PharmaSUG 2016 - Paper DS07

Considerations and Conventions within the Therapeutic Area User Guides (TAUGs)

Jerry Salyers, Accenture Accelerated R&D Services, Berwyn, PA
Kristin Kelly, Accenture Accelerated R&D Services, Berwyn, PA
Fred Wood, Accenture Accelerated R&D Services, Berwyn, PA

ABSTRACT

One of the major initiatives in the pharmaceutical industry involving standards is the development of Therapeutic Area User Guides (TAUGs) under the CFAST (Coalition for Accelerating Standards and Therapies) initiative. CFAST is a collaborative effort among CDISC, with participation from TransCelerate Biopharma, Inc,, FDA, and the National Institute of Health. Currently, there are seventeen TAUGs (three at Version 2, fourteen at Version 1) available for download from the CDISC website. There are also a number currently in development that are moving towards public review and eventual publication. The development of these standards includes the design of case report forms and resulting metadata using the CDASH standard, the mapping of the collected data to SDTM-based datasets, and examples of how the SDTM-based data would be used in the production of ADaM (analysis) datasets. This design and mapping of specialized data to current standards provides many opportunities to implement new and different submission strategies, while remaining compliant with the published standard. With this effort, it is the hope that data across sponsors, from clinical trials within these therapeutic areas, will be more standardized thus allowing for easier and potentially quicker regulatory review.

INTRODUCTION

The effort being made across Pharma in the development of Therapeutic Area (TA) User Guides is one of the more exciting and impactful initiatives currently in progress across the landscape of clinical trials and the submission of standardized data. As standards work takes place across the TA teams, there are often new collection and submission domains and/or variables identified that may be needed to support any given therapeutic area. These new entities are then assessed by the SDS and CDASH teams for inclusion in the next publication of the standard. In the meantime, any new domain or variable in a TAUG would be designated as “Provisional” until such time as they appear in a succeeding publication of the standard. This paper will highlight the current effort being made across the overall project and will focus on the data standards issues, both in collection and submission, and how they have been addressed across a number of therapeutic areas. As was foreseen, the TA initiative has “accelerated” the development of new approaches and novel thinking as we look to create standard CRFs and submission guidelines to support the study and ultimate regulatory review of data from clinical trials across many disease states.

THEREAPEUTIC AREA OVERVIEW

The CFAST initiative is a collaborative effort among the following groups with one of them designated as providing the project management resource in support of each Therapeutic Area User Guide:

- CDISC
- National Cancer Institute – Enterprise Vocabulary Services
- Critical Path Institute
- FDA
- Association of Clinical Research Organizations
- TransCelerate Biopharma, Inc.
- Innovative Medicines Initiative

The following represents the TAUGs that have been published to date and their current version:

- Alzheimer's Disease Version 2
- Asthma Version1
- Cardiovascular Version 1
- Chronic Hepatitis C Version 1
- COPD Version 1
- Diabetes Version 1
- Dyslipidemia Version 1
- Influenza Version 1
- Multiple Sclerosis Version 1
- Pain Version 1
- Parkinson's Disease Version 1
- Polycystic Kidney Disease Version 1
- QT Studies Version 1
- Schizophrenia Version 1
- Traumatic Brain Injury Version 1
- Tuberculosis Version 2
- Virology Version 2

New TAUGs projected to be published in 2016 include:

- Breast Cancer
- Diabetic Kidney Disease
- Rheumatoid Arthritis
- Prostate Cancer

New domains that were first developed in support of a TAUG are shown in the table below:

Domain	TA User Guide
Respiratory System Findings – RE *	Asthma, COPD (not published yet)
Cardiovascular System Findings – CV *	Cardiovascular
Procedure Agents – AG *	Asthma
Urinary Findings – UR *	PKD
Skin Response – SR	TB
Reproductive System Findings – RP	PKD
Subject Status – SS	CV, Parkinson's, PKD
Morphologic Findings - MO	CV, Parkinson's, PKD, TB
Microscopic Findings - MI	Parkinson's, TB
Death Details – DD	CV, Parkinson's, PKD
Musculoskeletal System Findings – MK *	Rheumatoid Arthritis (not published as yet)

* Denotes domains that are "Provisional" until they've appeared in a published version of the standard.

DATA STANDARD EXAMPLES FROM THE TAUGS

Below is presented a series of examples that illustrate the novel approaches and, at times, new data standard conventions as shown in a number of the TAUGs.

Example 1: The concept of "Disease Milestones" as first appearing in the Diabetes TAUG.

In the Diabetes TAUG, we first encounter explanations of the new study level disease milestone domain TM (Trial Milestones) which describes the events and activities that are to be treated as Trial Milestones. TM includes the variable MIDSTYPE that assigns a "type" to the trial milestone, such as "HYPOGLYCEMIC EVENT". At the subject level, the timing and the "instance" of the trial milestone (for those milestones that could occur multiple times) are recorded in the Subject Milestones (SM) dataset. The SM dataset includes the same MIDSTYPE variable as defined at the trial level and also includes the "MIDS" variable that records the name of the disease milestone along with a "number" that provides the "instance".

In the general observation class domains, records that collect assessments that are triggered by the occurrence of a disease milestone are linked back to the SM dataset by new "Timing" variables MIDS, RELMIDS, and MIDSDTC. MIDS again represents the specific Trial Milestone, RELMIDS provides the "temporal" relationship of the assessment to the disease milestone, while MIDSDTC provides the start date/time of the milestone instance. These new Timing variables may also be used to provide context and timing around an Intervention that may have preceded the milestone.

The Diabetes TAUG shows a number of examples of the use of these "milestone" timing variables. Below is an example of a subject's Exposure (EX) records where the CRF collects the overall start and end of the study treatment across the whole of their constant dosing interval and also collects the "most recent" dose prior to a subject experiencing a disease milestone (a Hypoglycemic Event). The SDTM EX dataset would appear as:

ROW	STUDYID	DOMAIN	USUBJID	EXSEQ	EXTRT	EXCAT	EXDOSE	EXDOSU	EXDOSFRQ
1	XYZ001	EX	001-020	1	DRGA		10	mg	BID
2	XYZ001	EX	001-020	2	DRGA	HILIGHTED DOSE	10	mg	BID
3	XYZ001	EX	001-020	3	DRGA	HILIGHTED DOSE	10	mg	BID

ROW	EXSTDTC	EXENDTC	RELMIDS	MIDS	MIDSDTC
1 (Cont)	2013-08-10	2013-11-05			
2 (Cont)	2013-09-01T07:00	2013-09-01T07:00	LAST DOSE PRIOR TO	HYPO 1	2013-09-01T11:00
3 (Cont)	2013-09-24T07:00	2013-09-24T07:00	LAST DOSE PRIOR TO	HYPO 2	2013-09-24T08:48

Example 2: Using FA to capture all "Occurrence" questions (employed in the Dyslipidemia TAUG)

We know that in SDTM, in the AE domain, we don't use the "OCCUR" variable to capture the "Yes or No" response for pre-specified AEs. The only records in AE are for those events that actually occurred to the subject. We use the "--PRESP" variable (set to 'Y') in AE to signify that an event was "pre-specified", however the "Yes or No" response itself to each pre-specified AE is captured in the FA domain with an FATESTCD of "OCCUR".

Considerations and Conventions within the Therapeutic Area User Guides

In the summer of 2015, prompted by work associated with the TAUGs, the decision was made to adopt this "AE Rule" to capture all "Yes or No" questions for pre-specified Events or Interventions. In effect, this took the "--OCCUR" variable out of the general observation class domains. Examples in the Dyslipidemia TAUG look like this:

For the FA domain:

ROW	STUDYID	DOMAIN	USUBJID	FASEQ	FATESTCD	FATEST	FAOBJ	FACAT
1	XYZ001	FA	001-020	1	OCCUR	Occurrence Indicator	SUPRASTATIN	Anti-Dyslipidemic
2	XYZ001	FA	001-020	2	OCCUR	Occurrence Indicator	VARASTATIN	Anti-Dyslipidemic
3	XYZ001	FA	001-020	3	OCCUR	Occurrence Indicator	TOPOSTATIN	Anti-Dyslipidemic

ROW	FAORRES	FASTRESC	FASTAT	FAREASND	FADTC
1 (Cont)	Y	Y			2013-10-27
2 (Cont)	N	N			2013-10-27
3 (Cont)	N	N			2013-10-27

For the CM domain:

ROW	STUDYID	DOMAIN	USUBJID	CMSEQ	CMTRT	CMCAT	CMPRESP
1	XYZ001	CM	001-020	1	SUPRASTATIN	Anti-Dyslipidemic	Y

ROW	CMDOSTOT	CMDOSU`	CMSTDTC	CMENDTC	CMENRF
1 (Cont)	20	mg	2012-01-01		AFTER

So as we see, the FA domain holds the "Occurrence" questions for all 3 pre-specified anti-dyslipidemic medications and the "Yes or No" response is found in FAORRES/FASTRESC. The FAOBJ variable holds the pre-specified medication. The CM domain holds the record for the sole pre-specified medication where the response is "Y". Note the absence of the --OCCUR variable in the CM domain and the value of 'Y' for CMPRESP.

Extending the "AE Rule" to encompass all pre-specified Interventions and Events has been met with mixed support and is being re-visited. In the recently published COPD TAUG, we see that the --OCCUR variable is back in the general observation class domain, MH.

ROW	STUDYID	DOMAIN	USUBJID	MHSEQ	MHTERM	MHEVTYP	MHCAT
1	XYZ001	MH	001-020	1	EMPHYSEMA		COPD HISTORY
2	XYZ001	MH	001-020	2	CHRONIC BRONCHITIS		COPD HISTORY
3	XYZ001	MH	001-020	3	COPD OVERLAP SYNDROME		COPD HISTORY

ROW	MHPRESP	MHOCCUR	MHDTC	MHSTDTC
1 (Cont)	Y	N	2012-09-08	
2 (Cont)	Y	N	2012-09-08	
3 (Cont)	Y	Y	2012-09-08	2011-10-31

Example 3: Findings About versus Clinical Events (example from the Diabetes TAUG)

There is much discussion about the best place to capture "signs and symptoms". Which of these domains would provide the greater flexibility and represent the data in the most correct way possible? Often there isn't a clear-cut answer to this question and it may just be a matter of choice. The following example is in the Diabetes TAUG where the presence or absence of hypoglycemic symptoms is being collected in CE. The Diabetes TAUG does acknowledge that the data could just as easily be represented in FA.

We know that there is a difference between a --DTC captured in a Findings domain versus a --DTC captured in an Events domain. In a Findings domain such as FA, the FADTC is the date of the assessment while in an Events domain, the --DTC is just the date of collection. FA, of course, while being able to show "results" along with result qualifiers on the same record, doesn't have the ability to represent the "start" of an event as does CE, for those symptoms where CEOCCUR = 'Y'.

ROW	STUDYID	DOMAIN	USUBJID	CESEQ	CETERM	CECAT
1	XYZ001	CE	001-020	1	HYPOGLYCEMIC EVENT	HYPOGLYCEMIC EVENTS
2	XYZ001	CE	001-020	2	SWEATING	HYPOGLYCEMIC SYMPTOMS
3	XYZ001	CE	001-020	3	TREMORS/TREMBLING	HYPOGLYCEMIC SYMPTOMS
4	XYZ001	CE	001-200	4	DIZZINESS	HYPOGLYCEMIC SYMPTOMS

ROW	CEPRESP	CEOCCUR	CESTDTC	RELMIDS	MIDS	MIDSDTC
1 (Cont)			2013-09-01T11:00		HYPO 1	
2 (Cont)	Y	Y		DURING	HYPO 1	2013-09-01T11:00
3 (Cont)	Y	N		DURING	HYPO 1	2013-09-01T11:00
4 (Cont)	Y	Y		DURING	HYPO 1	2013-09-01T11:00

Example 4: New Domain Specific Variable (MHEVTYP)

For quite some time, within the SDS Team, there was a subteam working on the issue of "multiple dates" and how best to characterize what the date contained in MHSTDTC referred to. Often, we need the ability to specify the aspect of the medical condition that MHSTDTC refers to. Depending on the data being collected, this date might reflect various aspects of timing, such as date of original diagnosis, date of relapse, date of the start of initial symptoms, or perhaps date of relapse.

Through work on various TAUGs, including Schizophrenia, Hepatitis C, and COPD, the domain specific variable MHEVTYP was defined and created. This variable provides the ability to specify, for the same MHTERM, what the dates in MHSTDTC refer to. Below is an example from the COPD TAUG.

ROW	STUDYID	DOMAIN	USUBJID	MHSEQ	MHTERM	MHEVTYP	MHCAT
1	XYZ001	MH	001-020	1	COPD	SYMPTOMS	COPD HISTORY
2	XYZ001	MH	001-020	2	COPD	DIAGNOSIS	COPD HISTORY

ROW	MHPRESP	MHOCCUR	MHDTC	MHSTDTC
1 (Cont)	Y	Y	2012-09-28	2011-11-10
2 (Cont)	Y	Y	2012-09-28	2012-05-30

Example 5: Body System Specific Morphology/Physiology Domains

This has been another area in which work on the Therapeutic Area User Guides has paved the way for the recent discussion and ultimate decision, made by the CDISC Foundational Teams, to combine the concepts of morphology and physiology by creating "body system specific" domains. These domains would represent both morphology and physiology Findings for a given body system. Already completed work on various TAUGs, where these body system specific "combination" domains have been modeled and drafted was instrumental in the vote to move in this direction. Domains such as OE (Ophthalmology), RP (Reproductive System Findings, and CV (Cardiovascular Physiology) were born out of the TAUGs. The net result of this decision is that the Morpholology (MO) domain will be deprecated in the next published version of the SDTMIG. Also, to make the transition clearer, definitions and assumptions for existing published or drafted body system specific domains will be modified to include both morphology and physiology concepts.

One of the newer domains that combines these two concepts is the Musculoskeletal System Findings domain MK that is being used in the Rheumatoid Arthritis TAUG that is approaching publication. Below is an example from the RA TAUG on how this domain will be used. This example details the "Joint Space Narrowing" assessment of joints in the hand and wrist.

ROW	STUDYID	DOMAIN	USUBJID	MKSEQ	MKTESTCD	MKTEST	MKCAT
1	XYZ001	MK	001-020	1	JSNSCORE	Joint Space Narrowing	SHARP/GENANT JOINT SPACE NARROWING ASSESSMENT
2	XYZ001	MK	001-020	2	JSNSCORE	Joint Space Narrowing	SHARP/GENANT JOINT SPACE NARROWING ASSESSMENT
3	XYZ001	MK	001-020	3	JSCSCORE	Joint Space Narrowing	SHARP/GENANT JOINT SPACE NARROWING ASSESSMENT

ROW	MKSCAT	MKORRES	MKSTRESC	MKSTRESN	MKLOC	MKLAT	MKMETHOD
1 (Cont)	HAND/WRIST JOINTS	MODERATE, 51-75% LOSS OF JOINT SPACE	2	2	INTERPHALANGEAL THUMB JOINT	RIGHT	X-RAY
2 (Cont)	HAND/WRIST JOINTS	MODERATE-SEVERE, 76-95% LOSS OF JOINT SPACE	2.5	2.5	INTERPHALANGEAL THUMB JOINT	LEFT	X-RAY
3 (Cont)	HAND/WRIST JOINTS	MILD, 25-50% LOSS OF JOINT SPACE	1.5	1.5	PROXIMAL INTERPHALANGEAL JOINT 2	RIGHT	X-RAY

ROW	VISITNUM	VISIT	MKDTC
1 (Cont)	4	VISIT 4	2013-08-12
2 (Cont)	4	VISIT 4	2013-08-12
3 (Cont)	4	VISIT 4	2013-08-12

CONCLUSIONS

As stated previously, the quickened pace and the need to provide novel approaches and strategies to data collection and submission have led to a number of modifications as we move closer to publishing the next version of the SDTMIG. We have summarized and highlighted just a few of these that have played an integral part in the development of numerous Therapeutic Area User Guides. Obviously, this work is continuing and will likely continue to break new ground as the overall effort extends into new therapeutic areas. This will continue to be one of the more exciting and fast-moving, as well as rewarding, efforts across the data standards community. The TAUGs will continue to be a mechanism by which sponsors can get targeted clinical trials up and running in a faster and more efficient manner by collecting and representing the trial data in a standard way.

RESOURCES

- Study Data Tabulation Model. Clinical Data Interchange Standards Consortium (CDISC) Submission Data Standards (SDS) Team, Version 1.4, December 2013.
- Study Data Tabulation Model Implementation Guide: Human Clinical Trials. Clinical Data Interchange Standards Consortium (CDISC) Submission Data Standards (SDS) Team, Version 3.2, December 2013
- Various published Therapeutic Area User Guides, available for download from the CDISC website

ACKNOWLEDGMENTS

The authors would like to acknowledge the management at Accenture Life Sciences who understand the importance of data standards in pharmaceutical research, and who have supported our participation in numerous CDISC teams.

CONTACT INFORMATION

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

Jerry Salyers
Senior Consultant, Data Standards Consulting
Accenture Life Sciences
1160 W. Swedesford Road Bldg One
Berwyn, PA 19312
513.668.8126
jerry.j.salyers@accenture.com

Kristin Kelly
Consultant, Data Standards Consulting
Accenture Life Sciences
1160 W. Swedesford Road Bldg One
Berwyn, PA 19312
267.250.5901
richard.m.lewis@accenture.com

Fred Wood
Senior Manager, Data Standards Consulting
Accenture Life Sciences
1160 W. Swedesford Road Bldg One
Berwyn, PA 19312
484.881.2297
f.wood@accenture.com