PharmaSUG 2020 - Paper DS-133

Is Your Data Set Analysis Ready?

Nancy Brucken, Clinical Solutions Group, Inc.;
Kapila Patel, Syneos Health™

ABSTRACT

One of the fundamental principles of the Analysis Data Model (ADaM) is that analysis data sets should be analysis-ready, which means that each item displayed on an output table, listing or figure can be generated from an analysis data set with minimal programming, according to the ADaM Implementation Guide (ADaMIG). The analysis data sets produced by assuming a 1-to-1 relationship between Study Data Tabulation Model (SDTM) and ADaM data sets may not be analysis-ready, especially if the actual analysis requirements have not been considered. This paper will provide several examples of how an SDTM domain can be split into more than one ADaM data set to meet analysis needs and show that the SDTM domain class does not have to drive the class of the resulting ADaM data sets.

INTRODUCTION

ADaM data sets should have a structure and content that allow statistical analyses to be performed with minimal changes. Such data sets are considered to be “analysis-ready”. That is, ADaM data sets should contain all of the data needed for the review and re-creation of specific statistical analyses performed for a clinical study.

RELATIONSHIP BETWEEN SDTM AND ADAM

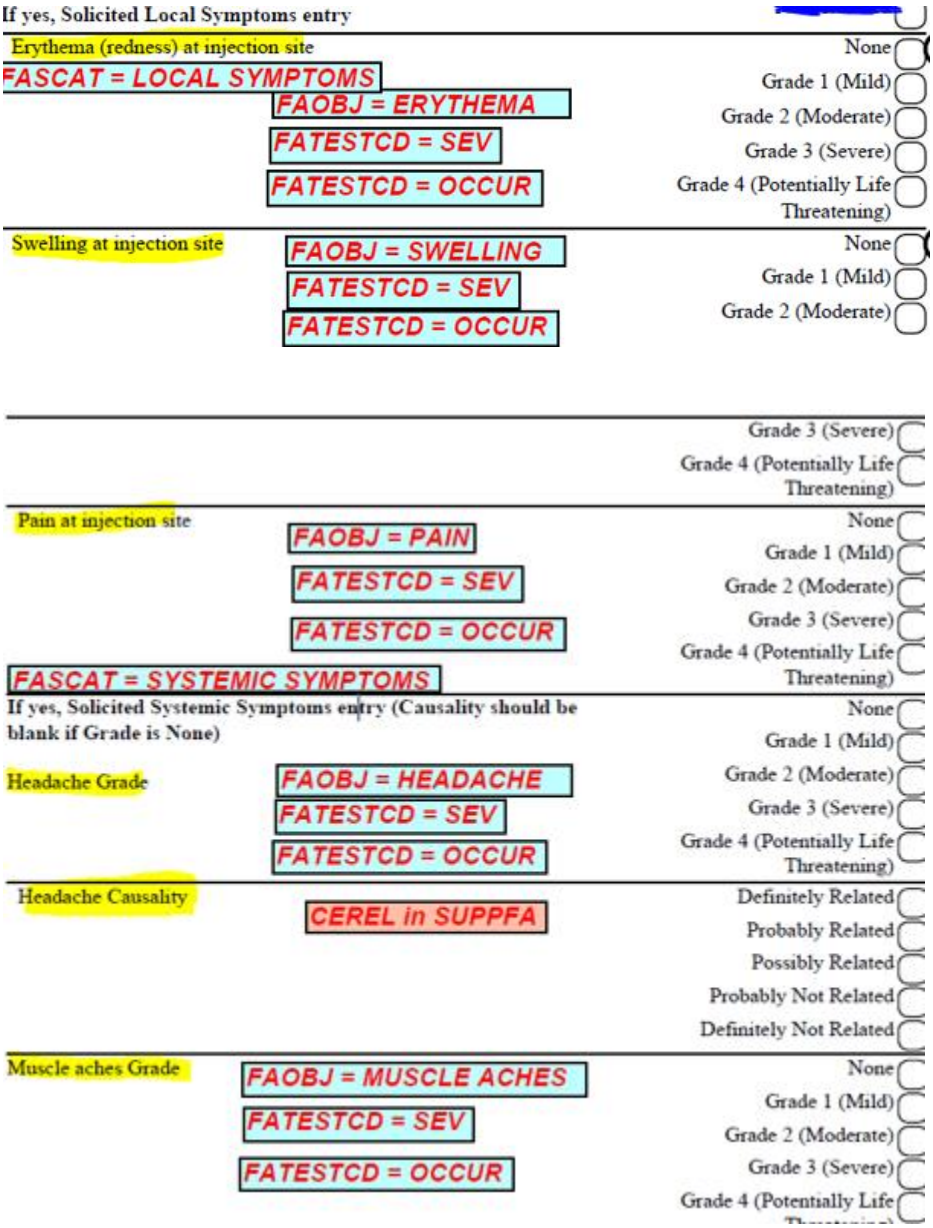
The ADaMIG does not mandate, or even suggest, that there should be a 1-to-1 relationship between SDTM domains and ADaM data sets. SDTM domains can be split into multiple ADaM data sets, and ADaM data sets can have multiple SDTM domains or other ADaM data sets as their source. Only data used for analysis, plus a few required variables, needs to be included in an ADaM data set, along with any other information that provides traceability back to the source of the record.

The recommended approach to designing ADaM data sets is to start at the end of the process, with the Tables, Listings and Figures (TLFs), and determine what data will be required in order to generate them. In other words, ADaM should pull from SDTM (or other ADaM data sets); SDTM does not push into ADaM. Following the latter strategy may result in ADaM data sets that are not analysis-ready, as we will demonstrate in subsequent examples.

FA- A CATCH-ALL DOMAIN

The FA (Findings About) domain in SDTM can contain data about multiple kinds of Findings in the study. Creating one ADaM data set called ADFA for all the TLFs related to FA domain in the study is often not the most efficient way to make the data set analysis ready. The example below shows a CRF used for an influenza vaccine study. Two ADaM data sets were created from the FA domain in SDTM, one to produce a table, and the other for a listing.

The CRF in Display 1 captures local and systemic symptoms in subjects after they received the vaccination. The CRF was mapped to FA domain.



Display 1: CRF for Local and Systemic Symptoms

To program table shown in Display 2 (next page), an ADaM data set called ADREACTN (Solicited Reactions Occur Analysis Data set) was created. Since the table counts the number of times a subject reported local and systemic symptoms, the ADaM Occurrence Data Structure (OCCDS) was used. In this data set, only the data required for the table and some common variables were taken from the FA domain in SDTM. The SDTM variable FA.FAOBJ was mapped to the ADaM variable ATERM (Analysis Term), and the SDTM variable FA.FAORRES was mapped to ADaM variable ASEV (Analysis Severity/Intensity). Occurrence flags were used to produce the counts in the table. AOCCFL was used to flag the first complaint within each subject, AOCC01FL indicated the first occurrence of a local or systemic reaction, AOCC02FL indicated the first occurrence of each complaint, and AOCCIFL marked the maximum severity of each complaint.

Table 1.1.1.1
Summary of Immediate Complaints:
Safety Analysis Set **SAFFL='Y'**

Age Stratum: Overall **AGEGR1**

	30 µg Drug A (N=xx) n (%)	Drug B (N=xx) n (%)
Subjects with at Least One Immediate Complaint		
Total **AOCCFL='Y'**	x (xx.x)	x (xx.x)
Subjects with at Least One Immediate Local Reaction		
AOCC01FL='Y' and ACAT1='LOCAL SYMP'	x (xx.x)	x (xx.x)
Erythema at Injection Site **ATERM where AOCC02FL='Y'**	x (xx.x)	x (xx.x)
Swelling at Injection Site	x (xx.x)	x (xx.x)
Pain at Injection Site	x (xx.x)	x (xx.x)
Subjects with at Least One Immediate Systemic Reaction		
AOCC01FL='Y' and ACAT1='SYSTEMIC SYMP'	x (xx.x)	x (xx.x)
Fever[a]	x (xx.x)	x (xx.x)
Headache	x (xx.x)	x (xx.x)
Muscle Ache	x (xx.x)	x (xx.x)
Joint aches	x (xx.x)	x (xx.x)
Fatigue	x (xx.x)	x (xx.x)
Chills	x (xx.x)	x (xx.x)
Malaise	x (xx.x)	x (xx.x)
Swelling in the Neck	x (xx.x)	x (xx.x)
Swelling in the Axilla	x (xx.x)	x (xx.x)

Fever is based on abnormal oral temperature (≥ 38.0 °C).
Note: Immediate Complaints is defined as any solicited local and systemic reactions occurring within 15 minutes after vaccination.
Repeat the table for Age Stratum: 65-74 and Age Stratum: >75.

Source data set: ADREACTN

Display 2: Summary of Immediate Complaints

USUBJID	TRTP	FASEQ	FATPT	ATERM	ACAT1	ADT	AREL	RELGR1	ATOXGR	AOCCFL	AOCC01FL	AOCC02FL	AOCCIFL
101	Drug A	72		FATIGUE	SYSTEMIC SYMP	14-Oct-18	POSSIBLY RELATED	RELATED	MILD			Y	
101	Drug A	66		HEADACHE	SYSTEMIC SYMP	14-Oct-18	POSSIBLY RELATED	RELATED	MILD	Y	Y	Y	Y
101	Drug A	70		JOINT ACHES	SYSTEMIC SYMP	14-Oct-18	POSSIBLY RELATED	RELATED	MILD			Y	
101	Drug A	68		MUSCLE ACHES	SYSTEMIC SYMP	14-Oct-18	POSSIBLY RELATED	RELATED	MILD			Y	
102	Drug B	16	DAY 0 - EVE	PAIN	LOCAL SYMP	10-Oct-18	DEFINITELY RELATED	RELATED	MILD	Y	Y	Y	Y
102	Drug B	29		PAIN	LOCAL SYMP	11-Oct-18	DEFINITELY RELATED	RELATED	MILD				
102	Drug B	42		PAIN	LOCAL SYMP	12-Oct-18	DEFINITELY RELATED	RELATED	MILD				
103	Drug A	4	DAY 0 - 15 MIN	PAIN	LOCAL SYMP	10-Oct-18	DEFINITELY RELATED	RELATED	MILD	Y	Y	Y	Y
103	Drug A	17	DAY 0 - EVE	PAIN	LOCAL SYMP	10-Oct-18	DEFINITELY RELATED	RELATED	MILD				
104	Drug A	16	DAY 0 - EVE	PAIN	LOCAL SYMP	11-Oct-18	DEFINITELY RELATED	RELATED	MILD	Y	Y	Y	Y
105	Drug A	17	DAY 0 - EVE	PAIN	LOCAL SYMP	11-Oct-18	DEFINITELY RELATED	RELATED	MILD		Y	Y	
105	Drug A	30		PAIN	LOCAL SYMP	12-Oct-18	DEFINITELY RELATED	RELATED	MILD				
105	Drug A	43		PAIN	LOCAL SYMP	13-Oct-18	DEFINITELY RELATED	RELATED	MILD				
105	Drug A	10	DAY 0 - 15 MIN	CHILLS	SYSTEMIC SYMP	11-Oct-18	PROBABLY RELATED	RELATED	MILD	Y	Y	Y	Y
106	Drug A	18	DAY 0 - EVE	HEADACHE	SYSTEMIC SYMP	11-Oct-18	PROBABLY RELATED	RELATED	MILD	Y	Y	Y	Y
106	Drug A	94		JOINT ACHES	SYSTEMIC SYMP	17-Oct-18	POSSIBLY RELATED	RELATED	MILD			Y	
106	Drug A	92		MUSCLE ACHES	SYSTEMIC SYMP	17-Oct-18	POSSIBLY RELATED	RELATED	MILD			Y	

Display 3: ADREACTN (Solicited Reactions Occur Analysis Data set)

Display 4 shows the listing of solicited local and systemic reactions over Days 0-7 for the study. The listing displayed age stratum and treatment for subjects in the safety population. Since these variables were not present in the FA domain, we created an ADaM dataset to pull in this information. In order to efficiently program the listing, the ADaM data set ADREACTL (Solicited Reactions Analysis data set) was created in a Basic Data Structure (BDS) format. The SDTM variable FA.FAOBJ was mapped to the ADaM variable PARAM (Parameter) and the SDTM variable FA.FASTRESC was mapped to the ADaM variable AVALC (Analysis Value (C)).

Listing 1.1.1.1
Solicited Local and Systemic Reactions (Days 0 to 7):
Safety Analysis Set **SAFFL='Y'**

				Day 0 **AVISIT**																			
							ATPT			Day 1				Day2		Day 3		Day 4		Day 5		Day 6		Day 7	
Subject	Planned /Actual Treatment	Age Stratum	Symptom	Pre[a]	C[b]	Post[a]	C[b]	PM[a]	C[b]		C[b]		C[b]		C[b]		C[b]		C[b]		C[b]		C[b]
USUBJID	**TRTA**	**AGEGR1**	**PARAM**	**AVAL**																			
101001	Drug A	65-74	Fever[c]	36.6		37.0		36.3		38.0	2	36.4		36.2		36.3		36.6		37.0		36.7	
			Erythema[d]			0		1	5	1	5	0		0		0		0		0		0	
			Swelling[d]			0		1	5	2	5	0		0		0		0		0		0	
			Pain[d]			1	5	2	5	1	5	0		0		0		0		0		0	
			Headache			0		0		0		2	3	1	3	0		0		0		0	
			Fatigue			0		0		0		0		0		0		0		0		0	
			Muscle Aches			0		0		0		0		0		0		0		0		0	
			Joint Aches			0		0		0		0		0		0		0		0		0	
			Chills			0		0		0		0		0		0		0		0		0	
			Malaise			0		0		0		0		0		0		0		0		0	
			Swelling Axilla			0		0		0		0		0		0		0		0		0	
			Swelling Neck			0		0		0		0		0		0		0		0		0	

…

Day 0 timepoints: Pre = pre-vaccination; Post = 15 minutes post-vaccination; PM = evening.
C=Causality
Fever is based on abnormal oral temperature (≥38.0 °C).
Local (injection site) reactions.
Note: Severity: 0-None, 1-Mild, 2-Moderate, 3-Severe, 4-Potentially Life-threatening; Causality (C): 1-Definitely Not Related, 2-Probably Not Related, 3-Possibly Related, 4-Probably Related, 5-Definitely Related.
Source dataset: ADREACTL. Display AVAL for severity parameters in the first column for each visit/timepoint, and AVAL for causality parameters in the 'C' column.

Display 4: Listing of Solicitated Local and Systemic Reactions

Display 5 shows the ADREACTL dataset used to produce this listing.

USUBJID	TRTA	AGEGR1	SAFFL	AVISIT	ADT	ADY	PARAM	PARAMCD	PARCAT1	AVALC
101	Drug A	65 - 74	Y	Day 1 Observation	14-Oct-18	2	ERYTHEMA	OCCERYTH	LOCAL SYMP	N
101	Drug A	65 - 74	Y	Day 1 Observation	14-Oct-18	2	PAIN	OCCPAIN	LOCAL SYMP	N
101	Drug A	65 - 74	Y	Day 1 Observation	14-Oct-18	2	SWELLING	OCCSWELL	LOCAL SYMP	N
101	Drug A	65 - 74	Y	Day 1 Observation	14-Oct-18	2	AXILLA SWELLING	OCCAXILL	SYSTEMIC SYMP	N
101	Drug A	65 - 74	Y	Day 1 Observation	14-Oct-18	2	CHILLS	OCCCHILL	SYSTEMIC SYMP	N
101	Drug A	65 - 74	Y	Day 1 Observation	14-Oct-18	2	FATIGUE	OCCFATIG	SYSTEMIC SYMP	Y
101	Drug A	65 - 74	Y	Day 1 Observation	14-Oct-18	2	FEVER	OCCFEVER	SYSTEMIC SYMP	N
101	Drug A	65 - 74	Y	Day 1 Observation	14-Oct-18	2	HEADACHE	OCCHEADA	SYSTEMIC SYMP	Y
101	Drug A	65 - 74	Y	Day 1 Observation	14-Oct-18	2	JOINT ACHES	OCCJOINT	SYSTEMIC SYMP	Y
101	Drug A	65 - 74	Y	Day 1 Observation	14-Oct-18	2	MALAISE	OCCMALAI	SYSTEMIC SYMP	N
101	Drug A	65 - 74	Y	Day 1 Observation	14-Oct-18	2	MUSCLE ACHES	OCCMUSCL	SYSTEMIC SYMP	Y
101	Drug A	65 - 74	Y	Day 1 Observation	14-Oct-18	2	NECK SWELLING	OCCNECK	SYSTEMIC SYMP	N
101	Drug A	65 - 74	Y	Day 2 Observation	15-Oct-18	3	ERYTHEMA	OCCERYTH	LOCAL SYMP	N
101	Drug A	65 - 74	Y	Day 2 Observation	15-Oct-18	3	PAIN	OCCPAIN	LOCAL SYMP	N
101	Drug A	65 - 74	Y	Day 2 Observation	15-Oct-18	3	SWELLING	OCCSWELL	LOCAL SYMP	N
101	Drug A	65 - 74	Y	Day 2 Observation	15-Oct-18	3	AXILLA SWELLING	OCCAXILL	SYSTEMIC SYMP	N
101	Drug A	65 - 74	Y	Day 2 Observation	15-Oct-18	3	CHILLS	OCCCHILL	SYSTEMIC SYMP	N
101	Drug A	65 - 74	Y	Day 2 Observation	15-Oct-18	3	FATIGUE	OCCFATIG	SYSTEMIC SYMP	N
101	Drug A	65 - 74	Y	Day 2 Observation	15-Oct-18	3	FEVER	OCCFEVER	SYSTEMIC SYMP	N
101	Drug A	65 - 74	Y	Day 2 Observation	15-Oct-18	3	HEADACHE	OCCHEADA	SYSTEMIC SYMP	N
101	Drug A	65 - 74	Y	Day 2 Observation	15-Oct-18	3	JOINT ACHES	OCCJOINT	SYSTEMIC SYMP	N
101	Drug A	65 - 74	Y	Day 2 Observation	15-Oct-18	3	MALAISE	OCCMALAI	SYSTEMIC SYMP	N
101	Drug A	65 - 74	Y	Day 2 Observation	15-Oct-18	3	MUSCLE ACHES	OCCMUSCL	SYSTEMIC SYMP	N
101	Drug A	65 - 74	Y	Day 2 Observation	15-Oct-18	3	NECK SWELLING	OCCNECK	SYSTEMIC SYMP	N

Display 5: ADREACTL (Solicited Reactions Analysis Data set)

The above example shows that it is okay to split one SDTM domain into multiple ADaM data sets in order to meet the analysis needs, and to maintain the traceability of data.

ANOTHER FA TRAP

Another case where simply copying FA over into an “ADFA” and adding variables from ADSL led to problems. The compound was being studied for its effect on the frequency and severity of migraine headaches. FA contained information about the subject’s history of migraines, along with data describing various attributes of the migraines that occurred during the study. There were multiple values for FACAT, and 57 different FATESTCD values in FA. Thus, the initial version of “ADFA” contained 57 parameters, where all FATEST/FATESTCD values were copied over directly into PARAM/PARAMCD, and ADSL variables were merged in. However, while developing programming specifications for the table, it was discovered that none of the required tables could be produced from this data set.

What happened? Upon closer inspection, it was determined that FA actually contained the following different types of data:

- Migraine history questionnaire
- Event-level information on each qualifying migraine attack, collected on the CRF
- Electronic diary (ePRO) data collected every 30 minutes for several hours following study drug administration after onset of a qualifying migraine attack

Each of these types of data was summarized and analyzed separately. The following information was reported on the tables:

- Summary of migraine history questionnaire responses
- Summaries of migraine event-level information
- Summaries of ePRO data by 30-minute interval
- Summaries of derived ePRO data by migraine attack number (first, fifth, etc.) and by study month

Since each of these tables summarized different data at a different level of detail, we decided to split the information contained in FA into four separate analysis data sets:

- BDS data set containing migraine history questionnaire responses, as one record per subject per question. This instrument was collected once for each subject, at the beginning of the study.
- BDS containing migraine attack-level CRF data. This data set consisted of one record per subject per migraine attack per migraine attribute.
- BDS data set containing ePRO migraine data collected every 30 minutes after study drug administration. This data set consisted of one record per subject per migraine attack per 30-minute interval per attribute.
- BDS data set containing summarized/derived ePRO migraine parameters. This data set consisted of one record per subject per migraine attack per parameter.

All of the required tables could be generated directly from one of these data sets, so at this point, we had analysis-ready data sets. However, in order to get there, we had to start at the end, with the tables, and work our way back through determining what types of collected data were needed in order to generate each table.

BEWARE OF CUSTOM FINDINGS DOMAINS, TOO

Let's go into the details of one more example of multiple ADAM data sets created from one SDTM domain. For an anti-epilepsy trial, a subject diary collected information about seizures and other non seizure related information from the subject. This data was maintained in a single custom SDTM domain. However, TLFs summarized the data by seizure type. Hence, separate ADaM data sets were created from one SDTM domain for ease of traceability between the TLFs and the analysis data sets.

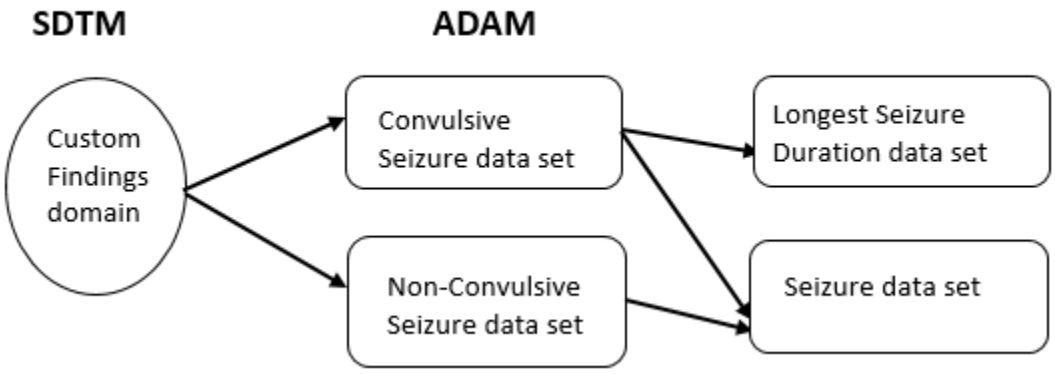
Display 6 is a screenshot of subject diary data in SDTM. As you see, the diary contains information about the subject's eligibility, seizure attributes, seizure type and seizure characteristics. 'Seizure code' in the diary refers to whether the seizure was convulsive or non-convulsive.

XATEST	XACAT	XAORRES
SUBJECT CONTINUING IN THE STUDY	ELIGIBILITY	Yes
TOTAL SEIZURES IN FIRST 3 WEEKS OF BASELINE	ELIGIBILITY	10
TOTAL SEIZURES IN SECND 3 WEEKS OF BASELINE	ELIGIBILITY	20
END OF DAY REVIEW COMPLIANCE	ELIGIBILITY	45
TOTAL CONVULSIVE SEIZURES IN BASELINE	ELIGIBILITY	30
RESPONDENT	SEIZURE QUIZ	Caregiver
SEIZURES LASTING LONGER TWO MINUTES	SEIZURE QUIZ	FALSE
SEIZURE CODE	SEIZURE TYPE	2
WAS RESCUE MEDICATION GIVEN	SEIZURE DIARY	Yes, rescue medication A
DATE SEIZURE BEGAN	SEIZURE DIARY	12/31/2019
TYPE OF SEIZURE	SEIZURE DIARY	Generalized bilateral symmetric clonic activity.
HOW LONG SEIZURE LASTED	SEIZURE DIARY	2 minutes
SINGLE OR REPEATED SEIZURE	SEIZURE DIARY	A single seizure
TIME SEIZURE BEGAN	SEIZURE DIARY	Early Morning (12:00am - 7:59am)
ANOTHER SEIZURE TO REPORT	END OF DAY	No

Display 6: Subject diary data

ADAM DATA SETS

Since the diary data domain included seizure as well as non-seizure data, we decided to pull out the seizure information separately, and create two intermediate data sets, one for convulsive seizures, and the other for non-convulsive seizures. These intermediate data sets serves as an input dataset for Seizure data set and Longest Seizure Duration data set.



a) Convulsive Seizures data set: This data set stores convulsive seizure data from the diary. This data set is structured as one record per subject per seizure per parameter.

PARAM	PARAMCD	PARCAT1	PARCAT2	AVALC	APERIODC	ASPERC
How Long Seizure Lasted	LASSEZ1	SEIZURE DIARY	Convulsive	Less than 2 minutes	TREATMENT	TITRATION
Date Seizure Began	SEZDAT1D	SEIZURE DIARY	Convulsive	1/1/2020	TREATMENT	TITRATION
Type Of Seizure	SEZTYP1L	SEIZURE DIARY	Convulsive	Abrupt eye opening, tonic	TREATMENT	TITRATION
Single Or Repeated Seizure	SNGRPT1L	SEIZURE DIARY	Convulsive	A single seizure	TREATMENT	TITRATION
Time Seizure Began	TIMSEZ1L	SEIZURE DIARY	Convulsive	Early Morning (12:00am - 7:59am)	TREATMENT	TITRATION
How Long Seizure Lasted	LASSEZ1	SEIZURE DIARY	Convulsive	Less than 2 minutes	TREATMENT	TITRATION
Date Seizure Began	SEZDAT1D	SEIZURE DIARY	Convulsive	1/15/2020	TREATMENT	TITRATION
Type Of Seizure	SEZTYP1L	SEIZURE DIARY	Convulsive	mouth pulls to the side, eyes	TREATMENT	TITRATION
Single Or Repeated Seizure	SNGRPT1L	SEIZURE DIARY	Convulsive	A single seizure	TREATMENT	TITRATION
Time Seizure Began	TIMSEZ1L	SEIZURE DIARY	Convulsive	Evening (6:00pm - 11:59pm)	TREATMENT	TITRATION
How Long Seizure Lasted	LASSEZ1	SEIZURE DIARY	Convulsive	Less than 2 minutes	TREATMENT	MAINTENANCE
Date Seizure Began	SEZDAT1D	SEIZURE DIARY	Convulsive	3/17/2020	TREATMENT	MAINTENANCE
Type Of Seizure	SEZTYP1L	SEIZURE DIARY	Convulsive	mouth pulls to the side, eyes	TREATMENT	MAINTENANCE
Single Or Repeated Seizure	SNGRPT1L	SEIZURE DIARY	Convulsive	A single seizure	TREATMENT	MAINTENANCE
Time Seizure Began	TIMSEZ1L	SEIZURE DIARY	Convulsive	Evening (6:00pm - 11:59pm)	TREATMENT	MAINTENANCE

Display 7: Convulsive Seizures data set

b) Non-Convulsive Seizures data set: This data set stores non-convulsive seizure data from the diary. This data set is structured as one record per subject per seizure per parameter.

PARAM	PARAMCD	PARCAT1	PARCAT2	AVALC	APERIODC	ASPERC
How Long Seizure Lasted	LASSEZ1	SEIZURE DIARY	Non-Convulsive	0:10	TREATMENT	TITRATION
Date Seizure Began	SEZDAT1D	SEIZURE DIARY	Non-Convulsive	12/31/2019	TREATMENT	TITRATION
Type Of Seizure	SEZTYP1L	SEIZURE DIARY	Non-Convulsive	Quick body twitch that can be felt in her body.	TREATMENT	TITRATION
Single Or Repeated Seizure	SNGRPT1L	SEIZURE DIARY	Non-Convulsive	A cluster of seizures back to back	TREATMENT	TITRATION
Time Seizure Began	TIMSEZ1L	SEIZURE DIARY	Non-Convulsive	Early Morning (12:00am - 7:59am)	TREATMENT	TITRATION
How Long Seizure Lasted	LASSEZ1	SEIZURE DIARY	Non-Convulsive	0:30	TREATMENT	TITRATION
Date Seizure Began	SEZDAT1D	SEIZURE DIARY	Non-Convulsive	1/15/2020	TREATMENT	TITRATION
Type Of Seizure	SEZTYP1L	SEIZURE DIARY	Non-Convulsive	Quick body twitch that can be felt in her body.	TREATMENT	TITRATION
Single Or Repeated Seizure	SNGRPT1L	SEIZURE DIARY	Non-Convulsive	A cluster of seizures back to back	TREATMENT	TITRATION
Time Seizure Began	TIMSEZ1L	SEIZURE DIARY	Non-Convulsive	Evening (6:00pm - 11:59pm)	TREATMENT	TITRATION
How Long Seizure Lasted	LASSEZ1	SEIZURE DIARY	Non-Convulsive	Less than 2 minutes	TREATMENT	MAINTENANCE
Date Seizure Began	SEZDAT1D	SEIZURE DIARY	Non-Convulsive	2/10/2020	TREATMENT	MAINTENANCE
Type Of Seizure	SEZTYP1L	SEIZURE DIARY	Non-Convulsive	Staring and unresponsive. Cannot be counted.	TREATMENT	MAINTENANCE
Single Or Repeated Seizure	SNGRPT1L	SEIZURE DIARY	Non-Convulsive	A single seizure	TREATMENT	MAINTENANCE
Time Seizure Began	TIMSEZ1L	SEIZURE DIARY	Non-Convulsive	Early Morning (12:00am - 7:59am)	TREATMENT	MAINTENANCE

Display 8: Non-Convulsive Seizures data set

c) Seizure data set: This data set combined data from the convulsive and non-convulsive seizure data sets to program numerous tables. Convulsive seizures were summarized by the percent change in convulsive seizure frequency from baseline. Non-convulsive seizures were summarized by duration and number of occurrences per subject and change from baseline seizure frequency. The resulting data set is structured as one record per subject per parameter.

Table 3.3.3.3
Convulsive seizure frequency per 28 days: Summary statistics
mITT Population **MITTFL='Y'**

	Placebo **TRTP** (N = XX)			XYZ 0.2 mg (N = XX)			XYZ 0.4 mg (N = XX)		
Period	Value	Change from Baseline	% Change from Baseline	Value	Change from Baseline	% Change from Baseline	Value	Change from Baseline	% Change from Baseline
PARAMCD='CSFB'	**AVAL**	**CHG**	**PCHG**						
Baseline CSFB									
N	XX			XX			XX		
Mean	XX.X			XX.X			XX.X		
SD	XX.XX			XX.XX			XX.XX		
Median	XX.X			XX.X			XX.X		
Min	XX			XX			XX		
Max	XX			XX			XX		
Week6									
N	XX	XX	XX	XX	XX	XX	XX	XX	XX
Mean	XX.X	XX.X	XX.X	XX.X	XX.X	XX.X	XX.X	XX.X	XX.X
SD	XX.XX	XX.XX	XX.XX	XX.XX	XX.XX	XX.XX	XX.XX	XX.XX	XX.XX
Median	XX.X	XX.X	XX.X	XX.X	XX.X	XX.X	XX.X	XX.X	XX.X
Min	XX	XX	XX	XX	XX	XX	XX	XX	XX
Max	XX	XX	XX	XX	XX	XX	XX	XX	XX

Display 9: Convulsive Seizure frequency per 28 days

PARAM	PARAMCD	AVAL	PARAMTYP
Number of days non-missing diary data during Baseline	NMDBLDY	42	DERIVED
Number of days non-missing diary data during M	NMDMDY	83	DERIVED
Number of days non-missing diary data during T+M	NMDTPMDY	97	DERIVED
Number of days non-missing diary data till Week 6	NMDW6DY	42	DERIVED
Number of days non-missing diary data till Week 10	NMDW10DY	70	DERIVED
Number of days non-missing diary data till Week 14	NMDW14DY	98	DERIVED
Convulsive Seizure frequency per 28 days during Baseline	CSFB	17	DERIVED
Convulsive Seizure frequency per 28 days during M	CSFM	25	DERIVED
Convulsive Seizure frequency per 28 days during T+M	CSFTPM	18	DERIVED
Convulsive Seizure frequency per 28 days till Week 14	CSFW14	20	DERIVED
Change in convulsive seizure frequency per 28 days till Week 14	CCSFW14	1.5	DERIVED
%Change in convulsive seizure frequency per 28 days during M	PCCSFM	18	DERIVED
%Change in convulsive seizure frequency per 28 days during T+M	PCCSFTPM	4	DERIVED
%Change in convulsive seizure frequency per 28 days till Week 14	PCCSFW6	9	DERIVED
25% reduction in convulsive seizures, yes/no during T+M	PCCS25YN	0	DERIVED
50% reduction in convulsive seizures, yes/no during T+M	PCCS50YN	0	DERIVED
100% reduction in convulsive seizures, yes/no during T+M	PCCS00YN	0	DERIVED
Non-convulsive seizure frequency per 28 days during Baseline	NCSFB	6	DERIVED
Non-convulsive seizure frequency per 28 days during M	NCSFM	1	DERIVED
Non-convulsive seizure frequency per 28 days during T+M	NCSFTPM	2	DERIVED
Change in non-convulsive seizure frequency per 28 days during T+M	CHGNCTPM	-3	DERIVED
%Change in non-convulsive seizure frequency per 28 days during M	PNCSFM	-20	DERIVED
%Change in non-convulsive seizure frequency per 28 days during T+M	PNCSFTPM	-30	DERIVED

Display 10: Seizure Data set

d) Longest Seizure Duration data set: This data set was created to find the longest seizure-free interval for a subject. The source for this data set is the convulsive seizure data set as shown is Display 7. Records were created for each seizure so that the longest interval between convulsive seizures could easily be determined.

Table 2.2.2.2
Longest interval (days) between convulsive seizures
mITT Population **MITTFL='Y'**

	Placebo (N = XX)	Drug XYZ 0.2 mg (N = XX)	Drug XYZ 0.4 mg (N = XX)
AVAL where PARAMCD='LONINTTM'			
Mean	XX.x	XX.x	XX.x
SD	XX.xx	XX.xx	XX.xx
Min	XX	XX	XX
25th Percentile	XX.X	XX.X	XX.X
Median	XX.X	XX.X	XX.X
75th Percentile	XX.X	XX.X	XX.X
Max	XX	XX	XX
Estimate of Median Treatment Difference*		XX	XX
95% CI for Treatment Difference*		(xx.x, xx.x)	(xx.x, xx.x)
p-value**		0.xxx	0.xxx

mITT = Modified Intent-to-treat; CI = Confidence Interval;
Programming notes:
For each subject, only 1 data point, their longest interval (days) between convulsive seizures, is used.
Source data set = ADLNGSZD

Display 11: Longest Seizure-Free Interval Table

PARAM	PARAMCD	AVAL	AVALC	DIARYY
Convulsive Seizure Present?	SEZYN	1	Y	Y
Convulsive Seizure Present?	SEZYN	0	N	Y
Convulsive Seizure Present?	SEZYN	0	N	Y
Convulsive Seizure Present?	SEZYN	0	N	Y
Convulsive Seizure Present?	SEZYN	0	N	Y
Convulsive Seizure Present?	SEZYN	1	Y	Y
Convulsive Seizure Present?	SEZYN	0	N	Y
Convulsive Seizure Present?	SEZYN	0	N	Y
Convulsive Seizure Present?	SEZYN	0	N	Y
Convulsive Seizure Present?	SEZYN	0	N	Y
Convulsive Seizure Present?	SEZYN	0	N	Y
Longest interval between convulsive seizures during T+M	LONINTTM	27		
Number of Seizure Occurrences during T+M	SEZNUMTM	10		

. **Display 12 : Longest Seizure Duration data set**

This example shows it is possible to form multiple ADaM data sets from one SDTM custom domain. This approach may serve to keep the data set simple and maintain the traceability, which would also help a reviewer replicate the data set, if needed.

CONCLUSION

When you have finished creating your analysis data sets, ask yourself these questions:

1. Can you create your TLFs from your analysis data sets without having to derive additional variables?
2. Can you easily get from your TLFs back to your analysis data sets, and from there back to the source records in SDTM?

If the answer to both of those questions is “Yes”, then your data sets are analysis-ready!

REFERENCES

CDISC Analysis Data Model Team. “Analysis Data Model Implementation Guide Version 1.2”. 2019. Available at https://www.cdisc.org/standards/foundational/adam.

ACKNOWLEDGMENTS

Special thanks to Vickie Zecca from Syneos Health™ for her help with the examples in this paper, and to Richann Watson for her thorough review.

CONTACT INFORMATION

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

Nancy Brucken
Clinical Solutions Group, Inc.
nbrucken@csg-inc.com

Kapila Patel
Syneos Health™
kapila.patel@syneoshealth.com