Summing up – SDTM Trial Summary Domain

Yi Liu[1], Jenny Erskine[2] and Stephen Read[2]
[1]Celerion, Lincoln, NE USA, [2] Celerion, Belfast, Northern Ireland, UK

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. 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.

TRIAL SUMMARY DOMAIN STRUCTURE

As part of the Trial Design Model, the Trial Summary (TS) 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 be presented 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.

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 should 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 TS 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 the 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 1-3 below.

TS PARAMETERS FROM THE PROTOCOLS

The TS Parameters listed in Table 1 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 collaborate with 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 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 1. TS Parameters from the Protocols

TS PARAMETERS FROM REAL DATA

Parameters in Table 2 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 the 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 2. TS Parameters Require Real Data

TS PARAMETERS FROM DICTIONARIES

Table 3 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	http://www.clinicaltrials.gov/
PCLAS	Pharmacological Class of Invest. Therapy	http://www.fda.gov/ForIndustry/DataStandards/StructuredProductLabeling
TDIGRP	Diagnosis Group	Conditionally Required: If healthy subjects = N then TDIGRP must be present and not null. SNOMED CT(Systematized Nomenclature of Medicine--Clinical Terms) http://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html

Table 3. TS Parameters from Additional Supporting Dictionaries

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

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 3. 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	J60AR2I90C	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

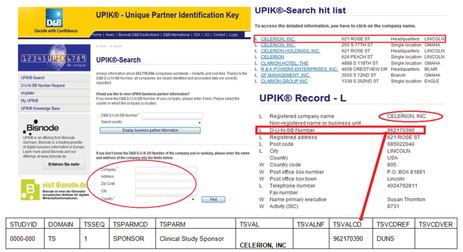
OTHER USEFUL RESOURCES AND LINKS

Table 4 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	http://www.upik.de/en/upik_suche.cgi
OUTMSPRI	Primary Outcome Measure	http://www.clinicaltrials.gov/
OUTMSSEC	Secondary Outcome Measures	http://www.clinicaltrials.gov/
TRT	Investigational Therapy or Treatment	http://fdasis.nlm.nih.gov/srs/

Table 4. Additional TS Parameters Source Link

Where previously it was sufficient to populate TSVAL with the name of the Sponsor (TSPARMCD=SPONSOR), now SDTM-IG requires the SDTM programmers to include the Data Universal Numbering System (DUNS) number found on the UPIK directory website. For example,entering the company name, Celerion Inc., 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.

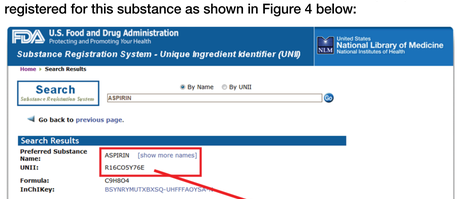


STUDYID	DOMAIN	TSSEQ	TSPARMCD	TSPARM	TSVAL	TSVALNF	TSVALCD	TSVCDREF	TSVCDVER
0000-000	TS	1	SPONSOR	Clinical Study Sponsor	CELERION, INC.		962170390	DUNS	

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 www.clinicaltrials.gov. 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:



STUDYID	DOMAIN	TSSEQ	TSPARMCD	TSPARM	TSVAL	TSVALNF	TSVALCD	TSVCDREF	TSVCDVER
0000-000	TS	1	TRT	Investigational Therapy or Treatment	ASPIRIN		R16CO5Y76E	UNII	

Figure 4. Investigational Therapy or Treatment from U.S. Food and Drug Administration

CONTROL TERMINOLOGY SOURCE LINK

Download the latest controlled Terminology (compatible for OpenCDISC) from the NCI FTP portal http://evs.nci.nih.gov/ftp1/CDISC/ . It is also available on the NCI home page, 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, programming 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		500						
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		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 the plasma pharmacokinetic profiles o							
OBJSEC	Trial Secondary Objective	To assess the safety and tolerability of 200 mg Aq							
OUTMSPRI	Primary Outcome Measure	Pharmacok	Pharmacokinetics						
PCLAS	Pharmacological Class of Invest. Therapy		Nonsteroidal Anti-inflammatory Dr		N0000175722	NDF-RT			
PLANSUB	Planned Number of Subjects		20						
RANDOM	Trial is Randomized		N		C49487	CDISC	2015-03-27		
REGID	Registry Identifier		NCT02286669		NCT02286669	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			
TITLE	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 EFE						
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 are now required or conditionally required in the current SDTM-IG 3.1.3 and 3.2. There are forty-three parameters listed on SDTM-IG 3.1.3 and 3.2.

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 required parameters must be included in the TS domain otherwise OpenCDISC will return an error message. For example, OpenCDISC report displays error message for "Missing REGID Trial Summary Parameter", "Missing OUTMSPRI Trial Summary Parameter", "Missing SSTDTC Trial Summary Parameter", "Missing ACTSUB Trial Summary Parameter", and "Missing HLTSUBJI Trial Summary Parameter" if required parameters of REGID, OUTMSPRI, SSTDTC, ACTSUB and HLTSUBJI are missing.

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. OpenCDISC 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'. Previous versions of the validation tool did not go this far.

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 SDTMIG 3.2 states, the proposal to include a null flavor variable in the TS model needs a better way to illustrate these variables with no protocol-defined value; for example, no maximum age of study subjects. Therefore, when there is no upper limit described in the protocol, the Null Flavor value for parameter of AGEMAX has to be presented 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 a text field, 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 trial 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 the SDTM level. TS data repositories can also serve as valued-added data warehouse/reference repositories when appropriately classified and managed across particular institutions or regulatory authorities.

REFERENCES

1. Liu, Y. and Read, R. (2014) Adopted Changes for SDTMIG v3.1.3 and 2013 OpenCDISC Upgrades. PharmaSUG, May 2014
2. CDSIC SDTM Implementation Guide 3.1.2 amendment. Available at http://www.cdisc.org/sdtm
3. CDSIC SDTM Implementation Guide 3.1.3. Available at http://www.cdisc.org/sdtm
4. CDSIC SDTM Implementation Guide 3.2. Available at http://www.cdisc.org/sdtm
5. CDISC Website. http://www.cdisc.org/
6. OpenCDISC Website. http://www.opencdisc.org/

ACKNOWLEDGE

The authors would like to acknowledge Celerion Inc. for providing the opportunity to work on this paper and some of our colleagues in Celerion's Data Programming group for their valued input.

We would also like to take this opportunity to acknowledge the many contributors to CDISC data standard initiatives including the SDTM SDS group, the OpenCDISC community and the PHUSE SDRG working group.

CONTACT INFORMATION

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

Name: Yi Liu
Enterprise: Celerion Inc.
Address: 621 Rose Street
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
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
Belfast, Northern Ireland, BT9 6AD
Work Phone: +44 (0)28 90 554091
E-mail: stephen.read@celerion.com