Paper 294-2017

A Critique of Implementing the Submission Data Tabulation Model (SDTM) for Drugs and Medical Devices

Carey G. Smoak, DataCeutics, Inc.

ABSTRACT

The Clinical Data Interchange Standards Consortium (CDISC) encompasses a variety of standards for medical research. Amongst the several standards developed by the CDISC organization are standards for data collection (Clinical Data Acquisition Standard Harmonization - CDASH), data submission (Study Data Tabulation Model - SDTM) and data analysis (Analysis Data Model - ADaM). Drug development was the original impetus for developing these standards. Therapeutic Area User Guides (TAUGs) have been a recent focus to provide advice, examples and explanations for collecting and submitting data for a specific disease. Non-subjects even have a way to collect data using the Associated Persons Implementation Guide (SDTMIG-AP). SDTM domains for medical device were published 2012. Interestingly, the use of Device domains in the TAUGs occurs in seventeen out of twenty-four of TAUGs providing examples of the use of the various Device domains. Drug-device studies also provide a contrast on adoption of CDISC standard for drug submissions versus device submissions. Adoption of SDTM (in general and the seven Device domains) by the medical device industry has been slow. Reasons for this slow adoption and suggestions for solutions adoption will be discussed.

INTRODUCTION

This paper will cover a variety of topics starting with a brief background on Clinical Data Interchange and Standards Consortium (CDISC) standards, but will focus mainly on the Study Data Tabulation Model (SDTM), which is one of the CDISC standards. SDTM applies to both pharmaceutical (drug) and medical device products. Therapeutic Area User Guides (TAUGs) assist pharmaceutical / biotech companies with implementing CDISC standards for a specific disease. Interestingly, the TAUGs also use the seven SDTM Device domains. The seven SDTM Device domains were published in 2012, yet adoption of these seven SDTM Device domains by medical device companies have has been slow. The collection of non-subject data (the Associated Persons Implementation Guide – SDTMIG-AP) applies to both pharmaceutical and medical devices. This paper will conclude with a discussion about how the various parts (SDTM for pharmaceutical and medical device products, TAUGs and SDTMIG-AP) of these standards are available for sponsor companies to use in submissions to the FDA. CDISC standards other than SDTM, which apply to FDA submissions, are out of scope for this paper. Examples of other CDISC standards that are important in FDA submissions include (but are not limited to) the Clinical Data Acquisition Standard Harmonization (CDASH), Analysis Data Model (ADaM) and Define.xml.

CDISC BACKGROUND

The CDISC Mission Statement (www.cdisc.org/about) is:

> *The CDISC mission is to develop and support global, platform-independent data standards that enable information system interoperability to improve medical research and related areas of healthcare.*

Thus, CDISC is an organization that develops standards for the medical research industry to facilitate review of the clinical and nonclinical data by regulatory authorities. CDISC has a global perspective and some regulatory authorities now require their standards. The FDA now requires industry sponsor companies to comply with submission of standardized electronic data (CDISC standards mentioned in non-binding documents) for studies starting after 17Dec2016 (Nelson 2016). The Japanese PDMA also requires similar electronic data standards as of 01Oct2016 with a 3.5-year transition period. Thus, adoption of CDISC standards by industry sponsor companies needs to happen quickly.

Other than being required to submit medical research data by following CDISC standards, there are other reasons why industry should adopt CDISC standards (see www.cdisc.org/resources/business-case).

These reasons include:

- Communication among project teams and partners is easier.
- A greater level of accuracy and less training with a constant process.
- Decision-making is simplified.
- Scientists can do the science rather than being concerned with the data.
- Easier transfer of data between partners.
- Opens up a wider choice of tools/technology (as long as they are standards compliant).

Furthermore, research has shown that adopting CDISC standards at the beginning of a research study that companies can save 70-90% of the time and resources in the Study Start-up Stage (time to first patient enrolled). Companies can also save approximately 75% of the time and resources on the non-patient time for Study Conduct and Analysis. Furthermore, for a typical 12-year investment in drug development, savings of up to two years of time results from the adoption of CDISC standards at the earliest phase of clinical research. Thus, about $180 million dollars could be saved per drug submission (see Executive Summary of the Business Case for CDISC Standards, Stage V, 2014 Update at http://www.cdisc.org/system/files/all/article/PDF/2014%20Business%20Case_Executive%20Summary.pdf).

CDISC FOUNDATIONAL STANDARDS

The CDISC foundational standards have been covered elsewhere (Minjoe 2013). Briefly, the foundation standards (located at http://www.cdisc.org/standards/foundational) include:

- PRM – Protocol Representation Model
- SDM-XML – Study/Trial Design in XML
- LAB – Laboratory Data Model
- CDASH – Clinical Data Acquisition Standard Harmonization
- SDTM – Study Data Tabulation Model
- SEND – Standard Exchange of Non-Clinical Data
- ADaM – Analysis Data Model
- DataSet-XMl – an alternative to SAS® version 5 transport files
- ODM-XML – Operational Data Model in XML
- CTR-XML – Clinical Trial Registry in XML
- NCI EVS – NCI Enterprise Vocabulary Services (where all CDISC Controlled Terminology standards are stored)
- Define.xml – is used to describe metadata; is also called Case Report Tabulation Data Definition Specification (CRT-DDS)
- Pharmacogenomics / Genetics
- Questionnaires, Ratings and Scales

STUDY DATA TABULATION MODEL (SDTM)

SDTM is the most well-known of the CDISC standards since it describes the format for submitting data tabulations to a regulatory authority. The SDTM Implementation Guide (SDTMIG) organizes and formats data to help streamline data collection and data analysis. While there are several versions of the SDTM and its corresponding implementation guide (SDTMIG), the FDA will require clinical-data submissions to follow certain SDTMIGs for studies that start after 17Dec2016

(http://www.fda.gov/ForIndustry/DataStandards/StudyDataStandards/default.htm). See the FDA Standards Catalog (see previous link) for a list of when the various versions of SDTMIGs will specifically be required. The latest versions of the SDTM and SDTMIGs are available on the CDISC website (http://www.cdisc.org/sdtm).

Wood (2008) has described the basics of SDTM and SDTMIG. The basic components of the SDTMIG are domains, observations, and observation classes. Domains are group of observations that have a common topic. Usually, but not always, SAS® datasets and domains are equivalent. Observations are a series of named variables that typically correspond to columns in a dataset. An example of a domain is Vital Signs, and within this domain are tests such as heart rate. A test within this domain could be the collection of heart rate at the baseline visit for a particular subject enrolled in the study.

Observation classes fall into three categories: Findings, Events and Interventions:

- Findings are observations that result from planned evaluations during the conduct of a study. Examples of Findings data are vital signs, labs, and ECGs. The Findings domains contains one record (row in a dataset) per finding result or measurement. Thus if temperature, heart rate, systolic blood pressure, and diastolic blood pressure were collected at a particular visit for each subject then there would be four records (one record per vital sign measurement) per subject for that visit.
- Events are occurrences or incidents that happen independent of planned evaluations during the conduct of a clinical trial. For example, the occurrence of an adverse event may not occur at the time of scheduled visit during the clinical trial. In Event domains, there is one record per event.
- Interventions are investigational treatments, therapeutic treatments or procedures that given to or taken by the subjects during the conduct of a clinical trial. For example, if a subject had an adverse event, and administration of a drug alleviated the symptoms then the drug (non-investigational) would be represented as a record in the domain called Concomitant Medications. Interventions domains contain one record per intervention or constant-dosing interval.

In addition to these three general observation classes, there are Special-purpose domains, Findings About, Trial Design and Relationship domains (Table 1. SDTM-Based Domains in SDTMIG 3.2).

Interventions	Events	Findings	Findings About	Special Purpose	Trial Design	Relationship
CM – Concomitant and Prior Medications	AE – Adverse Events	DA – Drug Accountability	FA – Findings About	DM - Demographics	TA – Trial Arms	Supplemental Qualifiers (SUPPQUAL)
EX – Exposure	CE- Clinical Events	DD – Death Details		CO - Comments	TE – Trial Elements	Related Records (RELEC)
EC – Exposure as Collected	DS - Disposition	EG – ECG Test Results		SE – Subject Elements	TV – Trial Visits	
PR - Procedures	DV – Protocol Deviations	IE – Inclusion / Exclusion Criteria Not Met		SV – Subject Visits	TI – Trial Inclusion / Exclusion Criteria	
SU – Substance Use	HO – Healthcare Encounters	IS – Immunogenicity Specimen Assessments			TS – Trial Summary	
	MH – Medical History	LB – Laboratory Test Results				
		MB – Microbiology Specimen				
		MS – Microbiology Susceptibility				

	Test
	MI – Microscopic Findings
	MO - Morphology
	PC – PK Concentrations
	PP – PK Parameters
	PE – Physical Examination
	QS - Questionnaires
	RP – Reproductive System
	SC – Subject Characteristics
	SS – Subject Status
	TU – Tumor Identification
	TR – Tumor Response
	SR – Disease Response
	VS – Vital Signs

Table 1. SDTM-Based Domains in SDTMIG 3.2

THERAPEUTIC AREA USER GUIDES

The purpose of the Therapeutic Area User Guides (TAUGs) is to facilitate solutions using the various CDISC standards for specific diseases or conditions. Typically, the TAUGs provide advice, examples, and explanations regarding the use of CDASH, Controlled Terminology, SDTM, and/or ADaM standards within the context of the specific therapeutic area.

Wood et al (2014) described the process for developing these TAUGs. The process involves finding companies interested in the specific Therapeutic Area (TA), and asking them to provide examples of data elements collected in that TA (e.g., Case Report Forms (CRFs)). These data elements are then mapped to CDASH domains and questions, as well as SDTM-based domains and variables. The TA team may develop new variables and/or domains. These new variables and/or domains then go through the SDTM Governance process and the CDISC Standards Review Council. Once approved, future SDTMIGs will incorporate the new variables and/or domains.

An important concept that has come out of the TAUGs is the Disease Milestone concept (Wood et al 2014; Salyers, Kelly, Wood 2016). The Diabetes TAUG provides a good example of Disease Milestones. Hypoglycemic events in diabetes clinical trials may cause data to be collected across numerous domains. For example, a diabetic subject that has a hypoglycemic event would have data in the following SDTM domains:

- data about the hypoglycemic event as a whole would be in the CE domain,
- the blood glucose level at the time of the hypoglycemic event would be in the LB domain,

- the last dose of study medication prior to the hypoglycemic event would be in the EX domain,
- the last meal prior to the hypoglycemic event would be in the ML domain, and
- any medications taken as a result of the hypoglycemic event would be in the CM domain.

Prior to the Diabetes TAUG Time-Point Variables and Reference Time Points would have handled this data in the SDTM-based datasets. However, the Diabetes TAUG introduced a new variable, MIDS, which when added to the relevant domains to link the data from each hypoglycemic event together. MIDS is somewhat analogous to VISIT in that it is a Timing variable that allows data to be grouped. Rather than a scheduled visit, however, MIDS is a "trigger event" that, when it occurs triggers data collection across multiple domains.

The current TAUGs that are available on the CDISC website (http://www.cdisc.org/standards/therapeutic-areas) are:

- Alzheimer's Disease
- Asthma
- Breast Cancer
- Chronic Obstructive Pulmonary Disease (COPD)
- Cardiovascular
- Diabetes
- Diabetic Kidney Disease
- Dyslipidemia
- Ebola
- Hepatitis C
- Influenza
- Kidney Transplant
- Major Depressive Disorder
- Malaria
- Multiple Sclerosis
- Polycystic Kidney Disease
- Pain
- Parkinson's Disease
- QT Studies
- Rheumatoid Arthritis
- Schizophrenia
- Traumatic Brain Injury
- Tuberculosis
- Virology

MEDICAL DEVICES

MEDICAL DEVICES BACKGROUND

Medical devices are an important and growing part of the medical world, both on their own and in combination with drugs or biologic agents. The ISO 14155 Medical Devices Good Clinical Practices standard defines a "device" as follows:

> *Any instrument, apparatus, appliance, material or other article, whether used alone or in combination, including the software necessary for its proper application, intended by the manufacturer to be used on human beings for the purpose of:*
>
> - *diagnosis, prevention, monitoring, treatment or alleviation of disease,*
> - *diagnosis, monitoring, treatment, alleviation or compensation for an injury or handicap,*
> - *investigation, replacement or modification of the anatomy or of a physiological process,*
> - *control of conception,*
>
> *and which does not achieve its principal intended action in or on the human body by pharmacological, immunological or metabolic means, but which may be assisted in its function by such means.*

While different types of medical devices have widely varying data-submission requirements, most Class II and III devices requiring regulatory data submissions share some fundamental characteristics (see Smoak 2010a; Smoak 2010b for definitions of Classes I, II and III). Regulatory submissions may now use these SDTM-based Device domains for the clinical sections involving devices under study. Therefore, these Device domains were developed to assist device companies in the collection and submission data for Premarket Applications (PMAs), 510(k)s, and Biological License Applications (BLAs).

These Device domains were developed to either to answer the protocol questions, to address associated safety questions, or to associate specific devices to subjects. The investigative sites enter some device data from Case Report Forms (CRFs); however, electronically captured data directly from the device is also an important source of data. To capture data about devices, it was necessary to develop the Device domains that have data that are different from SDTM subject-based domains used in drug studies. They must also accommodate a more complex and variable set of data than those in typical drug development studies. Therefore, this necessitated developing domains based on entities (devices) that is not typically required in most subject-related data (e.g., including the Device-Subject Relationship domain). Figure 1 (Figure 1. Device and Subject Data in Different Domains) illustrates the similarities and differences in device and drug data and demonstrates the need for the development of the Device domains.

For further information on Device domains, please refer to previously published papers on CDISC for medical devices that include the following:

- Differences between medical devices and pharmaceutical products and the goals of the CDISC devices team (Smoak 2007).
- Early domain design of device properties and the unique device identifier (Smoak 2008a; Smoak 2008b).
- The importance of device submissions and approvals (Smoak 2009).
- FDA approval/clearance process for medical devices, the growing importance of medical devices in the healthcare industry and types of medical devices studies needed for approval/clearance (Smoak 2010a; Smoak 2010b).
- Comparison of medical device CRFs with CDASH standards (Shiralkar et al 2010).
- The basics of device regulatory submissions (Smoak 2011).
- A detailed description of the seven medical Device domains (Smoak et al 2012).
- An example of implementing the Device domains for an implantable device (Bullock et al 2013).
- A discussion of the FDA Safety and Innovation Act (FDASIA) of 2012 and its potential impact on SDTM being required for FDA submissions (Smoak et al 2013).
- An example of capturing device data in non-Device domains, e.g., PR and AE (Bullock, Krishnamurthy 2014).

- An example of implementing the SDTM and SDTM Medical Device domains for various types of device therapeutic areas (Yang 2014).
- An example of a Device domain (ADDL) for analysis of device data (Gopal 2015).
- A practical example of the need for implementing ADDL and other analysis features for medical device studies (Yang 2015).

MEDICAL DEVICE SDTM DOMAINS

The seven SDTM domains have been previously described (Smoak 2012). The medical device SDTM Implementation Guide (STDMIG-MD) is available on the CDISC website (www.cdisc.org/standards/foundational/sdtmig). The following sections provide a brief description of the seven Device domains.

device identifiers (DI)

This special-purpose domain contains the data that identifies a specific device unit under study. The primary purpose of this domain is to provide a consistent sponsor-defined variable for a specific device (SPDEVID) for linking data across Device domains, independent of the level of granularity by which a device is identified by a sponsor in a study. The data that uniquely identifies a device is the information that is contained in DI. The domain does not contain information about items that can change without affecting the identification of the device, such as dial settings (e.g., imaging devices). Device Identifiers data exists independently from subjects, and therefore the DI domain does not contain USUBJID.

device properties (DO)

The Device Properties domain is a Findings domain and reports the characteristics of the device that are important to include in the submission, and that do not vary over the course of the study, but do not uniquely identify the device. Examples include expiration date or shelf life. Device Properties data exists independently from subjects and therefore the DO domain does not contain USUBJID.

device-in-use (DU)

Device-In-Use is a Findings domain that contains the values of measurements and settings that are intentionally set on a device when it is used, and may vary from subject to subject or other target. They are characteristics that exist for the device, and have a specific setting for a use instance. This is distinct from Device Properties, which describes the static characteristics of the device. For example: Device Properties would capture that an MRI machine's field strength has a range from 0.2 to 3 Tesla, whereas the Device In-Use domain would capture that the field strength for the MRI scan for Subject 123 was 0.5 T.

device exposure (DX)

Device Exposure is an Interventions domain that records the details of a subject's exposure to a medical device under study. This device is prospectively defined as a test article within a study (via SPDEVD) and may be used by the subject, on the subject, or be implanted into the subject. Examples include but are not limited to stents, drug delivery systems, and any other item under study that is defined as a device according to applicable regulations.

device events (DE)

Device Events is an Events domain that contains information about various kinds of device-related events, such as malfunctions. A device event may or may not be associated with a subject or a visit. The Adverse Event (AE) domain (see SDTMIG v3.2, Section 6.3) records an AE if the device event, such as a malfunction, resulted in an adverse event to a subject. The SDTM RELREC table records the relationship between the AE and the device malfunction. The relationship between AEs and multiple devices is a work in progress and solution will be forthcoming.

device tracking and disposition (DT)

The Device Tracking domain is an Events domain that represents a record of tracking events for a given device. This could include initial shipment, deployment, return, destruction, etc. Different events would be relevant to different types of devices. The last record represents the final disposition of the device. The sponsor decides upon the level of granularity that is appropriate for this domain, based on the type of device and agreements with the regulatory agencies.

device-subject relationships (DR)

The Device-Subject Relationships domain is a special-purpose domain that links each subject to the devices used in the study. Information in this table may have been initially collected and submitted in other domains (e.g., Device Exposure, Device Tracking and Device Events); however, this domain provides a single, consistent location to represent the relationship between a subject and a device, regardless of the device or the domain in which the subject-related data may have been submitted.

Figure 1. Device and Subject Data in Different Domains shows some examples of the relationship between device and subject data for SDTM-based domains. As this figure illustrates, device data can exist independent of subject data. While this may be a novel concept in the drug clinical trials, it is typical of device studies. For example, approval of a heart stent requires collection of data (such as make, model, and lot number) that is not directly connected to a subject. For example, a lot (each lot will have a lot number) of heart stents may be shipped to a particular clinical site for use in the device clinical trial. Putting the lot number in a subject-based domain would be inefficient. However, storing information about the heart stent in domains like DI, DO and DU would be an efficient way to store the data about the device. On the other hand, it would be important to include the serial number in a subject-based domain such as DX. This figure also illustrates places where device and subject data overlap and where subject data (e.g., Demographics) may not have any relationship to the device being used in the clinical trial.

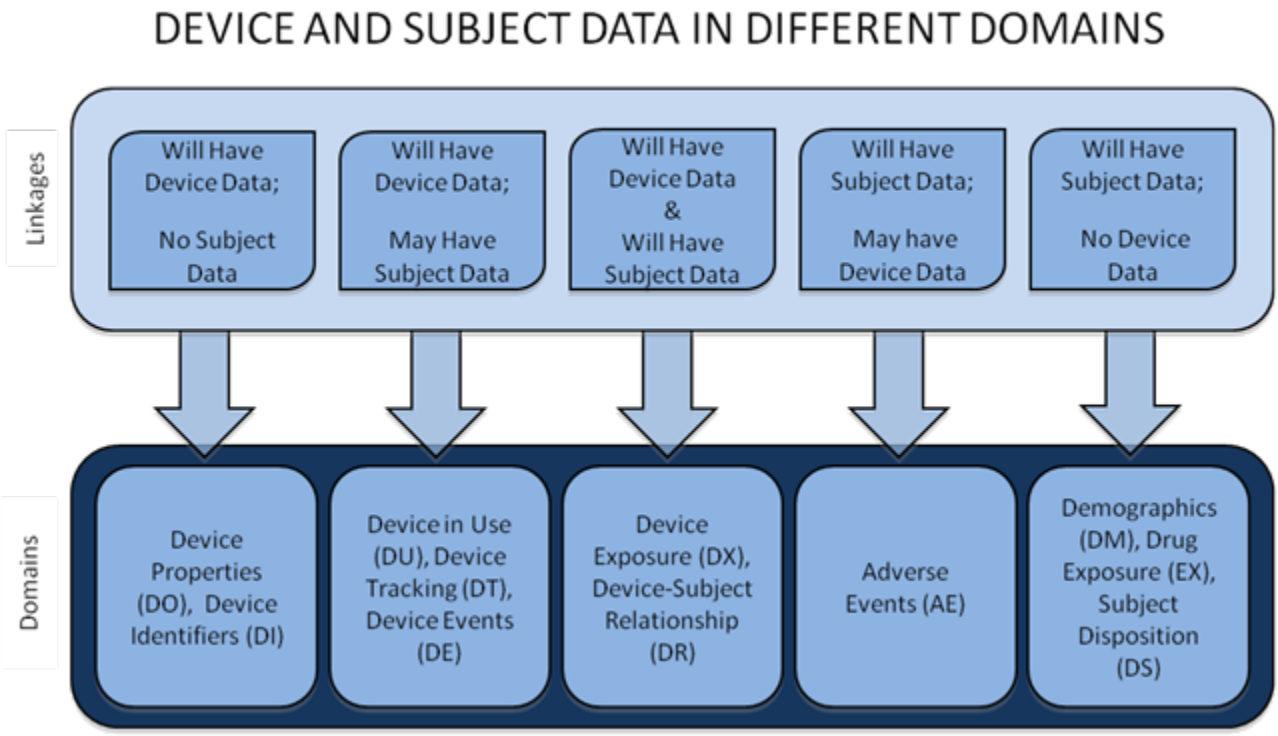

Figure 1. Device and Subject Data in Different Domains

DRUG-DEVICE STUDIES

Examples drug-device studies include combination products (e.g., patches that deliver a drug; drug eluting heart stents, etc.) and companion diagnostic devices which are approved along with a targeted

drug. An example of the latter is a companion diagnostic test that predicts which subjects will benefit from a targeted therapy (Smoak 2016). Drug-device studies are complicated because different branches of the FDA will normally review the drug and device. For example, CDER/CBER may review the drug application while CDRH reviews the device. It also possible that CDRH would want to review data about the drug and that CDER/CBER would want to review data about the device. This may cause a problem with the review since SDTM-based datasets are currently well-established for drugs, but still a work in progress for devices. Based upon this author's experience with companion diagnostics, when CDER looked at companion diagnostic data, it was sent as an analysis file that contained both source and analysis data. Thus, submission of the diagnostic data to CDER did not conform to any type of CDISC standard. While standards for in-vitro diagnostic data have been proposed (Smoak 2014b, Smoak 2014c), adoption of these standards is still a long way off.

THERAPEUTIC AREA USER GUIDES AND MEDICAL DEVICES

Seventeen of the 24 TAUGs published as of this writing use Device domains use one or more of the seven Medical Device domains (Table 2. TAUGs and Device Domains). The TAUG call out all seven of the Device domains and the TAUG examples illustrate the complexity of device data. For example, the DO domain captures the software version of a device for Parkinson's, Polycystic Kidney Disease, and QT TAUGs, while DU captures software version for the Traumatic Brain Injury TAUG. The distinction between DO and DU would be whether it is a static property (DO) of a device or varies based upon the subject or use instance of use (DU).

TAUG (Version)	Device Domains Mentioned in TAUGs	Examples
Alzheimer's Disease (v2)	DI / DO / DU	(1) Device information from lumbar procedure to collect CSF such as spinal needles, tube lots used to store samples, freezer number and microwell plate ID for lab instrument.(2) Imaging devices such as MRIs, PETs and CTs.
Asthma (v1)	DI / DU	(1) Peak flow meter and spirometry for pulmonary function tests. (2) Reference equation for the spirometry device.
Breast Cancer (v1)	DI / DO / DT	Planting a tracer chip implanted for subsequent surgery.
Cardiovascular (v1)	DI	Balloon angioplasty and pacemaker implantation.
COPD (v1)	DI*	Peak flow meter and spirometry for pulmonary function tests.
Diabetes (v1)	DI	Glucose meters and lancet devices used to measure blood glucose levels.
Diabetic Kidney Disease (v1)	None	
Dyslipidemia (v1)	None	
Ebola (v1)	DI	The name of the rapid Ebola diagnostic test kit and the instrument used to test for Ebola IgM antibodies.
Hepatitis C (v1)	None	
Influenza (v1)	DI	Diagnostic test kits.
Kidney Transplant (v1)	DI	Flow cytometry and multiplex assay kits.
Major Depressive Disorder (v1)	None	
Malaria (v1)	DI	Diagnostic test kits, fluorescent spot assay kit and flow cytometry.

Multiple Sclerosis (v1)	DI / DU	Optical Coherence Tomography and Visual Evoked Potential equipment.
Pain (v1.1)	None	
Parkinson's Disease (v1)	DI / DO / DU / DX / DE / DR	Lead hardware from neurosurgery, diagnostic imaging (MRI and PET-SPECT) and software version
Polycystic Kidney Disease (v1)	DI / DO / DU / DR	Imaging devices (MRI, CT and Ultrasound) and software version.
QT Studies (v1)	DI / DO / DR	ECG devices and software version
Rheumatoid Arthritis (v1)	None – will be addressed in v2	
Schizophrenia (v1)	None	
Traumatic Brain Injury (v1)	DI / DO / DU	Imaging devices (CT), protective devices (helmet) and software version.
Tuberculosis (v2)	DI	Diagnostic test kits, mycobacterial detection system, sputum decontamination kits and x-rays.
Virology (v2)	DI	Diagnostic test kits.

*****Not specifically mentioned in TAUG, but should be mentioned since devices are referred to in the TAUG.

Table 2. TAUGs and Device Domains

Khaja (2015) mentions the use of device data in the Diabetes TAUG. For example, the Device Identifier data (DI) domain captures the glucose meter used to measure glucose levels. The key variable in the DI domain is the Sponsor-Defined Device Identifier (SPDEVID) to link the glucose meter to other domains. Another example from Khaja's paper is that of a continuous glucose monitoring (CGM), which used a device inserted under the skin to monitor glucose levels for 72 hours. In this instance, not only would identifying the device (DI) be important, but also the SPDEVID would be an important link to the Device Exposure (DX) domain. Although not mentioned in the paper by Khaja, the Device Events (DE) domain would capture malfunctions of a device.

ASSOCIATED PERSONS IMPLEMENTATION GUIDE (SDTMIG-AP)

The Associated Persons Implementation Guide (SDTMIG-AP) was developed to model the submission of data collected about persons who are not directly enrolled in either drug or device clinical trials. The SDTMIG-AP is available on CDISC website (http://www.cdisc.org/standards/foundational/sdtm) with the SDTM v1.4 and SDTIMIG v3.2 download. The data could be collected on family members or caregivers of a person enrolled in a clinical trial, or the data could be collected on persons who handle investigative devices in a clinical trial. Examples of associated persons data include:

- a subject's family members are associated persons, and data collected about them are associated person's data;
- the original owners of donated organs, blood, tissues, etc. would have associated person's data;
- a questionnaire administered to the caretaker of a study subject is associated person's data;
- the demographics, sexual history, and/or pregnancy history of the sexual partner are associated person's data;
- the collection of Adverse Events on lab operators of an investigational device.

As a principal contributor to the SDTMIG-AP, this last example is the one that I am most familiar with. In certain types of device studies, there is a regulatory requirement to collect adverse events on people who operate the investigational device (e.g., a lab instrument).

Another concept with the SDTMIG-AP is that the associated person can represent an individual person or a group of persons. The above examples encompass the concept of an individual non-subject. However, an example of the associated person representing a group of subjects is pooled lab data. In this case,

pooling of a single lab sample from multiple subjects is performed; which is often the case in blood screening. In these cases, a pool identifier contains the list of associated persons who are in each pool.

Since associated-person data is collected on non-subjects (subjects not enrolled in a clinical trial), it does not belong in the subject-based SDTM domains. Thus non-subject data now has a home in SDTM via the SDTMIG-AP (Wittle, Stackhouse 2016). The data collected on these non-subjects is important to the regulatory submission. Thus, creation of associated person domains handles this non-subject data. The key variable to link the non-subject data to SDTM is the Associated Person Identifier (APID). Other important variables are the Related Subject (RSUBJID), Related Device (RDEVID), and Subject, Device or Study Relationship (SREL).

PUTTING IT ALL TOGETHER

SDTM and SDTMIGs for both drug and device regulatory submissions exist. However, drug regulatory submissions are far ahead of device regulatory submission in the adoption of CDISC standards such as SDTM. The pharmaceutical / biotech industry has been preparing for the day when the FDA will require CDISC standards for regulatory submissions and that day is here. For studies that started after 17Dec2016, the FDA now requires standardized electronic study data for submissions to CDER and CBER (Nelson 2016). As of 01Oct2016, the PDMA (Japan) now accepts CDISC standards such as SDTM. Mandatory submission in this format will be in effect 3.5 years later.

Additionally, the CDISC organization (in conjunction with other organizations) have developed specific TAUGs to assist with the advice, examples and explanations for submitting data in a particular therapeutic area. The use of these TAUGs by sponsor companies should speed-up the setup of clinical trials, data collection and representing the data in a more standardized fashion (Salyers, Kelly, Wood 2016). Moreover, many other CDISC standards exist to assist drug studies with submission of data to the FDA (see section on CDISC Foundational Standards in the Introduction). Thus, the pieces of the puzzle for the implementation of CDISC standards for drug studies is definitely falling into place. In fact, submission of data following CDISC standards has been going on for many years, but more pieces (such as the TAUGs) are helping to bring more pieces of the puzzle together. Thus, CDISCs new motto "Smarter Research to Unlock Cures" (http://www.cdisc.org/debuting-smarter-research-unlock-cures) is being realized.

Submitting device data to the FDA is an interesting situation. While the TAUGs use all of the seven SDTM Device domains for ancillary devices (not an investigational device), it seems that little device data (for investigational devices) is being submitted using SDTM-based domains (including the seven SDTM Device domains). My observation as being a co-founder and co-leader of the CDISC Medical Device Team. In Device Team meetings, my personal impression is CDRH is waiting for the device industry to step forward and submit device data following CDISC standards and the device industry is waiting for CDRH to require them to submit data following CDISC standards. Currently, it appears that CDRH is not moving towards requiring CDISC standards in the next MDUFA (the medical device equivalent of PDUFA) cycle. What needs to be done? The CDISC Medical Device Team needs to continue to educate both CDRH and the device industry about the benefits of standardized data. Standardized data should make reviews easier for CDRH. Adoption of standards by device companies would also benefit them. We know from the pharmaceutical / biotech industry that adoption of standards at the time of study start can reduce time and resources for a study by 70-90% (see Executive Summary of the Business Case for CDISC Standards, Stage V, 2014 Update at http://www.cdisc.org/system/files/all/article/PDF/2014%20Business%20Case_Executive%20Summary.pdf).

So, what specifically needs to be done to educate CDRH and the device industry about CDISC standards? More than a dozen papers on this topic have been presented at SAS® conferences (all of them are referenced in this paper). A workshop on "Management of Data in Medical Device and Diagnostics Studies" at the Society for Clinical Data Management (SCDM) included a presentation on CDISC standards and then need for implementing CDISC standards in medical device studies (http://www.scdm2016.org/program-overview/). At the annual AdvaMed/FDA MTLI Medical Device and Diagnostics Statistical Issues Workshop in 2014, this topic of CDISC standards for medical devices had

two presentations. Smoak (2014a) presented the basics of the seven new SDTM medical Device domains at this conference. An FDA statistician at CDRH also presented the need for CDRH to receive data that conforms to CDISC standards (Nair 2014). Nair (2014) showed specific issues that CDRH has with medical device submission and what CDRH reviewers would like to see in medical device submissions. The following table (Table 3. CDRH Issues and CDISC Solutions) is based on a poster which shows how CDISC standards can help CDRH (Nair et al 2015). Thus, implementation of CDISC standards by CDRH would benefit them in reviewing device submissions.

	CDRH Issue	CDRH Reviewer Request	CDISC Solution
Protocol Deviations	Hard to identify, determine impact	• Summary tables by type of deviation (major / minor) • Protocol deviations by investigational site	• SDTM: designed to facilitate summary table production • CDASH: defines deviation data capture, including narratives; facilitates categorization
Data Traceability	Lack of data traceability means cannot assess data validity	• Provide mechanism to trace each data point from the study report back to the CRF	• ADaM, SDTM, associated define-xml and CDASH-conformant CRFs are specifically designed for this: hyperlink each variable to associated algorithm(s), source dataset(s), controlled terms and annotated CRF(s)

Missing Data	May impact validity of conclusions, choice of statistical model	• Show why and when data are missing (missed visits, value not recorded, etc.)
		• No undisclosed data omissions; justify all data omissions
		• Clearly note all imputed data
			• SDTM and ADaM define-xml:
			  • Origin of each variable is defined as collected, derived or imputed
			  • Algorithms for all derivations and imputations included
			  • Can show what data were included or omitted and why
			• CDASH can indicate what data were missing, with associated dates

Patient Accountability	Hard to determine accountability for all subjects	• Provide patient accountability charts with discussions of missing data	• CDASH and SDTM: Subject Disposition domain captures status of each subject at each defined time point, which can be used to

			produce accountability charts; see also “Missing Data” box
Missing Coding Tools	Hard to identify, determine impact	• Include PROC FORMAT program that creates the format catalog	• Controlled Terminology contains standard "formats" • define-XML contains customized ones and external terms
Trial Data Issues		• Include electronic datasets in PMA submission • Adverse Event listings for medical reviewers • Study endpoints analysis dataset(s) and raw data to minimize complicated manipulations and merges required to validate results • Analysis datasets to support key effectiveness/safety analyses • Include basic demographic variables and important covariates in analysis datasets • Define/README file for datasets and program files • Document datasets and code sufficiently	• SDTM and ADaM provide subject- and device-level tabulation and analysis datasets • Data transmitted in SAS transport files • Standardized AE data support listings from data visualization tools • ADaM defines key effectiveness / safety analyses and datasets, and permits inclusion of any/all relevant variables • ADaM datasets are “one proc away” from running analyses • Define-xml provides structure to document all datasets

Table 3. CDRH Issues and CDISC Solutions

CONCLUSION

For many years’ pharmaceutical companies have been anticipating the requirement of CDISC standards by the FDA. That time has arrived! Studies that started after 17Dec2016 now require (by the FDA) the submission of standardized electronic data and non-binding guidance will recommend the use of CDISC standards as a part meeting this requirement. SDTM has evolved from the base foundational standards into rapidly developing Therapeutic Area User Guides (TAUGs). These TAUGs provide sponsors with advice, examples and explanations for submitting data to the FDA for specific disease areas. Thus, the TAUGs will provide even more information for sponsor companies than can be found in the SDTMIGs to help them with submission of their data to the FDA. Non-subject data will also have a place in regulatory submissions using the SDTMIG-AP. The SDTMIG-AP affects both drug and device clinical trials. Medical Device domains were published in 2012, and interestingly, adoption of the seven Device domains in the TAUGs is common. However, while adoption of these seven Device domains is being done for ancillary devices (the TAUGs), adoption for investigational devices seems to be slow. This problem is

further contrasted by drug-device studies where submission of drug data following CDISC standards is well established, but the use of CDISC standards for devices has been minimal. Therefore, much work remains for the adoption of CDISC standards for investigational devices. Stay tuned! Work is ongoing to help medical device companies and CDRH into the CDISC world.

ACKNOWLEDGMENT

I would like to thank Fred Wood (Senior Manager and Lead Data Standards Consulting, Accenture Accelerated R&D Services) for his review of this paper.

REFERENCES

Bullock TL, Nair S, Krishnamurthy R, Gross TM. "Mapping Unique Aspects of Implantable Medical Device Study Data to CDISC SDTM Medical Device Domains." *Proceedings of the Pharmaceutical Industry SAS® Users Group*, Chicago, IL, May 12-15, 2013. http://www.lexjansen.com/pharmasug/2013/DS/PharmaSUG-2013-DS16.pdf

Bullock TL, Krishnamurthy R. "Referencing Medical Device Data in Standard SDTM domains." *Proceedings of the Pharmaceutical Industry SAS® Users Group*, San Diego, CA, June 1-4, 2014. http://www.lexjansen.com/pharmasug/2014/DS/PharmaSUG-2014-DS06.pdf

Gopal P. "ADDL Model for Device Analysis." *Proceedings of the Pharmaceutical Industry SAS® Users Group*, Orlando, FL, May 17-20, 2015. http://www.lexjansen.com/pharmasug/2015/DS/PharmaSUG-2015-DS04.pdf

Khaja NM. "DIABETES: Submission Data Standards and Therapeutic End Points." *Proceedings of the Pharmaceutical Industry SAS® Users Group*, Orlando, FL, May 17-20, 2015. http://www.lexjansen.com/pharmasug/2015/DS/PharmaSUG-2015-DS12.pdf

Minjoe S. "Introduction to the CDISC Standards." *Proceedings of the Pharmaceutical Industry SAS® Users Group*, Chicago, IL, May 12-15, 2013. http://www.lexjansen.com/pharmasug/2013/IB/PharmaSUG-2013-IB06.pdf

Nair R, Lu L. "Good Practice in PMA Submissions for Efficient Regulatory Decision Making." *Presented at the 7th Annual FDA/MTLI Medical Device and Diagnostics Statistical Issues Workshop*, Washington, DC, April 29-30, 2014. http://www.cdisc.org/system/files/all/Education/Good%20Practice%20in%20PMA%20Submissions%20for%20Efficient%20Regulatory.pdf

Nair R, Howard K, Smoak C, Wood F. "How Can CDISC Help CDRH?" Poster presented at the PhUSE Annual Computational Science Symposium, Silver Springs, MD, March 15-17, 2015. http://www.phusewiki.org/docs/CSS2015Presentations/PP16FINAL.pdf

Nelson B. "FDA Binding Guidance: A Pivotal Milestone for CDISC Standards." Applied Clinical Trials Volume 25, Number 12, 2016. http://www.appliedclinicaltrialsonline.com/fda-binding-guidance-pivotal-milestone-cdisc-standards

Salyers J, Kelly K, Wood F. "Considerations and Conventions within the Therapeutic Area User Guides (TAUGs)." *Proceedings of the Pharmaceutical Industry SAS® Users Group*, Denver, CO, May 8-11, 2016. http://www.lexjansen.com/pharmasug/2016/DS/PharmaSUG-2016-DS07.pdf

Shiralkar P, Tedrow J, Howard K, Fazio L, Luzi D, Facile R. "CDASH Standards for Medical Device Trials: CRF Analysis." *Proceedings of the Pharmaceutical Industry SAS® Users Group,* Orlando, FL, May 23-26, 2010. http://www.lexjansen.com/pharmasug/2010/HS/HS01.pdf

Smoak C. "CDISC for the Medical Device and Diagnostic Industry." *Proceedings of the Pharmaceutical Industry SAS® Users Group,* Denver, CO, June 3-6. 2007. http://www.lexjansen.com/pharmasug/2007/fc/fc10.pdf

Smoak C. "Mission Possible: A Proposed SDTM Domain for the Medical Device and Diagnostic Industry." *Proceedings of the Pharmaceutical Industry SAS® Users Group,* Atlanta, GA, June 1-4, 2008a. http://www.lexjansen.com/pharmasug/2008/rs/rs03.pdf

Smoak CG. "A CDISC SDTM Model for the Medical Device and Diagnostic Industry." *Pharmaceutical Programming 1:28-30,2008b.*

Smoak C. "CDISC for the Medical Device and Diagnostic Industry: An Update." *Proceedings of the Western Users of SAS® Software,* San Jose, CA, September 1-3, 2009. http://www.lexjansen.com/wuss/2009/cdi/CDI-Smoak.pdf

Smoak C. "Medical Device and Diagnostic Industry 101." *Proceedings of the Pharmaceutical Industry SAS® Users Group,* Orlando, FL, May 23-26, 2010a. http://www.lexjansen.com/pharmasug/2010/ib/ib01.pdf

Smoak C. "Medical Device and Diagnostic Industry 101 for SAS® Professionals." *Pharmaceutical Programming 3:12-15,2010b.*

Smoak, C. "Regulatory Submissions for Medical Devices and Diagnostics: the Basics." October, 2011, http://www.cdisc.org/system/files/all/article/application/pdf/cdisc_journal_smoak_p3.pdf

Smoak C, Wood F, Facile R, Howard K. "Seven New SDTM Domains for Medical Devices." *Proceedings of the Pharmaceutical Industry SAS® Users Group,* San Francisco, CA, May 13-16, 2012. http://www.lexjansen.com/pharmasug/2012/DS/PharmaSUG-2012-DS05.pdf

Smoak C, Howard K, Wood F, Facile R. "Data Standards Will Be Required: Challenges for Medical Device Submissions." *Proceedings of the Pharmaceutical Industry SAS® Users Group*, Chicago, IL, May 12-15, 2013. http://www.lexjansen.com/pharmasug/2013/DS/PharmaSUG-2013-DS08.pdf

Smoak C. "Breaking New Ground for Developing Standards for Medical Device Submission." *Presented at the 7th Annual FDA/MTLI Medical Device and Diagnostics Statistical Issues Workshop*, Washington, DC, April 29-30, 2014a.

Smoak C, Krishnamurthy S, Singh M, Truong S. "Forging New SDTM Standards for In-Vitro Diagnostic (IVD) Devices: A Use-Case." *Proceedings of the Pharmaceutical Industry SAS® Users Group*, San Diego, CA, June 1-4, 2014b. http://www.lexjansen.com/pharmasug/2014/DS/PharmaSUG-2014-DS02.pdf

Smoak C, Shamas, S, Chowdagam C, Lim D, Rajeev G. "Route to SDTM Implementation in In-Vitro Diagnostic Industry: Simple or Twisted." *Proceedings of the Pharmaceutical Industry SAS® Users Group*, San Diego, CA, June 1-4, 2014c. http://www.lexjansen.com/pharmasug/2014/PO/PharmaSUG-2014-PO14.pdf

Smoak C. "An Intersection of Pharma and Medical Devices - Development of Companion Diagnostics in Conjunction with Targeted Therapies." *Proceedings of the Pharmaceutical Industry SAS® Users Group*, Denver, CO, May 8-11, 2016. http://www.lexjansen.com/pharmasug/2016/BB/PharmaSUG-2016-BB03.pdf

Wittle A, Stackhouse M. "Associated Persons Domains – Who? What? Where? When? Why? How?" *Proceedings of the Pharmaceutical Industry SAS® Users Group*, Denver, CO, May 8-11, 2016. http://www.lexjansen.com/pharmasug/2016/DS/PharmaSUG-2016-DS10.pdf

Wood F. "The CDISC Study Data Tabulation Model (SDTM): History, Perspective, and Basics." *Proceedings of the Pharmaceutical Industry SAS® Users Group,* Atlanta, GA, June 1-4, 2008. http://www.lexjansen.com/pharmasug/2008/rs/RS10.pdf

Wood F, Wold D, Facile R, Kubick W. "Data Standards Development for Therapeutic Areas: A Focus on SDTM-Based Datasets." *Proceedings of the Pharmaceutical Industry SAS® Users Group*, San Diego, CA, June 1-4, 2014. http://www.lexjansen.com/pharmasug/2014/DS/PharmaSUG-2014-DS05.pdf

Yang J. "SAS® as a Tool to Manage Growing SDTM+ Repository for Medical Device Studies." *Proceedings of the Pharmaceutical Industry SAS® Users Group*, San Diego, CA, June 1-4, 2014. http://www.lexjansen.com/pharmasug/2014/DS/PharmaSUG-2014-DS19.pdf

Yang J. Considerations in Conforming Data from Multiple Implantable Medical Devices to CDISC Standards Using SAS®. Proceedings of the Pharmaceutical Industry SAS® Users Group, Orlando, FL, May 17-20, 2015. http://www.lexjansen.com/pharmasug/2015/DS/PharmaSUG-2015-DS03.pdf

CONTACT INFORMATION

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

Carey G. Smoak
Senior Consultant, Clinical SAS Programmer Analyst
DataCeutics, Inc.
E-mail: smoakc@dataceutics.com