEMAX has to present as PINF (positive infinity). This new procedure still

needs to be tested, reviewed and developed since null flavor might create an unnecessary category for the missing data. Simply providing the true reason as text field as REASND in SDTM might be another better option.

CONCLUSION

While the scope and content of the SDTM TS domain has increased significantly over the past number of years, so too has the potential value and usefulness of this particular dataset. TS can offer a quick, one stop 'go to' location to gain a good understanding of a trail design in the whole. Defining TS requirements as early as possible via protocol driven appendixes or configurable metadata files can greatly simplify the work involved with the TS build at SDTM level. TS data repositories can also serve as valued added data warehouse/reference repository when appropriately classified and management across particular institutions or regulatory authorities.

REFERENCE

Liu, Y. and Read, R. (2014) Adopted Changes for SDTMIG v3.1.3 and 2013 OpenCDISC Upgrades. PharmaSUG, May 2014

CDSIC SDTM Implementation Guide 3.1.2 amendment
Available at http://www.cdisc.org/sdtm

CDSIC SDTM Implementation Guide 3.1.3
Available at http://www.cdisc.org/sdtm

CDSIC SDTM Implementation Guide 3.2
Available at http://www.cdisc.org/sdtm

CDISC Website. http://www.cdisc.org/

OpenCDISC Website. http://www.opencdisc.org/

ACKNOWLEDGE

CONTACT INFORMATION

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

Name: Yi Liu
Enterprise: Celerion Inc.
Address: 621 Rose Street
City, State ZIP: Lincoln, NE 68502
Work Phone: 402-437-4778
E-mail: yi.liu@celerion.com

Name: Jenny Erskine
Enterprise: Celerion Inc.
Address: 22-24 Lisburn Road
City, State ZIP: Belfast, Northern Ireland, BT9 6AD
Work Phone: +44 (0)28 90 554091
E-mail: Jenny.erskine@celerion.com

Name: Stephen Read
Enterprise: Celerion Inc.
Address: 22-24 Lisburn Road
City, State ZIP: Belfast, Northern Ireland, BT9 6AD

Work Phone: +44 (0)28 90 554091
E-mail: stephen.read@celerion.com