PharmaSUG China 2019 - Paper DS-069

A Systematic Review of CDISC TAUGs

Angelo Tinazzi, Cytel Inc.

ABSTRACT

The number of available CDISC TAUGs (Therapeutic Area User Guides) are continuously evolving now covering a wide range of indications, such as diabetes, oncology (Prostate, Colon and Breast), Rheumatoid Arthritis, etc. Overall 30 guidance have been released since 2013, with 5 more planned to be finalized in 2019.

In 2015 Johannes Ulander and Niels Both did a review of the content of the existing TAUGs available at that time ("Therapeutic Area standards and their impact on current SDTM implementations", PhUSE, CD03, 2015). The main focus was on SDTM and in particular the analysis of differences of similar aspects covered by the different TAUGs, for example the way 'Primary Diagnosis' information are handled in SDTM.

The objective of this presentation is to introduce TAUGs concept and to give some insights on what's covered in all TAUGs. All foundational standards covered by the different TAUGs are assessed, either CDASH or SDTM or ADaM or the Controlled Terminology.

INTRODUCTION: WHAT IS A THERAPEUTIC AREA USER GUIDE (TAUG)?

A TAUG is a guide for the implementation of CDISC standards in a specific disease area**.** Each TAUGs is based on biomedical concepts identified by subject matter experts and it includes examples from across CDISC foundational standards. This includes the following:

- SDTM mapping of key Therapeutic Area (TA) concepts such as disease background, endpoints
- CDASH with SDTM annotations
- additional examples of situations not covered by the current Implementation Guideline (Ig)
- new Controlled Terminology and New Domains / New Variables might be proposed
- Identification of Regulatory and Medical References

FOUNDATIONAL CONCEPT	THERAPEUTIC AREA CONCEPT
How to model data from labs **(LB domain, SDTMIG)**	How to model data from the complete concept of cerebrospinal fluid biomarker labs **(Alzheimer's/Huntington's TAUGs)**
How to collect labs data **(CDASHIG, CDASH Example CRF Library)**	How to collect data on Nadir CD4+ T-cell counts **(HIV TAUG)**
How to structure time-to-event analysis datasets **(ADTTE dataset, ADaMIG)**	How to structure analysis of time to kidney allograft rejection using ADTTE **(Kidney Transplant TAUG)**

Specifications for how to structure data of certain types Vs. How to implement those specifications in disease-specific use cases

From « Review of Therapeutic Areas for Newcomers », Bess LeRoy, CDISC, CDISC EU Interchange, 2018

TAUGs are developed in collaboration with TA opinion leaders / organizations; for example for the Cardiology TAUG the American College of Cardiology and Duke Clinical Research Institute, and for Multiple Sclerosis TAUG the National Institute for Neurological Disorder and Stroke

Each TAUG when released is considered Provisional[1]; some of the TAUG are «Validated» and accepted by the FDA and therefore listed in the FDA Study Data Technical Conformance Guide (SDTCG).

CONCEPT MAPS

Most of the TAUGs make use of Concept Maps to illustrate Biomedical Concepts. A concept map is a useful graphics way to illustrate relationship among concepts and attributes.

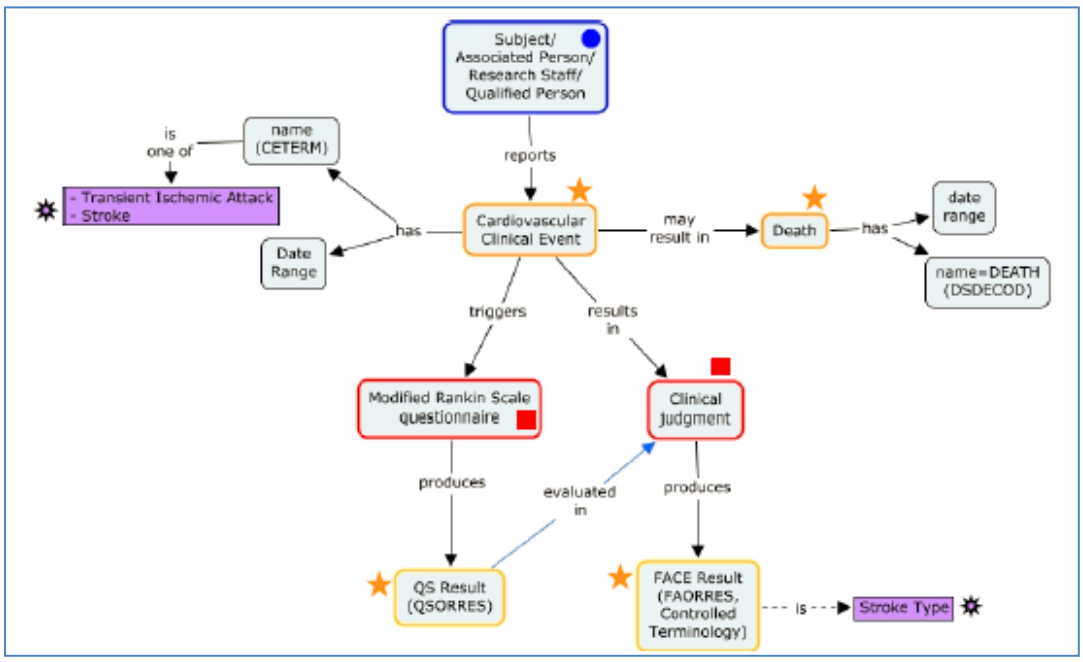

Figure 1. Concepts Maps: Example from the Cardiovascular Transient Ischemic Attack (TIA) TAUG

A TYPICAL TABLE OF CONTENTS OF A TAUG

Despite the specificity of each TAUG, most of the TAUGs address the following topics:

- Clinical Overview / Disease Background
- Trial Design
- Subject and Disease Characteristics i.e. Diagnosis
- Disease Assessments i.e. Symptoms, QRS, Response Measurements
- Routine Data i.e. Concomitant Medications
- Analysis Data (when covered) i.e. key efficacy endpoints with ADaM examples
- Known Issues
- Questionnaires, Rating and Scales (list and approval status)

[1] Provisional standards are published for initial use but they dependent upon completion of other standards and thus may involve risk of upcoming change

- Some provide some sort of excel metadata summarizing what the TAUG is covering i.e. SDTM domains mentioned in the TAUG

Figure 2 is an example of Table of Contents from the Schizophrenia TAUG.

User Guide Section Number	Section Name	Examples	SDTM Domains used	"New" SDTM variables used	Comments
2.0	Clinical Overview				
2.1	Clinical Case Examples				
2.1.1	Clinical Case 1. Patient with Acute Schizophrenia				
2.1.2	Clinical Case 2. Patient with Negative Symptoms in Adjunctive Treatment Trial				
3.0	Trial Design				
3.1	Example Trial 1. Acute Schizophrenia				
3.1.1	Acute Schizophrenia Trial Design Datasets	2.1 Example 1	TA, TE, TV, TI, TS		
3.2	Example Trial 2. Randomized Withdrawal to Assess Maintenance of Response				
3.2.1	Randomized-Withdrawal to Assess Maintenance of Response Trial Design Datasets	2.1 Example 2	TA, TE, TV, TI, TS		
3.3	Example Trial 3. Adjunctive Treatment				
3.3.1	Adjunctive Treatment Trial Design Datasets	2.1 Example 3	TA, TE, TV, TI, TS		
4.0	Subject and Disease Characteristics				
4.1	Schitzophrenia Diagnosis				
4.1.1	Diagnosing Schizophrenia_ DSM-IV-TR versus DSM-5	4.1.1 Example 1	MH	EVDTYP, DSMAXS, DSM4CD, DSM5CD, DIAMTH	
		4.1.1 Example 2	MH, FA	EVDTYP, CRNORD	Includes CRF: Psychiatric History
		4.1.1 Example 3	MH, FA, RELREC	EVDTYP, CRNORD, DSMAXS, DSM4CD, DIAMTH, MHAGE	
4.2	Medical History of Special Interest to Schizophrenia	4.2 Example 1	MH		
4.3	Additional Disease History	4.3 Example 1	MH, HO, FA	EVNUM, REAS,	
					Includes CRF: Psychiatric

Figure 2. TAUG Table of Contents Example (Schizophrenia TAUG)

REFERENCING A TAUG

If your SDTM data package makes use of some TAUG specific recommendations, the TAUG can be referenced in the SDTM TS (Trial Summary) dataset with a specific TSPARMCD/TSPARM (CTAUG/CDISC Therapeutic Area User Guide) available in the CDISC Standard Controlled Terminology; its use is also recommended by the FDA Study Data Technical Conformance Guide (October 2018 on)

THE STATE OF THE ART

The first TAUG was released in 2011 (Alzheimer) and at the end of 2018 overall 30 TAUGs were released in 10 different area of specialty:

- Autoimmune (1) e.g. Rheumatoid Arthritis
- Cardiovascular (2) e.g. QT Studies
- Endocrine (5) e.g. Diabetes
- Infectious (6) e.g. Influenza
- Mental Health (3) e.g. Schizophrenia

- Neurology (4) e.g. Multiple Sclerosis, Parkinson
- Oncology (3) e.g. Breast, Colorectal, Prostate
- Rare Disease (2) e.g. Huntington's Disease
- Respiratory (2) e.g. Asthma
- Treatments (2) e.g. Pain

Available TAUGs are shown in figure 3 and they are available at the following CDISC address: https://www.cdisc.org/standards/therapeutic-areas.

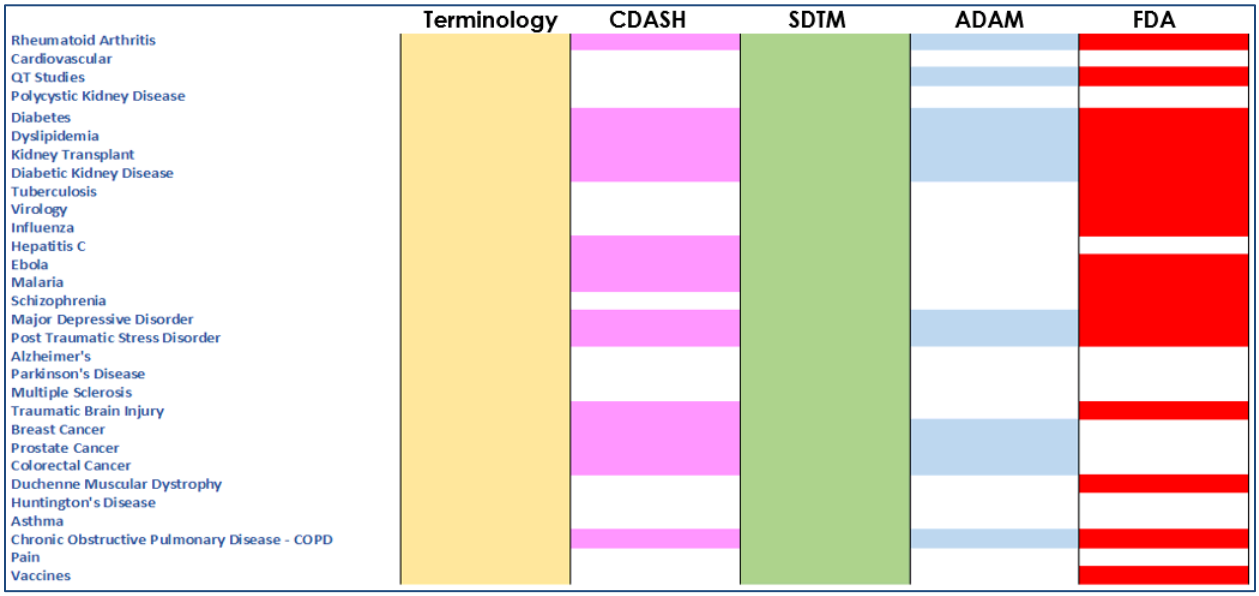

Figure 3. Standards Covered by the available TAUGs – Ordered by Date of First Released Version

A DEEPER INSIGHT

At Cytel, like any other CRO, we have to deal with several different scenario working with different sponsors, different therapeutic area and trial phases. This makes sometime difficult the work of the Statistical Programmer when you either have to take decision on which domain to map a specific CRF form or which best ADaM “modelling” to choose for a particular analysis endpoint.

In order to have a library of examples, an internal project was launched to review all available TAUGs and track all addressed topics, such as SDTM domains discussed by each individual TAUG.

MATERIALS AND METHODS

The project started by getting all available TAUGs from the CDISC website. A number of Cytel CDISC Subject Matter Expert (SME) started to review each TAUG and tracked in a shared excel file the following items:

- which SDTM domains are discussed
- ADaM and CDASH examples
- new proposed CDISC Controlled Terminology
- any non-standard proposed SDTM domain

Furthermore, while tracking each used SDTM domain, we also identified sections in each individual TAUGs further clarifying the SDTM Ig and major difference between TAUG.

RESULTS: SDTM

Figure 4 summarizes the main SDTM domains discussed in the available TAUGs highlighting some specificities covered by some of the TAUG when using each individual SDTM domain. The most discussed topic is the mapping of the disease diagnosis related information (MH). This was one of the key topic discussed in 2015 by Johannes Ulander and Niels Both (see the reference in the reference section).

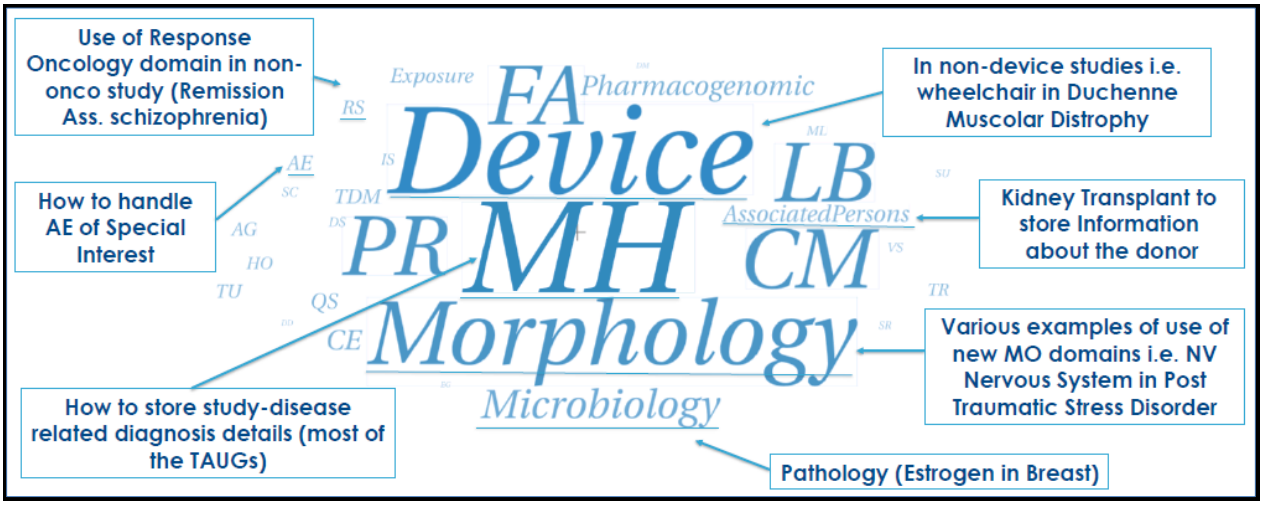

Figure 4. Key SDTM domains used in the available TAUGs

Variations between TAUGs

When Ulander at al at PhUSE 2015 presented their work, one of the main “criticism” to the TAUGs was the fact that for similar topics each individual team working in each individual TAUG, came with different solutions and recommendations. This is the case for example on how to map primary diagnosis in MH (Medical History).

Three years later such a difference has been reduced and as shown in figure 5 the current recommendations from the different TAUGs are as follows:

- MHCAT to group all disease diagnosis related records so that you can distinguish such records from the usual GENERAL MEDICAL HISTORY
- MHTERM to contain the disease diagnosis term, with MHSTDTC being the date when the diagnosis was made
- MHSCAT to identify the type of disease diagnosis information, for example SYMPTOMS vs DIAGNOSIS
- FA to collect details about the disease diagnosis or that make the diagnosis final or to further classify the diagnosis. For example the number of occurrences, age at diagnosis, stage of cancer, etc.

Some of the TAUGs also propose the use of a new SDTM MH variable, MHEVDTYP, to distinguish MH records containing DIAGNOSIS, EPISODE, EXACERBATION and SYMPTOM ONSET records (these are also the allowed terms as per CDISC Controlled Terminology). This new variable has been added to the SDTM standard version 1.7 and its use is discussed in the SDTM Ig 3.3. SDTM Ig 3.3 has also introduced a new dataset SM (Subject Disease Milestone), a domain designed to record the timing, for each subject,

of the disease milestones that have been identified in the Trial Disease Milestone (TM) domain, a domain that has been also introduced in the SDTM Ig 3.3 (figure 6 shows an example).

TAUG	Year	MHTERM	MHCAT	MHSCAT	MHEVDTYP New in IG 3.3	FAMH
Chronic Obstructive Pulmonary Disease [COPD]	2016	CHRONIC OBSTRUCTIVE PULMONARY DISEASE	COPD HISTORY		SYMPTOMS vs DIAGNOSIS	Nr of Occurrences
Diabetic Kidney Disease	2016	• TYPE 1 DIABETE • TYPE 2 DIABETES	DIABETES		DIAGNOSIS	
Ebola	2016	EVD for Ebola Diagnosis	• EVD SYMPTOMS • Empty when diagnosis with MHTERM=EVD		SYMPTOMS ONSET when MHTERM=EVD	
Hepatitis C	2015	HEPATITIS C	• HEPATITIS C • COMORBIDITIES OF INTEREST FOR HEPATITIS C		INFECTION vs DIAGNOSIS	
Malaria	2017		MALARIA SYMPTOMS			
Major Depressive Disorder	2016	MAJOR DEPRESSIVE DISORDER	PSYCHIATRIC HISTORY	MDD SYMPTOM	DIAGNOSIS	
Post Traumatic Stress Disorder	2018	POST TRAUMATIC STRESS DISORDER	PTSD HISTORY	PTSD HISTORY / PTSD SYMPTOMS	DIAGNOSIS vs SYMPTOM ONSET	
Prostate	2017	PROSTATE CANCER			DIAGNOSIS	
Traumatic Brain Injury	2015		TRAUMATIC BRAIN INJURY	• RELATED INJURY vs QUALIFYING EVENT i.e. TBI episode • SIGNS AND SYMPTOMS		Nr of Events

Figure 5. Variations in Primary Diagnosis mapping in MH

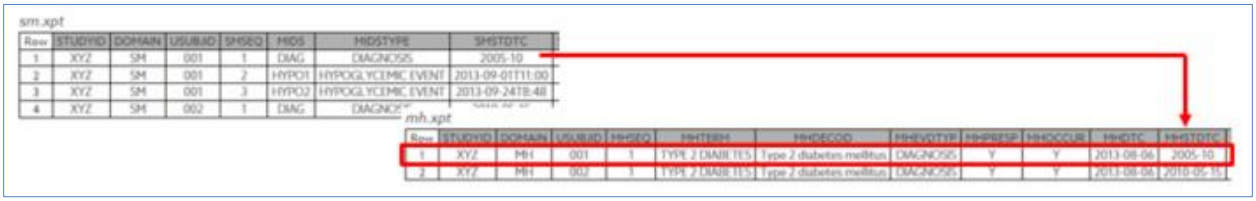

Figure 6. Use of the new SM (Subject Disease Milestone) domain in SDTM Ig 3.3

New Standard SDTM Domains

Some TAUGs have also proposed new domains that now are also part of either the latest CDISC-CT Domain, such as CV (Cardiovascular). The QT (ECG QT Correction Model Data) and the ER (Environmental Risk Factor) domains respectively proposed by the “QT Studies TAUG” and “Ebola, Malaria and Tuberculosis” TAUG, are not yet officially part of the CDISC-CT.

Other SDTM Points of Interest

The following are other interesting points introduced by some of the TAUGs that can be also applied to any other TAs:

- make use of CMGRPID variable to group drugs making the same regimen; for example to group medications making a chemotherapy regimen (“Breast” and “Prostate” TAUGs)
- Use of RS (Response) in other non-oncology TAUGs (for example Schizophrenia, Traumatic Brain Injury)
- Use of TU/TR in other non-oncology TAUGs (for example Tuberculosis and Cardiovascular)

- Use of Devices Ig domains in a non-device study, for example DX for ”Wheelchair, Powered” in Duchenne Muscular Dystrophy or DI for “Protective Device” such as Airbag in Traumatic Brain Injury TAUG
- Laboratory Parameters of Specific Interest, for example HIV Antibody and CD4 for Tuberculosis
- Some TAUGs have also a rich set of aCRF examples which again can be applied or “inspire” SDTM modeling in studies of other TA (for example Malaria, Parkinson’s, Diabetes, Major Depressive Disorder, Rheumatoid Arthritis, Oncology TAUGs)

RESULTS: ADAM

Among the 30 available released TAUGs, 12 TAUGs provide details about analysis topics specific to the TA. These are the most significant examples:

- Breast and Colon Cancer TAUGs: Time to Event and Intermediate ADaM
- Diabetes, Diabetic Kidney Disease, Dyslipidemia
- Chronic Obstructive Pulmonary Disease – COPD: Composite Endpoints
- QT Studies
- Rheumatoid Arthritis: Use of pre-ADSL

Most of them describe key efficacy endpoints with some ADaM mapping examples.

Like for SDTM, also for ADaM the TAUGs have some good example that could be applied to other TA with similar type of endpoints or needs.

Breast TAUG

- Use of Intermediate ADaM datasets prior to Best Overall Response-BOR (ADRESP) and Progression Free Survival (ADTTE) and other related TTE Endpoints
- Identification of Cancer Related ‘Baseline Characteristics’ in ADSL such as Staging
- Key Efficacy Endpoints discussed, for example Progression Free Survival, Disease Free Survival
- Colorectal Cancer reference Breast TAUG for Best Overall Response (BOR) and Time-to-Event (TTE) Endpoints modelling in ADaM

Diabetic Kidney Disease TAUG

- Good Example of Composite Endpoints, an event that is triggered by the occurrence of one of several events, that could be the value of a lab parameter and its ‘persistence’ i.e. confirmation xx weeks after
- Use of AP-- suffix for variables belonging to Associated Persons

Rheumatoid Arthritis TAUG

- The concept of pre-ADSL (some wordings and examples will be also introduced in the draft ADaM Ig 1.2)

OPD-Chronic-Obstructive-Pulmonary-Disease TAUG

This TAUG is rich of examples of ADaM modelling which include not only the identification of some key disease baseline variables in ADSL, but most important examples of intermediate analysis datasets derived from either multiple ADaM and/or SDTM datasets. Figure 7 is an example on how in the TAUG they have proposed to model in ADaM the primary effficacy endpoint through the use of several ADaM datasets each one covering one aspect in the derivation of such a complet endpoint. Splitting such endpoint into several steps and ADaM datasets, add clarity on the way the derivation was made, with interim steps being themselves the source of some analysis.

- Exacerbation

Identification of Individual Events: ADEXAC – Exacerbation Events (ADaM OCCDS, from SDTM.AE, CM, FA, etc.)

↓

Total Number of Events: ADXACSUM – Exacerbation Events Summary (ADaM BDS from ADEXAC)

- Bode Index

ADFEV, ADVS (for BMI), AD6MWT (6-minutes walk test), ADMMRC (Dyspnea Test)

↓

ADBODE: Sum individual score parameters from ADFEV/ADVS/AD6MWT/ADMMRC

Figure 7. Example of complex efficacy endpoint derivation

Diabetes TAUG – AdaM Supplement

The Diabetes TAUG is one of the most complete TAUG. It has also a specific document (supplement) for ADaM (about 38 pages). This supplement provide also some examples of Analysis Results Metadata (ARM) and it has introduced a standard way of representing randomization stratification factors in ADSL (see figure 8) ; this idea has been also adopted by the ADaM team and it will be proposed in the ADaM Ig 1.2.

Variable Name	Variable Label	Type	Length/Display Format	Codelist/Controlled Terms	Source/Derivation/Comment
USUBJID	Unique Subject Identifier	text	$15		DM.USUBJID
STRATA	Randomized Strata	text	$30	>7-<9% | Metformin alone; >=9% | Metformin alone; >7-<9% | Metformin + insulin; >=9% | Metformin + insulin	Obtained from QVAL in SUPPDM where QNAM = "STRATA" *Note: At present there is not a standard approach for capturing stratification factors in SDTM-based datasets. This variable represents the combination of individual stratum values used for randomization. The above text is an example and uses a pipe (|) as a delimiter between individual stratum values. These data could come from other sources as well depending on methodologies used for the design and the management of the randomization schedule.*
STRATAN	Randomized Strata (N)	integer	1	1; 2; 3; 4	= 1 when ADSL.STRATA = ">7-<9% | Metformin alone"; = 2 when ADSL.STRATA = ">=9% | Metformin alone"; = 3 when ADSL.STRATA = ">7-<9% | Metformin + insulin"; = 4 when ADSL.STRATA = ">=9% | Metformin + insulin"
STRAT1NM	Description of Stratum 1	text	$20	HbA1c at Baseline	Assigned based on stratification factors defined a given trial. The value is the same across all subjects and is intended to provide a full text description of the first stratification factor.
STRAT1	Randomized Value of Stratum 1	text	$6	>7-<9%; >=9%	Derived from ADSL.STRATA and is the text string up to the first delimiter of " ".
STRAT1N	Randomized Value of Stratum 1 (N)	integer	1	0; 1	= 0 when ADSL.STRAT1 = ">7-<9%"; = 1 when ADSL.STRAT1 = ">=9%"

Figure 8. ADSL standard variables to represent randomization stratification factors

CONCLUSION

The TAUGs are a great addition of CDISC to their set of standards and they have the main purpose to reduce variability and space of interpretation when applying the CDISC standards in different Therapeutic Area by different sponsors, thus reducing the variability across studies of the same type.

From our analysis there are still some aspects that should be solved or improved:

- There are still some variations between TAUGs, but for sure less than what it was in 2015
- We recommend CDISC to revise older TAUGs, for example those older than 2 years, and align with Ig 3.3 enhancements, for example the use of MHEVDTYP in MH

At Cytel we will complete the mapping of the topics covered by each individual TAUGs with the aim of having more examples the different team can make use of:

- CDASH and SDTM examples
- Good mapping examples by groups of domains e.g. morphology domains
- Additional ADaM Implementation Examples

REFERENCES

Leroy, Bess. 2018. “Review of Therapeutic Areas for Newcomers” CDISC EU Interchange, Berlin

Ulander Johannes and Both Niels.2015. “Therapeutic Area standards and their impact on current SDTM implementations” PhUSE-EU, Vienna

Langendorf, Kirsten Walther. 2018. “Easing Your Pain with Biomedical Concepts” PhUSE-EU, Berlin

CONTACT INFORMATION

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

Angelo Tinazzi
Cytel Inc
angelo.tinazzi@cytel.com
www.cytel.com
https://www.cytel.com/blog/topic/statistical-programming check for my blog series “The Good Data Submission Doctor”