Data Standards, Considerations, and Conventions within the Therapeutic Area User Guides

Jerry Salyers and Kristin Kelly

Data Standards Consulting
Accelerated R&D Services
Accenture

High performance. Delivered.

Philadelphia Univ. PharmaSUG SDE
October 22 2015

accenture

consulting | technology | outsourcing

Agenda

- CFAST TAUG Overview
- New Domains in TAUGs
- Domain Specific Variables - MHEVTYP
- CE vs. FA
- Pre-specified Events (--OCCUR vs. FA)
- Morphology vs. Physiology
- Adjudicated Data Representation
- Non-Standard Variables Representation
- Metadata Development Forum
- Conclusions

Therapeutic Area Collaboration

- CDISC
- National Cancer Institute - EVS
- Critical Path Institute
- FDA
- Association of Clinical Research Organizations
- TransCelerate Biopharma Inc.
- Innovative Medicines Initiative

TransCelerate Member Companies

Founding Members

- Abbvie
- AstraZeneca
- Boehringer-Ingelheim
- Bristol-Meyers Squibb
- Lilly
- GlaxoSmithKline
- Johnson & Johnson
- Pfizer
- Roche
- Sanofi

Additional Members

- Allergan
- Astellas
- Biogen Idec
- Cubist Pharmaceuticals
- EMD Serono, Inc
- Forest Laboratories
- Medgenics
- Shionogi
- UCB

Accenture providing project management and SMEs

Clinical Data Standards
Workstream Mid-Level Milestones & Activities

CFAST Program Overview – September 2015 CDISC

Therapeutic Area	Organization/PM	Stage 0 Charter Approval	Stage 1 Check of Concepts Completed	Stage 3a Posted for Internal Review	Stage 3b Posted for Public Review	*Stage 3c **Projected Publication
Traumatic Brain Injury v1	CDISC Amy Palmer	Oct 13	Sep 14	Mar	Jul	Q415
Breast Cancer v1	CDISC John Owen	Oct 14	Oct 14	Mar	Oct	Q415
COPD v1	CDISC/Quintiles Sherwood Barbee	Sep 14	Dec 14	Jul	Oct	Q415
ADaM Supplement to Diabetes v1	TCB Rachael Zirkle	NA	NA	Mar	Jul	Q415
Virology v2	C-Path Laura Butte	Mar	Apr	May	Jul	Q315
Diabetic Kidney Disease v1	TCB Rachael Zirkle	May	Aug	Sep	Dec	Q116
Tuberculosis v2	C-Path Laura Butte	Apr	Apr	Sep	Oct	Q116
Rheumatoid Arthritis v1	TCB Trisha Simpson	Jun	Sep	Oct		Q216
CV Imaging v1	CDISC/DCRI Amy Palmer	May	Jul	Oct	Dec	Q216
Prostate Cancer v1	CDISC John Owen	Oct				Q316
Major Depressive Disorder	CDISC Amy Palmer	Oct				Q316
General Anxiety Disorder v1	CDISC Amy Palmer					
Bi-polar Disorder v1	CDISC Amy Palmer					

Stage 0 – Scoping , Stage 1 – Concept Modeling, Stage 2 – Standards Development, Stage 3a – Internal Review, Stage 3b – Public Review, Stage 3c – Publication

September 8, 2015 Key: Stage completed | Stage ongoing | All months reflect when stage is, or is projected to be, completed.

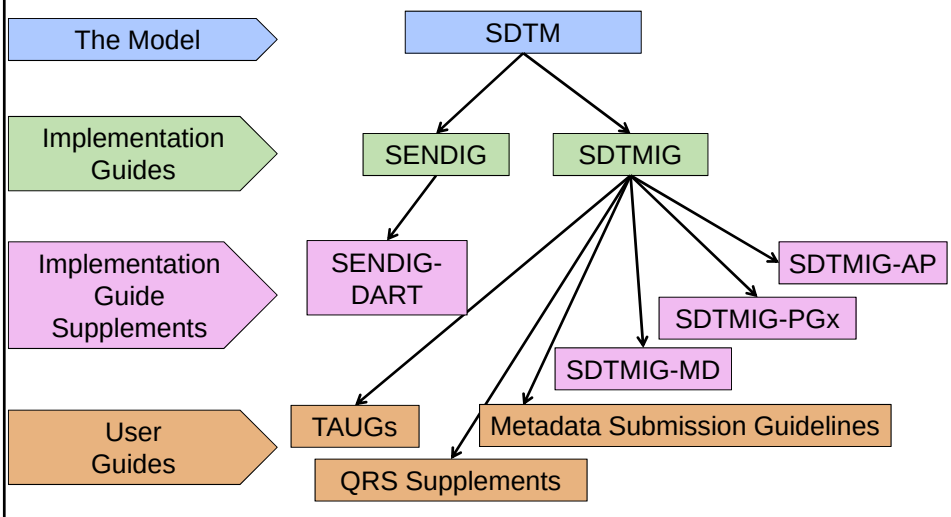
Completed Therapeutic Area User Guides (TAUGs)

- Alzheimer's Disease v2
- Asthma v1
- Cardiovascular v1
- Chronic Hepatitis C v1
- Diabetes v1
- Dyslipidemia v1
- Influenza v1
- Multiple Sclerosis v1
- Pain v1
- Parkinson's Disease v1
- Polycystic Kidney Disease v1
- QT Studies v1
- Schizophrenia v1
- Tuberculosis v1
- Virology v1

New TAUGs in Next 12 Months

- Breast Cancer v1
- COPD v1
- Diabetic Kidney Disease v1
- Prostate Cancer v1
- Rheumatoid Arthritis v1
- Traumatic Brain Injury v1
- Tuberculosis v2
- Virology v2

SDTM, Implementation Guide, and User Guide Relationships



A Note on Provisional Standards and Domains

In general, when anything CDISC publishes (after going through all the necessary reviews) is dependent upon something else being published, it's **provisional**. For example:

- Implementation guides are published as provisional when some of the tables have not yet appeared in an SDTM version.
  - Such was the case for the first versions of the SDTMIG-AP and SDTMIG-MD.
- User Guides are provisional when domains are not in an IG or variables are not in the SDTM.
  - The first versions of the TAUGs for Asthma and Diabetes are examples.
- In general, any domains where CT has not yet been fully developed are also provisional.

New Domains in TAUGs

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)

Subject Disease Milestones (SM)

- New table (2.2.10) identified in SDTM v1.5
- Designed to record the timing of Disease Milestones identified in the Trial Disease Milestones (TM) dataset (also new in SDTM v1.5)
- Introduces new variables MIDS and MIDSTYPE
  - MIDS records the specific name of the Disease Milestone and the "instance" (for milestones that can occur multiple times)
  - MIDSTYPE records the type of Disease Milestone such as 'HYPOGLYCEMIC EVENT'
- Used in the Diabetes TAUG
- Also new timing variables in Table 2.2.5 of the model
  - MIDS: as above, value being unique within a subject
  - RELMIDS: Describes the "temporal" relationship of the observation as shown in MIDS, e.g., 'IMMEDIATELY BEFORE'
  - MIDSDTC: The start date/time of the observation as detailed in MIDS

Subject Disease Milestones (2)

- Example from the Diabetes TAUG which shows a subject's last dose prior to a hypoglycemic event (as identified in the MIDS variable).

Row	STUDYID	DOMAIN	USUBJID	EXSEQ	EXTRT	EXCAT	EXDOSE	EXDOSU	EXDOSFRQ
1	XYZ	EX	XYZ-001-001	1	DRUG A		10	mg	BID
2	XYZ	EX	XYZ-001-001	2	DRUG A	HIGHLIGHTED DOSE	10	mg	
3	XYZ	EX	XYZ-001-001	3	DRUG A	HIGHLIGHTED DOSE	10	mg	

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

Using FA for Occurrence Questions

- The SDS Team developed “The AE Rule” for pre-specified adverse events
  - The occurrence question (Yes or No response) is submitted in Findings About
  - If the response is “Y’, CRF instructs site to enter an AE on the summary Adverse Event page.
    - AEPRESP = Y
    - The --OCCUR variable is not populated in AE
- Last summer, a decision was made to adopt this rule for other pre-specified Events and Interventions modeled in the TAUGs.
  - Implementation has not been entirely consistent
  - Opinions have been mixed
  - Special team (a “scrum”) is revisiting this decision.

Using FA for Occurrence Questions (2)

- In the published Dyslipidemia TAUG, this “AE Rule” was applied when representing the responses to pre-specified anti-Dyslipidemia medications:

Row	STUDYID	DOMAIN	USUBJID	FASEQ	FATESTCD	FATEST	FAOBJ	FACAT
1	XYZ987	FA	XY87-123	1	OCCUR	Occurrence Indicator	SUPRASTATIN	ANTI-DYSLIPIDEMIC TREATMENT
2	XYZ987	FA	XY87-123	2	OCCUR	Occurrence Indicator	VARASTATIN	ANTI-DYSLIPIDEMIC TREATMENT
3	XYZ987	FA	XY87-123	3	OCCUR	Occurrence Indicator	TOPOSTATIN	ANTI-DYSLIPIDEMIC TREATMENT

Row	FAORRES	FASTRESC	FASTAT	FAREASND	FADTC
1	Y	Y			2013-10-27
2	Y	Y			2013-10-27
3	N	N			2013-10-27

Using FA for Occurrence Questions (3)

- In the CM domain, we only have records for the 2 medications that were actually taken.
- With this application of the “AE Rule” in other Interventions and Events domains, it effectively removes the “OCCUR” variable from the data domains.
- Note the inclusion of the --PRESP variable in the CM dataset.

Row	STUDYID	DOMAIN	USUBJID	CMSEQ	CMTRT	CMCAT	CMPRESP	CMDOSTOT
1	ABC123	CM	XYZ987-123	1	SUPRASTATIN	ANTI-DYSLIPIDEMIC TREATMENT	Y	20
2	ABC123	CM	XYZ987-123	2	VARASTATIN	ANTI-DYSLIPIDEMIC TREATMENT	Y	20

Row	CMDOSU	CMDTC	CMSTDTC	CMENDTC	CMENRTPT	CMENTPT
1 (cont)	mg	2013-10-27	2012-01-01	2013-06-01		
3 (cont)	mg	2013-10-27	2013-07-15		ONGOING	SCREENING VISIT

15

COPD TAUG – Bringing the –OCCUR variable back into the data domain

Shows the Yes or No response in MHOCCUR to a series of pre-specified MHTERMs.

Row	STUDYID	DOMAIN	USUBJID	MHSEQ	MHTERM	MHEVTYP	MHCAT	MHPRESP	MHOCCUR	MHDTC	MHSTDTC
1	ABC-123	MH	101	1	EMPHYSEMA		COPD HISTORY	Y	N	2012-09-28	
2	ABC-123	MH	101	2	CHRONIC BRONCHITIS		COPD HISTORY	Y	N	2012-09-28	
3	ABC-123	MH	101	3	COPD-ASTHMA OVERLAP SYNDROME		COPD HISTORY	Y	Y	2012-09-28	2011-10-31

16

Findings About vs. Clinical Events

Findings About

- Snapshots of observations within an event
  - e.g., severity at various times within an event
- Usually has --DTC
  - --DTC different from --STDTC and --ENDTC
- Can handle multiple results and units

Clinical Events

- Capture information about the event as a whole
  - e.g., overall severity of the event
- Often have --STDTC and/or --ENDTC

Clinical Events Example from Diabetes TAUG

- In this example, CE is used to collect a pre-specified set of associated symptoms for a collected hypoglycemic event. The Diabetes TAUG acknowledges that FA could also be used.

Row	STUDYID	DOMAIN	USUBJID	CESEQ	CETERM	CECAT	CEDECOD	CEPRESP	CEOCCUR	CESTDTC
1	XYZ	CE	XYZ-001-001	1	HYPOGLYCEMIC EVENT	HYPO EVENTS	Hypoglycaemia			2013-09-01T11:00
2	XYZ	CE	XYZ-001-001	2	SWEATING	HYPO SYMPTOMS	Hyperhidrosis	Y	Y	
3	XYZ	CE	XYZ-001-001	3	TREMORS/ TREMBLING	HYPO SYMPTOMS	Tremor	Y	N	
4	XYZ	CE	XYZ-001-001	4	DIZZINESS	HYPO SYMPTOMS	Dizziness	Y	Y	

Row	RELMIDS	MIDS	MIDSDTC
1 (cont)		HYPO 1	
2 (cont)	DURING	HYPO 1	2013-09-01T11:00
3 (cont)	DURING	HYPO 1	2013-09-01T11:00
4 (cont)	DURING	HYPO 1	2013-09-01T11:00

New Domain Specific Variable - MHEVTYP

- Concept being introduced in soon-to-be-published SDTM v1.5
- MHEVTYP – Medical History Event Type
  - Specifies the aspect of the medical condition as shown in MHTERM by which MHSTDTC is defined
    - Date of Diagnosis
    - Date of Relapse
    - Date of Initial Symptoms
    - Most recent episode
- Used in the Schizophrenia and Hepatitis C TAUGs
- Also modeled in the current version of the COPD TAUG – now being readied for public review

COPD – Date of Onset of Symptoms, Date of Diagnosis

Row	STUDYID	DOMAIN	USUBJID	MHSEQ	MHTERM	MHEVTYP	MHCAT
1	ABC-123	MH	101	1	CHRONIC OBSTRUCTIVE PULMONARY DISEASE	SYMPTOMS	COPD HISTORY
2	ABC-123	MH	101	2	CHRONIC OBSTRUCTIVE PULMONARY DISEASE	DIAGNOSIS	COPD HISTORY

Row	MHPRESP	MHOCCUR	MHDTC	MHSTDTC
1 (Con't)	Y	Y	2012-09-28	2010-04-01
2 (Con't)	Y	Y	2012-09-28	2011-10-31

COPD – Most recent Exacerbation (1)

- CRF captures the question “Has the subject experienced any COPD exacerbations in the previous 12 months?”
- CRF then captures the date that the most recent exacerbation ended as well as the number of “mild, moderate, and severe” exacerbations over that time period.

Row	STUDYID	DOMAIN	USUBJID	MHSEQ	MHLNKID	MHTERM	MHEVTYP	MHCAT	MHPRESP	MHOCCUR
1	XYZ123	MH	XYZ123-001	1	MH-03	COPD EXACERBATION	MOST RECENT	COPD EXACERBATION HISTORY	Y	Y

Row	MHDTC	MHDY	MHENDTC	MHENDY	MHEVLINT
1 (cont)	2013-10-28	-22	2013-09-13	-67	-P12M

COPD – Most recent Exacerbation (2)

- The number of exacerbations of each severity are represented in a *famh* dataset and then, using --LNKID as the IDVARVAL, a dataset to dataset RELREC relationship is defined (not shown).

Row	STUDYID	DOMAIN	USUBJID	FASEQ	FALNKID	FATESTCD	FATEST	FAOBJ	FACAT
1	XYZ123	FA	XYZ123-001	1	MH-03	EPIMILD	Number of Mild Episodes	COPD EXACERBATION	COPD EXACERBATION HISTORY
2	XYZ123	FA	XYZ123-001	2	MH-03	EPIMOD	Number of Moderate Episodes	COPD EXACERBATION	COPD EXACERBATION HISTORY
3	XYZ123	FA	XYZ123-001	3	MH-03	EPISEVR	Number of Severe Episodes	COPD EXACERBATION	COPD EXACERBATION HISTORY

Row	FAORRES	FASTRESC	FASTRESN	FADTC	FADY	FAEVLINT
1 (cont)	4	4	4	2013-10-28	-22	-P12M
2 (cont)	1	1	1	2013-10-28	-22	-P12M
3 (cont)	1	1	1	2013-10-28	-22	-P12M

Morphology vs. Physiology

- CDISC Foundational Teams (SDS, CDASH, ADaM, SEND) recently made a decision to combine these two concepts and to create separate domains for each body system that would contain both morphology and physiology findings
  - Difficult to separate morphology from physiology at times
  - Body system specific domains were already published/drafted (e.g. RP, OE, CV, NV, RE)
  - More implications than benefits to having just one morphology/physiology domain (e.g. difficult to manage assumptions for all body systems together, large CT lists)
- Impact to SDTMIG v3.3 and TAUGs:
  - Deprecate the Morphology Findings (MO) domain
  - Revise definitions/assumptions for existing body system domains to include both morphology and physiology

Musculoskeletal System Findings (MK)

- One of the first proposed body system domains that will combine morphology and physiology
- To be modeled in the Rheumatoid Arthritis TAUG (under development)

Row	STUDYID	DOMAIN	USUBJID	MKSEQ	MKTESTCD	MKTEST	MKCAT	MKSCAT	MKORRES
1	XYZ	MK	XYZ-002	1	JSN	Joint Space Narrowing	GENANT/SHARP JOINT SPACE NARROWING ASSESSMENT	HAND/WRIST JOINTS	MODERATE, 51-75% LOSS OF JOINT SPACE
2	XYZ	MK	XYZ-002	2	JSN	Joint Space Narrowing	GENANT/SHARP JOINT SPACE NARROWING ASSESSMENT	HAND/WRIST JOINTS	MODERATE- SEVERE, 76-95% LOSS OF JOINT SPACE
3	XYZ	MK	XYZ-002	3	JSN	Joint Space Narrowing	GENANT/SHARP JOINT SPACE NARROWING ASSESSMENT	HAND/WRIST JOINTS	

Row	MKSTRESC	MKSTRESN	MKSTAT	MKREASND	MKLOC	MKLAT	MKMETHOD	VISITNUM	MKDTC
1	2	2			INTERPHALANGEAL THUMB JOINT	RIGHT	X-RAY	4	2013-08-12
2	2.5	2.5			INTERPHALANGEAL THUMB JOINT	LEFT	X-RAY	4	2013-08-12
3			NOT DONE	AMPUTATION/MISSING ANATOMY/JOINT REPLACEMENT/ SURGICAL ALTERATION	PROXIMAL INTERPHALANGEAL JOINT 2	RIGHT	X-RAY	4	2013-08-12

Adjudication Data in TAUGs (1)

- Currently, the SDTMIG does not provide guidance for adjudicated events and findings
- Recently modeled in the CV TAUG v1 which took the following approach:
  - Separate records were created in the relevant domain in which the data occurred.
  - The following variables were added to Table 2.2.2 Events table in SDTM v1.5:
    - --EVAL: identifies the evaluator (e.g., INVESTIGATOR, CEC ADJUDICATOR), this can be sponsor-defined since the EVAL codelist controlled terminology is extensible.
    - --EVALID: Used to identify multiple evaluators that have the same role populated in --EVAL
    - --ACPTFL: Used to identify the accepted evaluation.
  - Separate sub-team to be formed to expand on modeling adjudication data and to address in a future publication.

Adjudication Data in TAUGs (2)

- From the CV TAUG

Row	STUDYID	DOMAIN	USUBJID	CESEQ	CEGRPID	CETERM	CEEVAL	CEACPTFL	CEDTC	CESTDTC
1	STUDY01	CE	40523	1	1	TRANSIENT ISCHEMIC ATTACK	INVESTIGATOR		2008-10-15	2008-10-15
2	STUDY01	CE	40523	2	1	TRANSIENT ISCHEMIC ATTACK	CEC ADJUDICATOR	Y	2008-11-15	2008-10-15

Non-Standard Variable Representation

- SDTMIG v3.3 Section 8.4.4 (Relating Non-Standard Variables Values to a Parent Domain) Revised
  - Incorporates content submitted as part of the publicly reviewed "SDS Proposal for Alternate Handling of Supplemental Qualifiers" (SDTMIG v3.3, Batch 2).
- Provides an alternative method for submitting NSVs in parent domain.
- Doesn't change the nature of the non-standard data submitted
- Complete metadata must be provided for each NSV
  - Variable lengths set to the appropriate length, as with all standard character variables
  - Roles to be used:
    - Non-Standard Identifier
    - Non-Standard Qualifier
    - Non-Standard Timing
- NSVs would be ordered after the standard variables, and ordered by Role

Submitting Non-Standard Variables: Example

Existing Method Using SUPPHO

ho.xpt

STUDYID	DOMAIN	USUBJID	HOSEQ	HOTERM	HOSTDTC	HOENDTC	HODUR
1999001	HO	0001	1	Hospital	2004-01-05	2004-01-12	P1W

suppho.xpt

STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL	QORIG	QEVAL
1999001	HO	0001	HOSEQ	1	HOAERPFL	AE Reported This Episode	Y	CRF	
1999001	HO	0001	HOSEQ	1	HOMEDSFL	Meds Prescribed	Y	CRF	
1999001	HO	0001	HOSEQ	1	HOPROCFL	Procedures Performed	Y	CRF	
1999001	HO	0001	HOSEQ	1	HOPROVNM	Provider Name	General Hosp	CRF	
1999001	HO	0001	HOSEQ	1	HOSPUFL	Any Time in Spec. Unit	Y	CRF	
1999001	HO	0001	HOSEQ	1	HOSPUTYP	Specialized Unit Type	ICU	CRF	
1999001	HO	0001	HOSEQ	1	HORLCNDF	Visit Related to Study Med Cond.	Y	CRF	

New Method Including Non-Standard Variables in Parent (HO) Domain

ho.xpt

STUDYID	DOMAIN	USUBJID	HOSEQ	HOTERM	HOSTDTC	HOENDTC	HODUR	HOAERPFL	HOMEDSFL
1999001	HO	0001	1	Hospital	2004-01-05	2004-01-12	P1W	Y	Y

HOPROCFL	HOPROVNM	HOSPUFL	HOSPUTYP	HORLCNDF
Y	General Hosp	Y	ICU	Y

Metadata Development Forum

- Forum to which TAUG team members across therapeutic areas can bring proposed domains, variables, and modeling strategies
  - Meets twice weekly
  - Any team member may add to the agenda and present at the meetings
  - Discuss modeling issues that may arise during the creation of SDTM examples for a specific therapeutic area
    - Depending on the data being collected, issues may apply to other therapeutic areas as well (e.g. –OCCUR vs FA, MHEVTYP)
  - Ensures that a proposed domain or variable is needed prior to the TAUG being published
  - Ensures better consistency in modeling SDTM data across the TAUGs

Conclusions

- The CFAST Initiative is a collaborative effort among CDISC, TransCelerate Biopharma, Inc, FDA, and the National Institute of Health in developing standards for the pharmaceutical industry via the publication of the TAUGs
- This has resulted in an acceleration of standards development in the TAUGs prior to being incorporated into the SDTM/SDTMIG
- Several new domains, variables, and modeling strategies have come from the development of the TAUGs
- In order to mitigate some of the risks of each TAUG team working independently, the Metadata Development Forum was created.

Questions?

Jerry Salyers
Accenture Accelerated R&D Services
E-mail: jerry.j.salyers@accenture.com

Kristin Kelly
Accenture Accelerated R&D Services
E-mail: kristin.c.kelly@accenture.com

Fred Wood
Accenture Accelerated R&D Services
E-mail: f.wood@accenture.com