PharmaSUG 2014 – Paper DS02

Forging New SDTM Standards for In-Vitro Diagnostic (IVD) Devices: A Use-Case

Carey Smoak, Roche Molecular Systems, Inc., Pleasanton, CA
Smitha Krishnamurthy, Roche Molecular Systems, Inc., Pleasanton, CA
Mansi Singh, Roche Molecular Systems, Inc., Pleasanton, CA
Sy Truong, Meta-Xceed, Inc., Fremont, CA

ABSTRACT

How does a new data standard get established for medical devices? Data standards for medical devices have made good progress recently with the development of seven new SDTM domains specifically intended for medical device submissions. These seven new domains address the requisite domains to capture the data that is unique to medical devices because medical device data can be distinct and different from pharmaceutical and biotechnology data. These seven medical device domains were intended to capture data that is commonly collected across various types of devices. Currently, in SDTM for drugs, there is an on-going effort to develop therapeutic specific standards (e.g., Alzheimer's, Parkinson's, etc.). Similarly, within medical devices there is a need to develop standards for various types of devices. This paper addresses one such need to design domains specifically for In-Vitro Diagnostic (IVD) devices which are different from other medical devices (e.g., implantable devices). This paper will present a use-case for IVD devices. The project was undertaken at Roche Molecular Systems by a team that identified data used in IVD studies, which can be generalized and implemented as an additional standard for IVD devices. The results are refinements to existing domains and creation of new domains along with variables that follow the standards established by CDISC. The goal of this paper and the team is to have these new standards be used in establishing the next set of SDTM and ADaM data models in support of IVD devices.

INTRODUCTION

In December of 2012, seven new SDTM domains were published for use in medical device submissions (Smoak et al 2012). Since the publication of these seven new SDTM domains, the CDISC Device Team has formed several sub-teams to work on the following projects including:

- CDASH/CRF Standards
- ADaM Standards
- Controlled Terminology Standards
- Granularity Issues
- In-Vitro Diagnostic (IVD) Devices

Additionally, the FDA (mainly for drugs and biologics) is moving towards requiring CDISC standards such as SDTM and ADaM for regulatory submissions. While the requirement for medical devices (including IVD devices) may be less pressing at this time than drugs and biologics, it is still important to continue to work on developing standards for medical devices to prepare for the eventual requirement of CDISC standards such as SDTM and ADaM for medical device regulatory submissions (Smoak et al 2013). Thus, the work of the CDISC Device team is important in preparing for this eventuality.

This paper describes the efforts of one company, Roche Molecular Systems (RMS) to begin developing standards for IVD submissions to the FDA. The authors recognize that their work is only a part of a much larger effort in the following areas:

- **IVD Devices Presentation -** Several years ago, a Biostatistician from an IVD company (not RMS) presented a use-case for CDISC for IVD devices to the FDA at an annual conference in Washington, DC

- **Multi-Divisional Effort within Roche Diagnostics -** The work presented in this paper was done with the support of upper management at RMS and under the auspices of the Clinical Operations Committee of Roche Diagnostics. Several divisions at Roche Diagnostics contributed instrument data to this project. So, while this paper focuses on instrument data from RMS, it has applicability to other divisions in Roche Diagnostics.

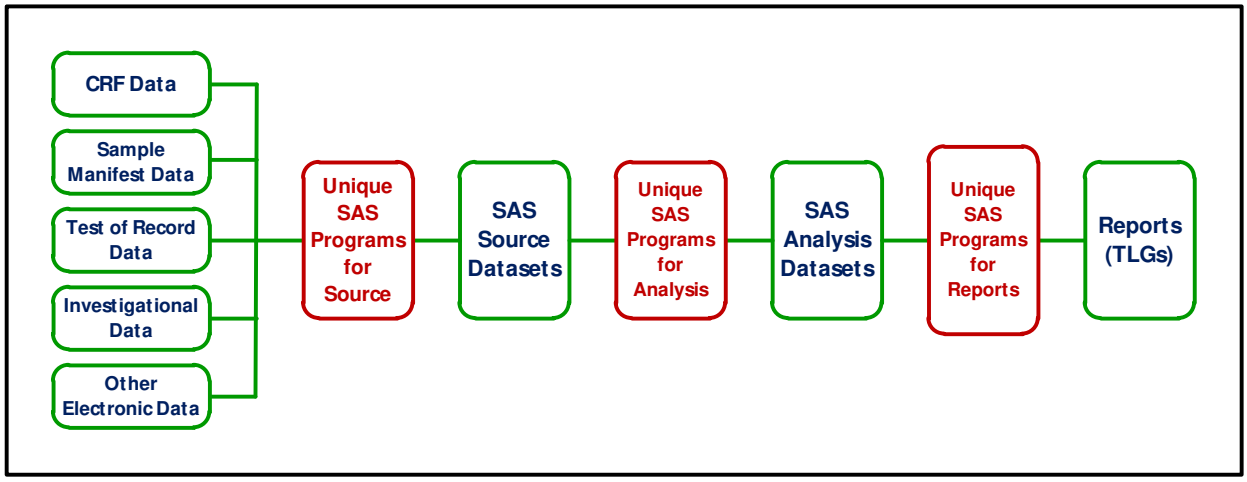
- **Developing Diagnostics Standards -** The CDISC Diagnostics (IVD) Team includes several IVD companies

plus CDISC experts and representatives from the FDA. Data from the IVD companies on this team (plus the effort described in this paper) will be evaluated to help develop a final CDISC SDTM standard for IVDs.

Thus the intent of this paper is not to present a final SDTM standard for IVDs, but rather to show the efforts of one IVD company (RMS) to begin the process of developing a standard which will require further evaluation and refinements.

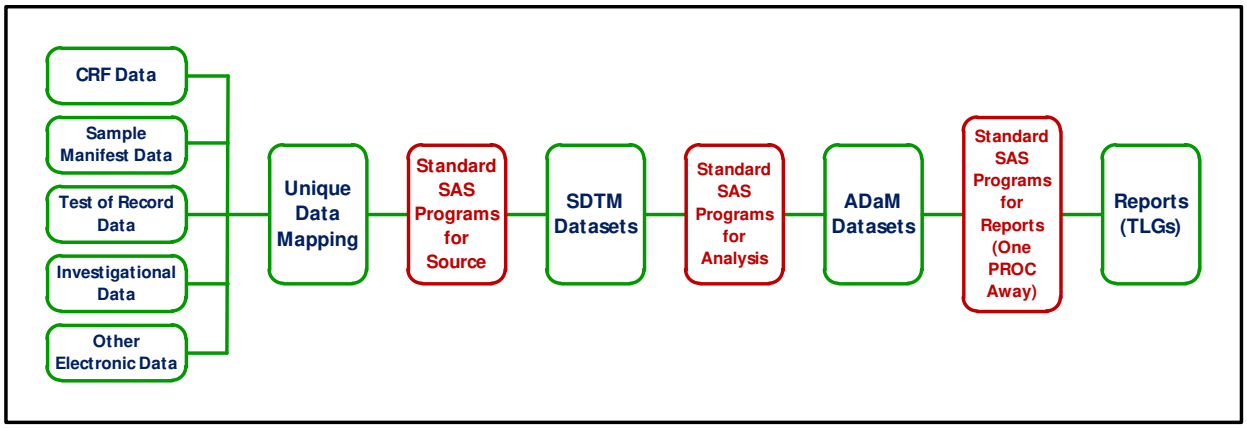
The current SAS® programming environment at RMS (Figure 1) is very labor intensive and unique SAS programs must be developed in order to create source SAS datasets, analysis datasets and TLGs for each study due to lack of standard data structures. These unique SAS programs are single use only for one study (i.e., SAS code is not intended for reuse from study to study). Parts of SAS programs may be used in other studies, but entire SAS programs are rarely reusable from study to study.

Figure 1. Current SAS Programming Environment



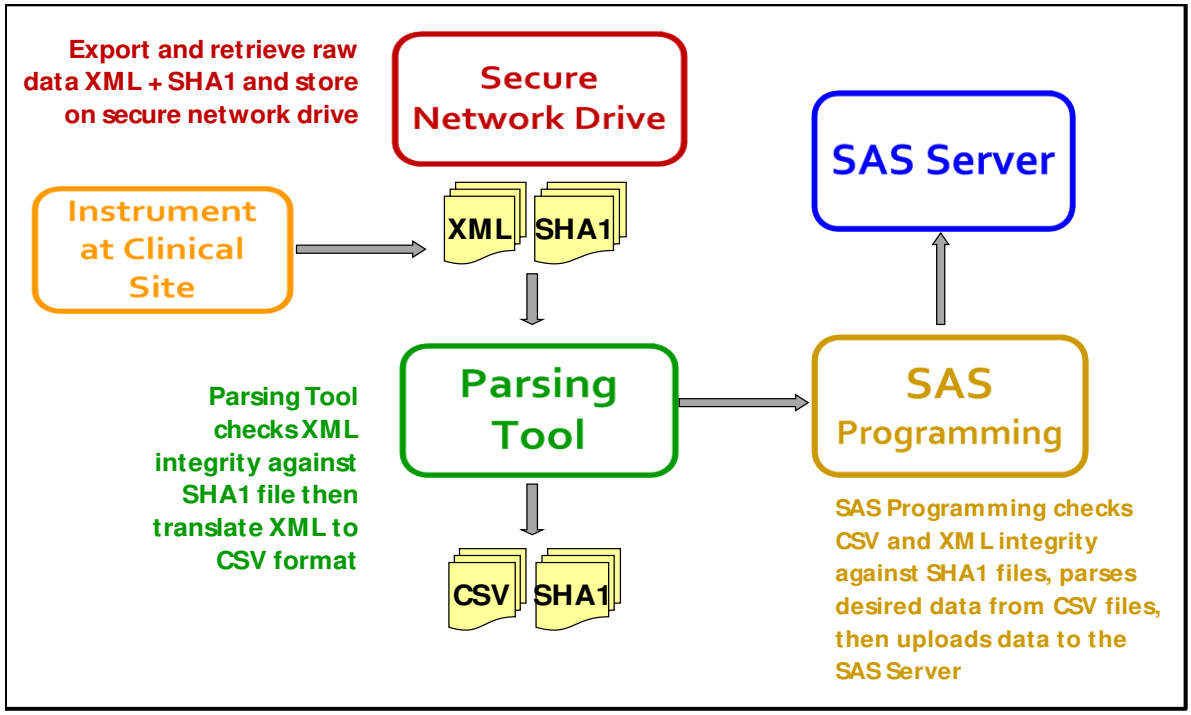
The reason for developing SDTM standards for IVD devices at RMS was to simplify SAS programming and increase the reusability of SAS code across different studies and maintain consistency of SAS code (Figure 2).

Figure 2. Proposed SAS Programming Environment



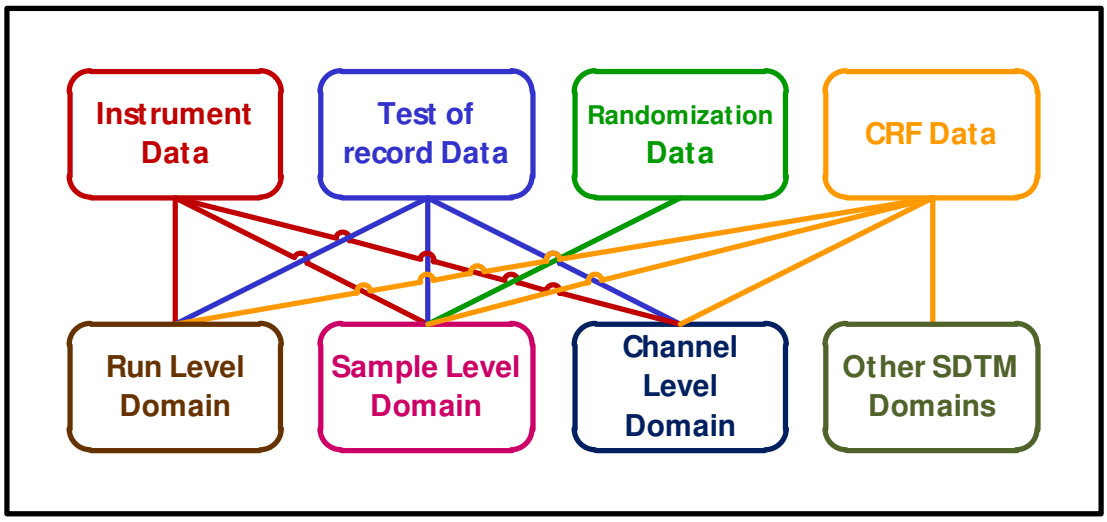
The need for an IVD standard comes from the fact that most of the data that we collect is electronic lab instrument data. Currently, most of the electronic data comes out of the lab instrument as an .xml file. We then have a tool which parses data from the .xml file into a .csv file which is then converted into SAS datasets (Figure 3).

Figure 3. Instrument Data Workflow



The most intensive part of this process of mapping the electronic instrument data involves deciding what data needs to go into different domains. The problem with electronic lab data is that multiple layers of data are included in the instrument output. At the most basic level, the electronic lab data has two levels: run level data (metadata about a run) and sample level data (result data from samples tested by the lab instrument). Some lab instruments can perform multiple lab tests in a single run. In this case, an additional level is referred to as a channel domain. Thus when one puts it all together – instrument lab data, CRF data and other types of data are mapped to the domains shown in Figure 4. Further details on this mapping can be found in Smoak et al 2014.

Figure 4. Mapping of Instrument and CRF Data for RMS IVD Data



The run level, sample level and channel level domains (using instrument data, test of record data, randomization data and CRF data) are explained using a use-case example.

USE-CASE EXAMPLE

The motivation to start this project was to:

- Start developing IVD domains which would work with RMS data
- To streamline our SAS programming processes
- To become consistent with the pharmaceutical industry in terms of use of CDISC standards

The benefit of standardizing our data was to:

- Have submission-ready data for regulatory agencies
- Restructured data format based on CDISC standards
- Foster reusability of SAS code
- Reduce validation time

The first step towards this project was to evaluate the existing standard domains from the SDTMIG v.3.1.3 and the SDTMMDIG (medical devices) v1.0. We identified the domains which could be used to fit our IVD data into existing standards from pharmaceutical SDTM domains and the medical devices domains. Another step was looking into the SDTM+ (a common SDTM approach to add variables prior to the creation of SUPPQUAL for submission) approach for some of the domains. The SDTM+ approach was needed to fit IVD data into some of the existing SDTM domains. Thus this mapping process allowed us to use the variables from existing SDTM domains in order to add additional variables that apply to IVD data (see Figure 5).

SDTM domains that we could use from pharmaceutical industry and did not need any change (not SDTM+) included:

- AE (Adverse Events)
- CM (Concomitant Medications)
- DS (Disposition)
- IE (Inclusion Exclusion)

SDTM domains that we could use from devices industry and did not need any change (not SDTM+) included:

- DU (Device In-Use Properties)
- DI (Device Identifiers)

SDTM+ approach was used to modify a few existing domains to accommodate our IVD data which included:

- DM (Demographics)
- MS (Microbiology Specimens)
- DE (Device Events)
- DV (Protocol Deviations)

Figure 5. Mapping of IVD Data Using SDTM and SDTM+ Approach

RMS Raw Datasets	Standard SDTM Domains										New		
	Pharmaceuticals							Devices					
	AE	CM	DM	DS	DV	IE	MS	DE	DI	DU	RN	CH	SM
AE	X												
DM			X										
DS				X									
DT					X								
DV					X								
EC							X						
IC					X								
ID								X					
IE						X							
IR								X					
IT					X								
MR							X						
RS											X		
SA							X						
TR		X											
Roche Instrument								X	X	X	X	X	X

The idea was to implement automatic direct mapping wherever possible, otherwise we had to derive logic to perform the mapping. In addition to the domains we identified and used above, we had to map the data coming from our lab

instruments. The data from these lab instruments had to be categorized based on the topic.

For RMS, lab instrument data usually consists of:

- **Run** – A run usually consists of samples and controls on a rack which the lab operator puts into the lab instrument for processing. For analysis purposes, each run must be uniquely identified by a sequential or a distinct identifier which, in our case, is called a run number.

- **Sample** – Samples are specimens (e.g., blood) from a subject which are tested by the lab instrument.

- **Channel** – An instrument which tests for multiple analytes requires one channel per analyte for test results. Thus compartmentalized test results for each analyte will come through different instrument channels.

Typically our data can be categorized into three main sources of data consisting of: investigational instrument, test of record and CRF data. For this SDTM project, we have begun to harmonize data from different types of lab instruments used at RMS. The key was to identify how each of our source dataset could fit into an existing Pharma SDTM domain or a Medical Device SDTM domain to determine if a new domain needed to be created. The idea was to follow the general guidelines of SDTM to create these new domains.

Once we identified that we required new domains to be created, the first step was to identify the class or topic that our IVD data would fit into. Based on our assessment, the data was similar to the FINDING class. So our new domains followed the rules of the FINDINGs domain. We created three new domains as:

- **RN (Run Level)** – This domain contains information about the run, such as start date and end date of the run, operator who performed the run, run number, etc. The RN domain contains multiple rows per instrument per run.

- **SM (Sample Level)** – There are many attributes of a sample, but the most important is the sample result. Thus the SM domain contains test results per sample per instrument as a row.

- **CH (Channel Level)** – An instrument which tests multiple assays has multiple channels (one channel per analyte). The CH domain contains metadata about the channel per instrument as a row.

The process of mapping RMS IVD data to SDTM was challenging and remains a work in progress. Further refinements are expected – especially as the CDISC Diagnostic (IVD) Team continues it work with other IVD companies.

MAPPING OF INSTRUMENT DATA: UPDATING TO SDTM+

Detail mappings pertaining to the three domains (RN, SM and CH) are described in another paper entitled “Route to SDTM Implementation in In-Vitro Diagnostic Industry, Simple or Twisted” (Smoak et al 2014).

CONCLUSION

Forging a new path for data standards within a highly regulated environment and within an organization which has entrenched legacy methodologies poses many challenges. An early discussion about data standards, which may potentially take time away from existing resources, is a common challenge. This paper described a bold step that our team did by to forging a new set of domains for IVD devices for our company. This effort was initially applied to SDTM data models (including the new medical device domains) and then it was extended to fit other IVD data that we routinely collect. Rather than taking an existing data standard from CDISC guidelines and apply them, this project had to perform a different and more difficult task. It had to leverage existing data domains and extend them in a use case example that did not fit into existing domains. At the time of this writing, medical device domains (including IVDs) are still not yet fully explored within CDISC. This paper illustrates during this stage that an effective approach is to take real use case examples of data to derive new data domains and related variables. This allowed for CDISC to be applied to IVD devices that did not fit to any existing CDISC domains before. This paper should be taken as a use-case example from one IVD company. The CDISC Diagnostic (IVD) Team (a sub-team of the CDISC Medical Device Team) is working on developing the actual domains which will be proposed for all IVD companies. Thus the use-case in this paper needs to be fully vetted by the CDISC Diagnostic Team before it becomes a standard for all IVD devices.

ACKNOWLEGEMENTS

We would like to gratefully acknowledge the hard work of the SAS Programming Team for Roche Molecular Systems

in Pleasanton, California. The SAS Programming Team also included: Sofia Shamas, Chaitanya Chowdagam, Girish Rajeev, Don Lim and Swarna Umesh. This project would not have been possible without their dedication and countless hours of work. The team who worked on these standards was initially led by Mario Widel who now works for Eli Lily. We are grateful for Mario's leadership in getting this project initiated.

We would also like to gratefully acknowledge the support of the Roche Diagnostics' Clinical Operation Committee. This work was done under their auspices and with their support.

We would also like to gratefully acknowledge the work of the CDISC Diagnostic Team. This team includes industry experts, CDISC experts and FDA representatives from both CDRH and CBER.

REFERENCES

Smoak C, Shamas S, Chowdagam C, Lim D, Rajeev G. 2014. Route to SDTM Implementation in In-Vitro Diagnostic Industry: Simple or Twisted. To appear in the Proceedings of the Annual Conference of the Pharmaceutical Industry SAS User Group. San Diego, California, June 1-4, 2014.

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

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

CONTACT INFORMATION

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

Name:	Carey Smoak	Name:	Mansi Singh
Enterprise:	Roche Molecular Systems, Inc	Enterprise:	Roche Molecular Systems, Inc
Address:	4300 Hacienda Drive	Address:	4300 Hacienda Drive
City, State ZIP:	Pleasanton, CA 95134	City, State ZIP:	Pleasanton, CA 95134
Work Phone:	(925) 730 8033	Work Phone:	(925) 730 8274
Fax:		Fax:	
E-mail:	carey.smoak@roche.com	E-mail:	mansi.singh@roche.com
Name:	Smitha Krishnamurthy	Name:	Sy Truong
Enterprise:	Roche Molecular Systems, Inc	Enterprise:	Meta-Xceed, Inc.
Address:	4300 Hacienda Drive	Address:	42978 Osgood Rd
City, State ZIP:	Pleasanton, CA 95134	City, State ZIP:	Fremont, CA 94539
Work Phone:	(925) 730 8313	Work Phone:	510-979-9333
Fax:		Fax:	510-440-8301
E-mail:	smitha.krishnamurthy .sk1@roche.com	E-mail:	sy.truong@meta-x.com
		Web:	http://meta-x.com

TRADEMARK