PharmaSUG 2023 - Paper DS-129

Have Meaningful Relationships: An Example of Implementing SDTM Special Purpose Dataset RELREC with a Many-to-Many Relationship

Kaleigh Ragan, Crinetics Pharmaceuticals;
Richann Watson, DataRich Consulting

ABSTRACT

The Related Records (RELREC) special purpose dataset is a tool, provided in the Study Data Tabulation Model (SDTM), for conveying relationships between records housed in different domains. Most SDTM users are familiar with a one-to-one relationship type, where a single record from one domain is related to a single record in a separate domain. Or even a one-to-many relationship type, where a single record from one domain may be related to a group of records in another. But what if there are two groups of records related to one another? How do you properly convey the relationship between these sets of data points? This paper aims to provide a clearer understanding of when and how to utilize the, not often encountered, many-to-many relationship type within the RELREC special purpose dataset.

INTRODUCTION

No data point is an island. A subject's data tells the story of their journey through a clinical trial. To fully understand that journey, it's important to consider all records for a given subject, as well as how those records might be related. That's where RELREC comes in. This dataset plays a critical role in capturing and reporting complex relationships between different study events, interventions, and outcomes. Understanding the relationships represented by RELREC is essential for accurately interpreting clinical trial data and ensuring regulatory compliance. In this paper, we will explore portraying these relationships, from the basic "one-to-one" relationship to far more complex "one-to-many", or indeed, "many-to-many" relationships in the SDTM dataset RELREC and their importance in clinical research.

UNDERSTANDING RELREC

PURPOSE OF RELREC

The purpose of the RELREC dataset is to illustrate the relationship between records across different domains for a single subject. The simplest example of related records is a concomitant medication (CM) given in response to an adverse event (AE) the subject experiences. To understand the full picture and context of these records within the subject's data, we need to be able to show the CM was given for, and thus related to, this AE. The RELREC dataset allows us to do this by capturing uniquely identifying information about the AE record and tying it to uniquely identifying information about the CM with a common identifier (RELID). The uniquely identifying information for the related records are captured in IDVAR, where IDVAR represents a variable in the domain that is used to identify the related record(s), and IDVARVAL, where IDVARVAL points to the record with the value of the identifying variable found in IDVAR. Thus, providing any reviewer a clear understanding of why that CM was administered to the subject.

DESCRIBING A MANY-TO-MANY RELATIONSHIP

In the simplest terms, a many-to-many relationship occurs when multiple values from a domain are related to multiple values in another domain. While these relationships are uncommon, most people will be familiar with this concept at the dataset level. For example, if the study has a single analyte concentration over time curve, then the pharmacokinetic concentration (PC) values collected for a subject would be related to all the pharmacokinetic parameters (PP) calculated for each subject as illustrated in section 6.3.5.9.3 in the SDTM Implementation Guide (SDTMIG) v3.4. For this paper, however, we will be discussing many-to-many relationships at the record level rather than the dataset level.

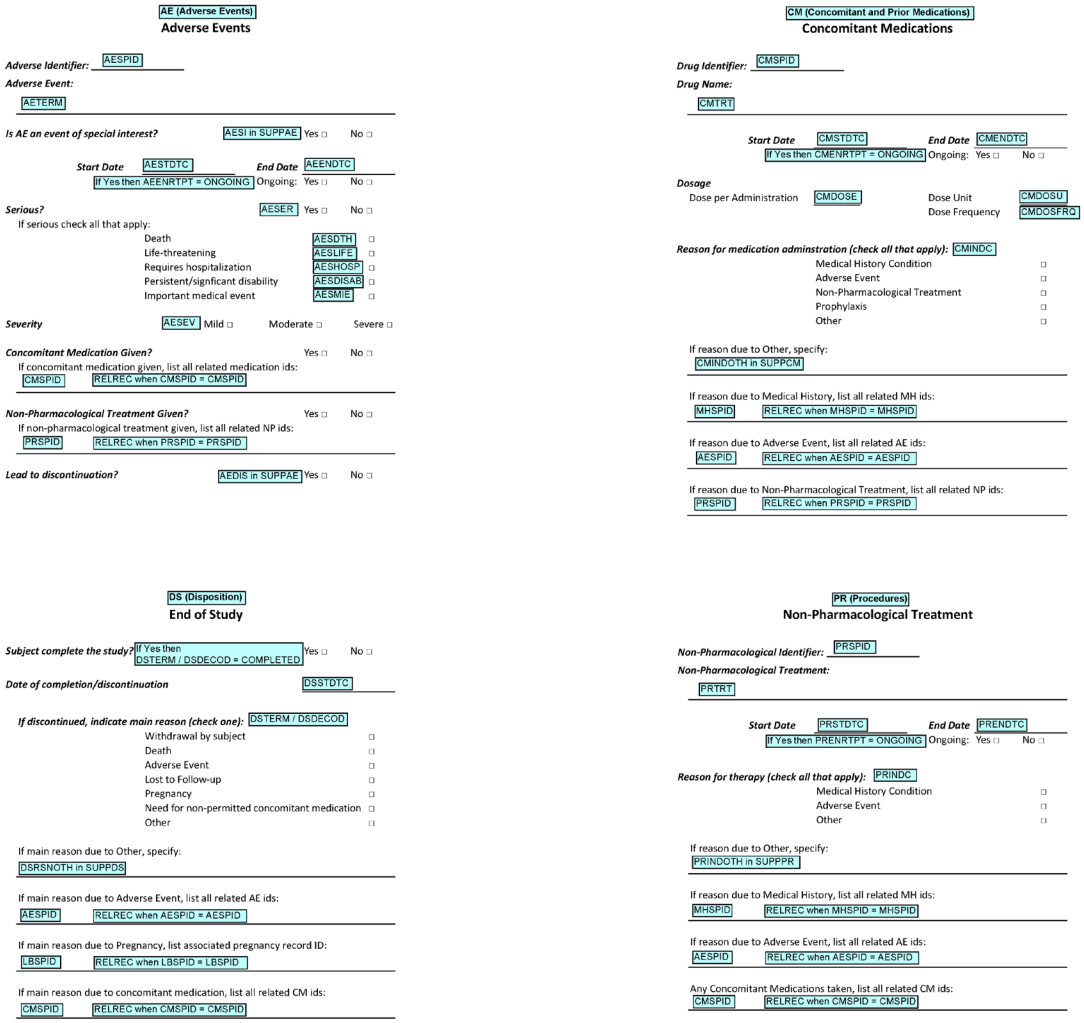
A many-to-many relationship at the record level may occur because the case report forms (CRFs) allow for collection of information that pertain to the relationship of data captured on other CRFs (e.g., medications

taken for an adverse event as well as any laboratory values related to that adverse event). This would mean that in the case of our example above, the adverse event(s) is related to the concomitant medications listed as well as the laboratory values reported, and the concomitant medications listed could potentially be related to the laboratory values. The additional correlation of the concomitant medication and the laboratory values are where conveying the relationship between all three domains in the RELREC dataset can become difficult. We need to include records from three different domains, with possibly multiple records per domain, for the relationship between them to be correctly reported.

SAMPLE DATASETS TO ILLUSTRATE A MANY-TO-MANY RELATIONSHIP

SAMPLE CASE REPORT FORM

Display 1 shows a sample of the CRFs that will be used to illustrate how the many-to-many relationship is established in the data. Notice that the adverse events (AE) CRF asks for a list of concomitant medications (CM) and non-pharmacological treatments (NP) given for the AE. In addition, the CRF has a question regarding whether the AE lead to discontinuation. Based on these questions, it is evident that there is going to be at least a 1-to-many relationship. However, if we look at the other CRFs, we see that similar questions are asked. For example, on the CM CRF, there is a question that asks for list of medical history (MH), AE and NP associated with the given medication. This is now establishing that many-to-many relationship, since an AE can be associated with multiple CMs, and a CM can be associated with multiple AEs.



Display 1: Sample of Annotated CRFs

SAMPLE SOURCE COLLECTED FROM CRF

Data Display 1 shows fictitious data for four subjects for the AE, CM, NP, end of study and pregnancy data. Since pregnancy data comes from a lab it does not have a corresponding CRF for this scenario. Each subject is color coded so that it is easy to see the relationship between the different data.

Note that only the data necessary to show the relationship between the different sources are displayed.

Row	STUDYID	SITEID	SUBJID	AENUM	AETERM	AESTDAT	AEENDAT	AEONGO	AESEV	AE_CMTRT	AE_CMID	AE_NPTRT	AE_NPID	AE_DISCONT
1	ABC	001	0001	1	Headache	2021-01-10	2021-05-24	No	Moderate	Yes	1	No		No
2	ABC	001	0001	2	Anxiety	2021-01-10	2021-05-24	No	Mild	Yes	2	No		No
3	ABC	001	0002	1	Anemia	2020-10-11	2020-11-20	No	Mild	No		No		No
4	ABC	001	0002	2	Post-partum bleeding	2021-11-15	2021-11-25	No	Moderate	Yes	1, 2	No		No
5	ABC	001	0003	1	Intermittent shortness of breath	2020-06-14	2021-02-05	No	Mild	Yes	2, 3	No		No
6	ABC	001	0003	2	Dyspnea	2021-03-04		Yes	Moderate	Yes	2, 5, 9	Yes	1, 2, 3, 4, 5	No
7	ABC	001	0003	3	Worsening of CAD	2021-04-09	2021-10-10	No	Moderate	Yes	4, 6, 7	Yes	3, 4, 5, 6	No
8	ABC	001	0003	4	Anxiety	2021-05-UN	2021-10-15	No	Mild	Yes	8	No		No
9	ABC	001	0004	1	Fatigue	2020-08-30	2021-01-19	No	Moderate	No		No		No
10	ABC	001	0004	2	Joint swelling right leg	2020-10-17	2021-08-16	No	Severe	Yes	8, 9	Yes	1, 2	No
11	ABC	001	0004	3	Vitamin B12 deficiency	2021-10-31	2022-09-12	No	Moderate	Yes	3	No		No
12	ABC	001	0004	4	Joint swelling right leg	2021-08-17		Yes	Mild	Yes	8, 9	Yes	1, 2	No
13	ABC	001	0004	5	Low bone mineral density	2022-08-29		Yes	Mild	Yes	4, 5, 6, 7	No		No
14	ABC	001	0004	6	Vitamin B12 deficiency	2022-09-13		Yes	Mild	Yes	3	No		No

Row	STUDYID	SITEID	SUBJID	CMNUM	CMTRT	CMSTDAT	CMENDAT	CMONGO	CMINDC	CM_MHID	CM_AEID	CM_NPID
1	ABC	001	0001	1	Ibuprofen	2021-01-10	2021-05-24	No	Adverse Event		1	
2	ABC	001	0001	2	Lexapro	2021-01-10	2021-05-24	No	Adverse Event		2	
3	ABC	001	0002	1	Iron Suplementation	2021-11-16	2021-11-24	No	Adverse Event		2	
4	ABC	001	0002	2	Folic Acid	2021-11-16	2021-11-24	No	Adverse Event		2	
5	ABC	001	0003	1	Naproxen	2020-02-15	2020-03-15	No	Prophylaxis			
6	ABC	001	0003	2	Albuterol	2020-06-14		Yes	Adverse Event		1, 2	
7	ABC	001	0003	3	prednisone	2020-06-14	2020-08-14	No	Adverse Event		1	
8	ABC	001	0003	4	plavix	2021-04-09		Yes	Adverse Event		3	
9	ABC	001	0003	5	Breo inhaler	2021-07-04	2021-11-08	No	Adverse Event		2	
10	ABC	001	0003	6	Aggrenox	2021-04-10	2021-10-10	No	Adverse Event		3	
11	ABC	001	0003	7	Aspirin	2021-04-10		Yes	Adverse Event		3	
12	ABC	001	0003	8	Lorazepam	2021-05-UN	2021-10-15	No	Adverse Event		4	
13	ABC	001	0003	9	Hydrochlorothiazide	2021-03-04	2021-04-19	No	Adverse Event		2	
14	ABC	001	0004	1	Metformin XR	2020-05-10	2020-11-14	No	Medical History Condition			
15	ABC	001	0004	2	Prednisone	2015-UN-UN		Yes	Medical History Condition			
16	ABC	001	0004	3	Vitamin B12	2021-10-31		Yes	Adverse Event		3, 6	
17	ABC	001	0004	4	Vitamin K2	2022-08-29		Yes	Adverse Event		5	
18	ABC	001	0004	5	Vitamin D3	2022-08-29		Yes	Adverse Event		5	
19	ABC	001	0004	6	Calcium malate citrate	2022-08-29		Yes	Adverse Event		5	
20	ABC	001	0004	7	Magnesium	2022-08-29		Yes	Adverse Event		5	
21	ABC	001	0004	8	Ibuprofen	2020-10-17		Yes	Adverse Event; Non-Pharmacological Treatment		2, 4	1, 2
22	ABC	001	0004	9	Meloxicam	2020-10-19		Yes	Adverse Event; Non-Pharmacological Treatment		2, 4	1, 2

Row	STUDYID	SITEID	SUBJID	NPNUM	NPTRT	NPSTDAT	NPENDAT	NPONGO	NPINDC	NPOTHSPEC	NP_MHID	NP_AEID	NP_CMID
1	ABC	001	0002	1	IUD	2014-UN-UN	2021-02-20	No	Prophylaxis				
2	ABC	001	0003	1	cardiac echocardiogram	2021-03-25	2021-03-25	No	Adverse Event			2	
3	ABC	001	0003	2	cardiac echocardiogram	2021-03-24	2021-03-24	No	Adverse Event			2	
4	ABC	001	0003	3	Pulmonary Function Test	2021-04-28	2021-04-28	No	Adverse Event			2, 3	
5	ABC	001	0003	4	Carotid ultrasound	2021-04-11	2021-04-11	No	Adverse Event			2, 3	
6	ABC	001	0003	5	cardiac catherization	2021-04-09	2021-04-09	No	Adverse Event			2, 3	
7	ABC	001	0003	6	outpatient cardiac rehab	2021-07-22		Yes	Adverse Event			3	
8	ABC	001	0004	1	Ankle-foot orthosis	2020-10-17		Yes	Adverse Event			2, 4	8, 9
9	ABC	001	0004	2	Compression socks	2020-10-17		Yes	Adverse Event			2, 4	8, 9

Row	STUDYID	SITEID	SUBJID	COMPLYN	DSSTDAT	DSREAS	DSREAS_OTH	DS_AEID	DS_PREGID	DS_CMID
1	ABC	001	0001	No	2021-05-23	Lost to Follow-up				
2	ABC	001	0002	No	2021-03-06	Pregnancy			UPREG6	
3	ABC	001	0003	Yes	2021-11-16					
4	ABC	001	0004	Yes	2022-12-15					

Row	STUDYID	SITEID	SUBJID	LBRECID	VISIT	LBPGDAT	PREG_RSLT	LB_DISCONT
1	ABC	001	0002	UPREG1	Day 1	2020-10-05	NEGATIVE	
2	ABC	001	0002	UPREG2	Week 4	2020-11-07	NEGATIVE	
3	ABC	001	0002	UPREG3	Week 8	2020-12-12	NEGATIVE	
4	ABC	001	0002	UPREG4	Week 12	2021-01-09	NEGATIVE	
5	ABC	001	0002	UPREG5	Week 16	2021-02-13	NEGATIVE	
6	ABC	001	0002	UPREG6	Week 20	2021-03-06	POSITIVE	Yes

Data Display 1: Sample Raw Data Source for Four Subjects

Note that the field names in the raw data source (Data Display 1) may not reflect CDASH. In addition, the sample CRFs and raw data do not reflect the typical way that a relationship between domains is captured. Normally, one CRF is the source for capturing across domain relationships and not both CRFs. For example, if there is a relationship between AE and CM, then either AE or CM CRF would capture the

necessary information to link the two domains. The information is not usually captured in both domains. However, in this illustration the relationship is captured on both CRFs which would require a reconciliation of the data to ensure that both raw data sources yield the same relationship.

SOURCE DATA AS SDTM DOMAINS

Table 1 contains a list of the SDTM domains that the CRF data could potentially be mapped to. Not all datasets contain examples. Only the datasets that are needed to illustrate the relationship between domains are provided. For example, MH, SUPPCM and SUPPPR are not shown but to fully map the data found on the four CRFs shown in Display 1, these domains would be needed. We will illustrate how the data from the source data is collected in the corresponding SDTM domain. Note that only the variables necessary to illustrate the concepts are displayed. However, all required and expected variables should be included.

DATASET	DESCRIPTION	CLASS	STRUCTURE	KEYS
AE	Adverse Events	EVENTS	One record per adverse event per subject	STUDYID, USUBJID, AEDECOD, AESTDTC
CM	Concomitant Medications	INTERVENTIONS	One record per recorded intervention occurrence or constant-dosing interval per subject	STUDYID, USUBJID, CMTRT, CMSTDTC
DS	Disposition	EVENTS	One record per disposition status or protocol milestone per subject	STUDYID, USUBJID, DSDECOD, DSSTDTC
LB	Laboratory Test Results	FINDINGS	One record per lab test per time point per visit per subject	STUDYID, USUBJID, LBTESTCD, LBSPEC, VISITNUM, LBTPTREF, LBTPTNUM
MH	Medical History	EVENTS	One record per medical history event per subject	STUDYID, USUBJID, MHDECOD
PR	Procedures	INTERVENTIONS	One record per recorded procedure per occurrence per subject	STUDYID, USUBJID, PRTRT, PRSTDTC
RELREC	Related Records	RELATIONSHIP	One record per related record, group of records or dataset	STUDYID, RDOMAIN, USUBJID, IDVAR, IDVARVAL, RELID
SUPPAE	Supplemental Qualifiers for AE	RELATIONSHIP	One record per supplemental qualifier per related parent domain record(s)	STUDYID, RDOMAIN, USUBJID, IDVAR, IDVARVAL, QNAM
SUPPCM	Supplemental Qualifiers for CM	RELATIONSHIP	One record per supplemental qualifier per related parent domain record(s)	STUDYID, RDOMAIN, USUBJID, IDVAR, IDVARVAL, QNAM
SUPPPR	Supplemental Qualifiers for PR	RELATIONSHIP	One record per supplemental qualifier per related parent domain record(s)	STUDYID, RDOMAIN, USUBJID, IDVAR, IDVARVAL, QNAM

Table 1: List of SDTM Domains Needed to Capture Source Data

The AE CRF is captured in the AE and SUPPAE datasets as shown in Data Display 2 and Data Display 3. Based on the AE CRF, we see there are three questions that indicate that the AE data may have a relationship to other data:

- Concomitant Medication Given?
- Non-Pharmacological Treatment Given?
- Lead to discontinuation?

The variable AECONTRT (Concomitant or Additional Trtmnt Given) is a Y (Yes) or N (No) variable that captures whether the subject was treated with a concomitant medication or non-pharmacological treatment for the specific AE. Capturing both in the same field, we are unable to identify if the subject took a concomitant medication or non-pharmacological treatment, we just know some intervention took place. With RELREC we will be able to see that relationship. However, before we can build RELREC we need to create the other domains.

Since there is no variable in the parent AE domain that captures leading to discontinuation, this data can be housed in the SUPPAE dataset as shown in Data Display 3.

Row	STUDYID	USUBJID	AESEQ	AESPID	AETERM	AESTDTC	AEENDTC	AESEV	AECONTRT
1	ABC	ABC-001-0001	1	1	Headache	2021-01-10	2021-05-24	MODERATE	Y
2	ABC	ABC-001-0001	2	2	Anxiety	2021-01-10	2021-05-24	MILD	Y
3	ABC	ABC-001-0002	1	1	Anemia	2020-10-11	2020-11-20	MILD	N
4	ABC	ABC-001-0002	2	2	Post-partum bleeding	2021-11-15	2021-11-25	MODERATE	Y
5	ABC	ABC-001-0003	1	1	Intermittent shortness of breath	2020-06-14	2021-02-05	MILD	Y
6	ABC	ABC-001-0003	2	2	Dyspnea	2021-03-04		MODERATE	Y
7	ABC	ABC-001-0003	3	3	Worsening of CAD	2021-04-09	2021-10-10	MODERATE	Y
8	ABC	ABC-001-0003	4	4	Anxiety	2021-05	2021-10-15	MILD	Y
9	ABC	ABC-001-0004	1	1	Fatigue	2020-08-30	2021-01-19	MODERATE	N
10	ABC	ABC-001-0004	2	2	Joint swelling right leg	2020-10-17	2021-08-16	SEVERE	Y
11	ABC	ABC-001-0004	3	3	Vitamin B12 deficiency	2021-10-31	2022-09-12	MODERATE	Y
12	ABC	ABC-001-0004	4	4	Joint swelling right leg	2021-08-17		MILD	Y
13	ABC	ABC-001-0004	5	5	Low bone mineral density	2022-08-29		MILD	Y
14	ABC	ABC-001-0004	6	6	Vitamin B12 deficiency	2022-09-13		MILD	Y

Data Display 2: Sample SDTM Adverse Event (AE) Dataset

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
1	ABC	AE	ABC-001-0001	AESEQ	1	AEDIS	AE Caused Study Discontinuation	N
2	ABC	AE	ABC-001-0001	AESEQ	2	AEDIS	AE Caused Study Discontinuation	N
3	ABC	AE	ABC-001-0002	AESEQ	1	AEDIS	AE Caused Study Discontinuation	N
4	ABC	AE	ABC-001-0002	AESEQ	2	AEDIS	AE Caused Study Discontinuation	N
5	ABC	AE	ABC-001-0003	AESEQ	1	AEDIS	AE Caused Study Discontinuation	N
6	ABC	AE	ABC-001-0003	AESEQ	2	AEDIS	AE Caused Study Discontinuation	N
7	ABC	AE	ABC-001-0003	AESEQ	3	AEDIS	AE Caused Study Discontinuation	N
8	ABC	AE	ABC-001-0003	AESEQ	4	AEDIS	AE Caused Study Discontinuation	N
9	ABC	AE	ABC-001-0004	AESEQ	1	AEDIS	AE Caused Study Discontinuation	N
10	ABC	AE	ABC-001-0004	AESEQ	2	AEDIS	AE Caused Study Discontinuation	N
11	ABC	AE	ABC-001-0004	AESEQ	3	AEDIS	AE Caused Study Discontinuation	N
12	ABC	AE	ABC-001-0004	AESEQ	4	AEDIS	AE Caused Study Discontinuation	N
13	ABC	AE	ABC-001-0004	AESEQ	5	AEDIS	AE Caused Study Discontinuation	N
14	ABC	AE	ABC-001-0004	AESEQ	6	AEDIS	AE Caused Study Discontinuation	N

Data Display 3: Sample SDTM Supplemental AE (SUPPAE) Dataset

Next, we look at the concomitant medication CRF and see that the reason for medication administration has several options that would indicate a possible relationship with other data:

- Medical History Condition
- Adverse Event
- Non-Pharmacological Treatment.

In Data Display 4, the variable CMINDC (Indication) is used to contain the reason(s) for medication administration. The remaining indications, Prophylaxis and Other, do not have a relationship to any other domain and thus are not included in the example RELREC data. Note that for subject ABC-001-0004 medication Ibuprofen (CMSPID = 8; CMSEQ = 3), this medication had two reasons listed, Adverse Event and Non-Pharmacological Treatment. This indicates that there is a relationship between the CM and the AE and PR data for this record. Note that NPs are captured in the PR domain.

Row	STUDYID	USUBJID	CMSEQ	CMSPID	CMTRT	CMINDC	CMSTDTC	CMENDTC
1	ABC	ABC-001-0001	1	1	Ibuprofen	ADVERSE EVENT	2021-01-10	2021-05-24
2	ABC	ABC-001-0001	2	2	Lexapro	ADVERSE EVENT	2021-01-10	2021-05-24
3	ABC	ABC-001-0002	1	1	Iron Suplementation	ADVERSE EVENT	2021-11-16	2021-11-24
4	ABC	ABC-001-0002	2	2	Folic Acid	ADVERSE EVENT	2021-11-16	2021-11-24
5	ABC	ABC-001-0003	1	1	Naproxen	PROPHYLAXIS	2020-02-15	2020-03-15
6	ABC	ABC-001-0003	2	2	Albuterol	ADVERSE EVENT	2020-06-14	
7	ABC	ABC-001-0003	3	3	prednisone	ADVERSE EVENT	2020-06-14	2020-08-14
8	ABC	ABC-001-0003	4	4	plavix	ADVERSE EVENT	2021-04-09	
9	ABC	ABC-001-0003	5	5	Breo inhaler	ADVERSE EVENT	2021-07-04	2021-11-08
10	ABC	ABC-001-0003	6	6	Aggrenox	ADVERSE EVENT	2021-04-10	2021-10-10
11	ABC	ABC-001-0003	7	7	Aspirin	ADVERSE EVENT	2021-04-10	
12	ABC	ABC-001-0003	8	8	Lorazepam	ADVERSE EVENT	2021-05	2021-10-15
13	ABC	ABC-001-0003	9	9	Hydrochlorothiazide	ADVERSE EVENT	2021-03-04	2021-04-19
14	ABC	ABC-001-0004	1	2	Prednisone	MEDICAL HISTORY CONDITION	2015	
15	ABC	ABC-001-0004	2	1	Metformin XR	MEDICAL HISTORY CONDITION	2020-05-10	2020-11-14
16	ABC	ABC-001-0004	3	8	Ibuprofen	ADVERSE EVENT; NON-PHARMACOLOGICAL TREATMENT	2020-10-17	
17	ABC	ABC-001-0004	4	9	Meloxicam	ADVERSE EVENT; NON-PHARMACOLOGICAL TREATMENT	2020-10-19	
18	ABC	ABC-001-0004	5	3	Vitamin B12	ADVERSE EVENT	2021-10-31	
19	ABC	ABC-001-0004	6	4	Vitamin K2	ADVERSE EVENT	2022-08-29	
20	ABC	ABC-001-0004	7	5	Vitamin D3	ADVERSE EVENT	2022-08-29	
21	ABC	ABC-001-0004	8	6	Calcium malate citrate	ADVERSE EVENT	2022-08-29	
22	ABC	ABC-001-0004	9	7	Magnesium	ADVERSE EVENT	2022-08-29	

Data Display 4: Sample SDTM Concomitant Medication (CM) Dataset

Similar to the CM data, there is a question on the CRF that ask for the reason for the therapy and this reason can indicate there is a relationship with other data:

- Medical History Condition
- Adverse Event.

The reason for the therapy is captured in PRINDC as see in Data Display 5.

Row	STUDYID	USUBJID	PRSEQ	PRSPID	PRTRT	PRINDC	PRSTDTC	PRENDTC
1	ABC	ABC-001-0002	1	1	IUD	Prophylaxis	2014	2021-02-20
2	ABC	ABC-001-0003	1	1	cardiac echocardiogram	Adverse Event	2021-03-25	2021-03-25
3	ABC	ABC-001-0003	2	2	cardiac echocardiogram	Adverse Event	2021-03-24	2021-03-24
4	ABC	ABC-001-0003	3	3	Pulmonary Function Test	Adverse Event	2021-04-28	2021-04-28
5	ABC	ABC-001-0003	4	4	Carotid ultrasound	Adverse Event	2021-04-11	2021-04-11
6	ABC	ABC-001-0003	5	5	cardiac catherization	Adverse Event	2021-04-09	2021-04-09
7	ABC	ABC-001-0003	6	6	outpatient cardiac rehab	Adverse Event	2021-07-22	
8	ABC	ABC-001-0004	1	1	Ankle-foot orthosis	Adverse Event	2020-10-17	
9	ABC	ABC-001-0004	2	2	Compression socks	Adverse Event	2020-10-17	

Data Display 5: Sample SDTM Procedures (PR) Dataset

Based on the AE CRF, we noted there is a potential for a relationship with the end of study. But looking at the End of Study CRF, there is also the potential for a relationship to other data besides the AE dataset. Looking at the main reason for discontinuation, we see that there a few options that clue us in to a possibility of a link to other data:

- Adverse Event
- Pregnancy
- Need for non-permitted concomitant medication.

DSTERM (Reported Term for the Disposition Event) captures the verbatim value as indicated on the CRF, while DSDECOD (Standardized Disposition Term) captures the value matched to CDISC controlled terminology.

Data Display 6 is a sample of the disposition dataset, which includes the informed consent as well as randomization data for all subjects. However, there are only two subjects that had a disposition event that led to discontinuation.

Row	STUDYID	USUBJID	DSSEQ	DSTERM	DSDECOD	DSCAT
1	ABC	ABC-001-0001	1	INFORMED CONSENT	INFORMED CONSENT OBTAINED	PROTOCOL MILESTONE
2	ABC	ABC-001-0001	2	RANDOMIZED	RANDOMIZED	PROTOCOL MILESTONE
3	ABC	ABC-001-0001	3	Adverse Event	LOST TO FOLLOW-UP	DISPOSITION EVENT
4	ABC	ABC-001-0002	1	INFORMED CONSENT	INFORMED CONSENT OBTAINED	PROTOCOL MILESTONE
5	ABC	ABC-001-0002	2	RANDOMIZED	RANDOMIZED	PROTOCOL MILESTONE
6	ABC	ABC-001-0002	3	Pregnancy	PREGNANCY	DISPOSITION EVENT
7	ABC	ABC-001-0003	1	INFORMED CONSENT	INFORMED CONSENT OBTAINED	PROTOCOL MILESTONE
8	ABC	ABC-001-0003	2	RANDOMIZED	RANDOMIZED	PROTOCOL MILESTONE
8	ABC	ABC-001-0003	2	COMPLETED	COMPLETED	DISPOSITION EVENT
9	ABC	ABC-001-0004	1	INFORMED CONSENT	INFORMED CONSENT OBTAINED	PROTOCOL MILESTONE
10	ABC	ABC-001-0004	2	RANDOMIZED	RANDOMIZED	PROTOCOL MILESTONE
10	ABC	ABC-001-0004	2	COMPLETED	COMPLETED	DISPOSITION EVENT

Data Display 6: Sample SDTM Disposition (DS) Dataset

In this illustration, although there is no CRF for labs, the data is coming from an electronic data transfer (eDT) and since the End of Study CRF indicated a possible relationship to labs due to a subject being able to discontinue due to pregnancy, the Laboratory Test Results (LB) dataset is needed to fully illustrate the relationship with other data.

Data Display 7 has a record placeholder for each of the subjects for all lab tests, only the lab tests pertaining to pregnancy are shown to illustrate the relationship between LB and DS domains.

Row	STUDYID	USUBJID	LBSEQ	LBSPID	LBTESTCD	LBCAT	LBORRES	LBSPEC	VISIT	LBDTC
1	ABC	ABC-001-0001	*All lab tests*							
2	ABC	ABC-001-0002	*All other lab tests*							
3	ABC	ABC-001-0002	33	UPREG1	HCG	PREGNANCY TEST	NEGATIVE	URINE	Day 1	2020-10-05
4	ABC	ABC-001-0002	34	UPREG2	HCG	PREGNANCY TEST	NEGATIVE	URINE	Week 4	2020-11-07
5	ABC	ABC-001-0002	35	UPREG3	HCG	PREGNANCY TEST	NEGATIVE	URINE	Week 8	2020-12-12
6	ABC	ABC-001-0002	36	UPREG4	HCG	PREGNANCY TEST	NEGATIVE	URINE	Week 12	2021-01-09
7	ABC	ABC-001-0002	37	UPREG5	HCG	PREGNANCY TEST	NEGATIVE	URINE	Week 16	2021-02-13
8	ABC	ABC-001-0002	38	UPREG6	HCG	PREGNANCY TEST	POSITIVE	URINE	Week 20	2021-03-06
9	ABC	ABC-001-0003	*All lab tests*							
10	ABC	ABC-001-0004	*All lab tests*							

Data Display 7: Sample SDTM Laboratory Test Results (LB) Dataset

CREATING RELREC TO SHOW THE MANY-TO-MANY RELATIONSHIP

The RELREC dataset contains the following variables: STUDYID, RDOMAIN, USUBJID, IDVAR, IDVARVAL, RELTYPE and RELID.

RDOMAIN is the domain abbreviation of the source of the related record. For example, RDOMAIN would equal “AE” if the record we’re showing a relationship for came from the Adverse Events domain.

IDVAR is used to show the name of the SDTM variable needed to uniquely identify the related record. This can be --SEQ, --GRPID, --SPID, --LNKID or any other variable that uniquely identifies the record or records of the relationship being shown.

IDVARVAL is the value of the variable mentioned in IDVAR from the domain specified in RDOMAIN. This variable would hold the actual value of the sequence number or identifier variable for the related record or records.

RELTYPE indicates the relationship type and can take the value of either “ONE” or “MANY” or null. This variable is only populated in the case where whole datasets are being related rather than individual records within a dataset.

RELID is used to group the related records together. Since the RELREC dataset is a vertical dataset and there is an entry to represent each related record or group of records, the purpose of RELID is to provide a common identifier used to show that the records identified in RELREC are related to each other. This value does not have to have a meaning, but it needs to be unique to each set of related records within a subject.

Now that we have all the SDTM datasets created, we can focus on building the dataset that will show the relationship amongst the datasets. Something to point out is that in this example, all of the datasets had the --SPID variable captured as part of the source data. This field is the Sponsor-Defined Identifier, so it does not change as more data is added. The --SEQ variable, which is the SDTM Sequence Number, can change as more data is added. This is important to distinguish because it is the --SPID that is used to link the different records across the different datasets. In addition, when looking at the CRFs in Display 1, there are fields that allow the person recording the information to enter the corresponding --SPID (or in the raw data source --NUM) that the record is related to. This information is what is used to build the relationship amongst the datasets.

Looking at one subject at a time, let’s start with the subject that had the least complicated relationship amongst the data and work our way towards the more complicated scenarios.

For subject ABC-001-0001, the relationship between AE and CM is a one-to-one. For each AE there is one concomitant medication associated with it as seen in Data Display 8.

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
1	ABC	AE	ABC-001-0001	AESPID	1		1
2	ABC	CM	ABC-001-0001	CMSPID	1		1
3	ABC	AE	ABC-001-0001	AESPID	2		2
4	ABC	CM	ABC-001-0001	CMSPID	2		2

Data Display 8: Sample RELREC for Subject ABC-001-0001

For subject ABC-001-0002, the subject discontinued due to pregnancy, and it is noted that the associated lab record is LBSPID = ‘UPREG6’. The subject was monitored throughout the pregnancy and for some time after the pregnancy and any AEs and CMs were recorded. The subject had an AE after pregnancy that led to taking of two concomitant medications. The study discontinuation is associated with the pregnancy and the AE is associated with the two medications and those relationships are demonstrated in Data Display 9 with each relationship having a unique value for RELID.

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
5	ABC	AE	ABC-001-0002	AESPID	2		1
6	ABC	CM	ABC-001-0002	CMSPID	1		1
7	ABC	CM	ABC-001-0002	CMSPID	2		1
8	ABC	DS	ABC-001-0002	DSSEQ	3		2
9	ABC	LB	ABC-001-0002	LBSPID	UPREG6		2

Data Display 9: Sample RELREC for Subject ABC-001-0002

Note that for most of the examples, it is one-to-many although it may look like it is a many-to-many scenario. For example, subject ABC-001-0003 has two AEs that are related to several medications and

several non-pharmacological treatments but there is no indication that the medications and non-pharmacological treatments are related as established by the source data shown in Data Display 10. Therefore, in this scenario the relationship between AE and CMs is one relationship, and the relationship between AE and NPs is another relationship, which give us six unique relationships for this subject's data (Data Display 11).

Row	STUDYID	SITEID	SUBJID	AENUM	AETERM	AESTDAT	AEENDAT	AEONGO	AESEV	AE_CMTRT	AE_CMID	AE_NPTRT	AE_NPID	AE_DISCONT
5	ABC	001	0003	1	Intermittent shortness of breath	2020-06-14	2021-02-05	No	Mild	Yes	2, 3	No		No
6	ABC	001	0003	2	Dyspnea	2021-03-04		Yes	Moderate	Yes	2, 5, 9	Yes	1, 2, 3, 4, 5	No
7	ABC	001	0003	3	Worsening of CAD	2021-04-09	2021-10-10	No	Moderate	Yes	4, 6, 7	Yes	3, 4, 5, 6	No
8	ABC	001	0003	4	Anxiety	2021-05-UN	2021-10-15	No	Mild	Yes	8	No		No

Row	STUDYID	SITEID	SUBJID	CMNUM	CMTRT	CMSTDAT	CMENDAT	CMONGO	CMINDC	CM_MHID	CM_AEID	CM_NPID
5	ABC	001	0003	1	Naproxen	2020-02-15	2020-03-15	No	Prophylaxis			
6	ABC	001	0003	2	Albuterol	2020-06-14		Yes	Adverse Event		1, 2	
7	ABC	001	0003	3	prednisone	2020-06-14	2020-08-14	No	Adverse Event		1	
8	ABC	001	0003	4	plavix	2021-04-09		Yes	Adverse Event		3	
9	ABC	001	0003	5	Breo inhaler	2021-07-04	2021-11-08	No	Adverse Event		2	
10	ABC	001	0003	6	Aggrenox	2021-04-10	2021-10-10	No	Adverse Event		3	
11	ABC	001	0003	7	Aspirin	2021-04-10		Yes	Adverse Event		3	
12	ABC	001	0003	8	Lorazepam	2021-05-UN	2021-10-15	No	Adverse Event		4	
13	ABC	001	0003	9	Hydrochlorothiazide	2021-03-04	2021-04-19	No	Adverse Event		2	

Row	STUDYID	SITEID	SUBJID	NPNUM	NPTRT	NPSTDAT	NPENDAT	NPONGO	NPINDC	NPOTHSPEC	NP_MHID	NP_AEID	NP_CMID
2	ABC	001	0003	1	cardiac echocardiogram	2021-03-25	2021-03-25	No	Adverse Event			2	
3	ABC	001	0003	2	cardiac echocardiogram	2021-03-24	2021-03-24	No	Adverse Event			2	
4	ABC	001	0003	3	Pulmonary Function Test	2021-04-28	2021-04-28	No	Adverse Event			2, 3	
5	ABC	001	0003	4	Carotid ultrasound	2021-04-11	2021-04-11	No	Adverse Event			2, 3	
6	ABC	001	0003	5	cardiac catherization	2021-04-09	2021-04-09	No	Adverse Event			2, 3	
7	ABC	001	0003	6	outpatient cardiac rehab	2021-07-22		Yes	Adverse Event			3	

Data Display 10: Sample Source Data for Subject ABC-001-0003

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
10	ABC	AE	ABC-001-0003	AESPID	1		1
11	ABC	CM	ABC-001-0003	CMSPID	2		1
12	ABC	CM	ABC-001-0003	CMSPID	3		1
13	ABC	AE	ABC-001-0003	AESPID	2		2
14	ABC	CM	ABC-001-0003	CMSPID	2		2
15	ABC	CM	ABC-001-0003	CMSPID	5		2
16	ABC	CM	ABC-001-0003	CMSPID	9		2
17	ABC	AE	ABC-001-0003	AESPID	2		3
18	ABC	PR	ABC-001-0003	PRSPID	1		3
19	ABC	PR	ABC-001-0003	PRSPID	2		3
20	ABC	PR	ABC-001-0003	PRSPID	3		3
21	ABC	PR	ABC-001-0003	PRSPID	4		3
22	ABC	PR	ABC-001-0003	PRSPID	5		3
23	ABC	AE	ABC-001-0003	AESPID	3		4
24	ABC	CM	ABC-001-0003	CMSPID	4		4
25	ABC	CM	ABC-001-0003	CMSPID	6		4
26	ABC	CM	ABC-001-0003	CMSPID	7		4
27	ABC	AE	ABC-001-0003	AESPID	3		5
28	ABC	PR	ABC-001-0003	PRSPID	3		5
29	ABC	PR	ABC-001-0003	PRSPID	4		5
30	ABC	PR	ABC-001-0003	PRSPID	5		5
31	ABC	PR	ABC-001-0003	PRSPID	6		5
32	ABC	AE	ABC-001-0003	AESPID	4		6
33	ABC	CM	ABC-001-0003	CMSPID	8		6

Data Display 11: Sample RELREC for Subject ABC-001-0003

While at first glance subject ABC-001-0003 may have appeared to have data that is many-to-many, looking at the data one AE was associated with many CMs or many PRs, but no relationship was established between the CMs and PRs. Thus, the relationships are one-to-many. However, for subject ABC-001-0004, there are two AEs that are related to one concomitant medication, which is a many-to-

one relationship. However, for that same subject there are two AEs that are related to the same concomitant medications and related to the same non-pharmacological treatments and there is an indication that the concomitant medications and non-pharmacological treatments are related (Data Display 12). In addition, the two AEs are related to each other. This leads us to a many-to-many relationship (i.e., many AEs related to many CMs and PRs that are related to each other) which is shown in Data Display 13.

Row	STUDYID	SITEID	SUBJID	AENUM	AETERM	AESTDAT	AEENDAT	AEONGO	AESEV	AE_CMTRT	AE_CMID	AE_NPTRT	AE_NPID	AE_DISCONT
9	ABC	001	0004	1	Fatigue	2020-08-30	2021-01-19	No	Moderate	No		No		No
10	ABC	001	0004	2	Joint swelling right leg	2020-10-17	2021-08-16	No	Severe	Yes	8, 9	Yes	1, 2	No
11	ABC	001	0004	3	Vitamin B12 deficiency	2021-10-31	2022-09-12	No	Moderate	Yes	3	No		No
12	ABC	001	0004	4	Joint swelling right leg	2021-08-17		Yes	Mild	Yes	8, 9	Yes	1, 2	No
13	ABC	001	0004	5	Low bone mineral density	2022-08-29		Yes	Mild	Yes	4, 5, 6, 7	No		No
14	ABC	001	0004	6	Vitamin B12 deficiency	2022-09-13		Yes	Mild	Yes	3	No		No

Row	STUDYID	SITEID	SUBJID	CMNUM	CMTRT	CMSTDAT	CMENDAT	CMONGO	CMINDC	CM_MHID	CM_AEID	CM_NPID
14	ABC	001	0004	1	Metformin XR	2020-05-10	2020-11-14	No	Medical History Condition			
15	ABC	001	0004	2	Prednisone	2015-UN-UN		Yes	Medical History Condition			
16	ABC	001	0004	3	Vitamin B12	2021-10-31		Yes	Adverse Event		3, 6	
17	ABC	001	0004	4	Vitamin K2	2022-08-29		Yes	Adverse Event		5	
18	ABC	001	0004	5	Vitamin D3	2022-08-29		Yes	Adverse Event		5	
19	ABC	001	0004	6	Calcium malate citrate	2022-08-29		Yes	Adverse Event		5	
20	ABC	001	0004	7	Magnesium	2022-08-29		Yes	Adverse Event		5	
21	ABC	001	0004	8	Ibuprofen	2020-10-17		Yes	Adverse Event; Non-Pharmacological Treatment		2, 4	1, 2
22	ABC	001	0004	9	Meloxicam	2020-10-19		Yes	Adverse Event; Non-Pharmacological Treatment		2, 4	1, 2

Row	STUDYID	SITEID	SUBJID	NPNUM	NPTRT	NPSTDAT	NPENDAT	NPONGO	NPINDC	NPOTHSPEC	NP_MHID	NP_AEID	NP_CMID
8	ABC	001	0004	1	Ankle-foot orthosis	2020-10-17		Yes	Adverse Event			2, 4	8, 9
9	ABC	001	0004	2	Compression socks	2020-10-17		Yes	Adverse Event			2, 4	8, 9

Data Display 12: Sample Source Data for Subject ABC-001-0004

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
34	*ABC*	*AE*	*ABC-001-0004*	*AESPID*	*2*		*1*
35	*ABC*	*AE*	*ABC-001-0004*	*AESPID*	*4*		*1*
36	*ABC*	*CM*	*ABC-001-0004*	*CMSPID*	*8*		*1*
37	*ABC*	*CM*	*ABC-001-0004*	*CMSPID*	*9*		*1*
38	*ABC*	*PR*	*ABC-001-0004*	*PRSPID*	*1*		*1*
39	*ABC*	*PR*	*ABC-001-0004*	*PRSPID*	*2*		*1*
40	*ABC*	*AE*	*ABC-001-0004*	*AESPID*	*3*		*2*
41	*ABC*	*AE*	*ABC-001-0004*	*AESPID*	*6*		*2*
42	*ABC*	*CM*	*ABC-001-0004*	*CMSPID*	*3*		*2*
43	ABC	AE	ABC-001-0004	AESPID	5		3
44	ABC	CM	ABC-001-0004	CMSPID	4		3
45	ABC	CM	ABC-001-0004	CMSPID	5		3
46	ABC	CM	ABC-001-0004	CMSPID	6		3
47	ABC	CM	ABC-001-0004	CMSPID	7		3

Data Display 13: Sample RELREC for Subject ABC-001-0004

An alternative approach for handling multiple AEs that are related to each other is to incorporate the use of AEGRPID. For AETERM = “Joint swelling right leg” (AESPID = 2 and 4), these are the same AE but there was a change in severity, thus it is captured across two records. These two AE records are tied together using AEGRPID = 2. In addition, AETERM = “Vitamin B12 deficiency” had a change in severity so it is also captured across two records, and they are linked together using AEGRPID = 3 (see Data Display 14).

Row	STUDYID	USUBJID	AESEQ	*AEGRPID*	AESPID	AETERM	AESTDTC	AEENDTC	AESEV	AECONTRT
9	ABC	ABC-001-0004	1	*1*	1	Fatigue	2020-08-30	2021-01-19	MODERATE	N
10	ABC	ABC-001-0004	2	*2*	2	Joint swelling right leg	2020-10-17	2021-08-16	SEVERE	Y
11	ABC	ABC-001-0004	3	*3*	3	Vitamin B12 deficiency	2021-10-31	2022-09-12	MODERATE	Y
12	ABC	ABC-001-0004	4	*2*	4	Joint swelling right leg	2021-08-17		MILD	Y
13	ABC	ABC-001-0004	5	*4*	5	Low bone mineral density	2022-08-29		MILD	Y
14	ABC	ABC-001-0004	6	*3*	6	Vitamin B12 deficiency	2022-09-13		MILD	Y

Data Display 14: Alternative SDTM AE for Subject ABC-001-0004

With the use of AEGRPID in the AE datasets, we can use that as an IDVAR within RELREC and have only one record to represent both AEs (see Data Display 15).

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
34	*ABC*	*AE*	*ABC-001-0004*	*AEGRPID*	*2*		*1*
35	*ABC*	*CM*	*ABC-001-0004*	*CMSPID*	*8*		*1*
36	*ABC*	*CM*	*ABC-001-0004*	*CMSPID*	*9*		*1*
37	*ABC*	*PR*	*ABC-001-0004*	*PRSPID*	*1*		*1*
38	*ABC*	*PR*	*ABC-001-0004*	*PRSPID*	*2*		*1*
39	*ABC*	*AE*	*ABC-001-0004*	*AEGRPID*	*3*		*2*
40	*ABC*	*CM*	*ABC-001-0004*	*CMSPID*	*3*		*2*
41	ABC	AE	ABC-001-0004	AESPID	5		3
42	ABC	CM	ABC-001-0004	CMSPID	4		3
43	ABC	CM	ABC-001-0004	CMSPID	5		3
44	ABC	CM	ABC-001-0004	CMSPID	6		3
45	ABC	CM	ABC-001-0004	CMSPID	7		3

Data Display 15: Alternative SDTM RELREC for Subject ABC-001-0004

CONCLUSION

To see the full story of each subject's clinical trial journey, we need to understand how the various datasets are related to each other. Whether the relationship is amongst datasets or amongst records, it needs to be conveyed in a manner that clearly shows how the data is tied together. While most people are familiar with the dataset-to-dataset relationship or the one-to-one or one-to-many relationship within RELREC, we took you on a journey where a subject had a many-to-many relationship. These relationships require special attention during the data mapping process to ensure accurate and complete representation of the data in the final dataset. However, with careful planning and attention to detail, it is possible to accurately map and represent these complex relationships with ease. We hope you have a new appreciation for meaningful relationships.

REFERENCES

CDISC Submission Data Standards and CDISC SDTM Governance Committee. (n.d.). *SDTM*. Retrieved 12 2022, from CDISC: https://www.cdisc.org/standards/foundational/sdtm

CDISC Submission Data Standards and CDISC SDTM Governance Committee. (n.d.). *SDTMIG*. Retrieved 12 2022, from CDISC: https://www.cdisc.org/standards/foundational/sdtmig

Karl Miller, J. J. (2012). Avoiding a REL-WRECK; Using RELREC Well. San Francisco: PharmaSUG. Retrieved from https://www.lexjansen.com/pharmasug/2012/DS/PharmaSUG-2012-DS08.pdf

Sreeraman, C. (2017). RELREC - SDTM Programmer's Bermuda Triangle. *PharmaSUG.* Baltimore. Retrieved from https://www.lexjansen.com/pharmasug/2017/DS/PharmaSUG-2017-DS09.pdf

Susan Fehrer, R. L. (2007). Relationships Between CDISC Variables Linking Domains (RELREC) (Part 3 of 3). NESUG. Retrieved from https://www.lexjansen.com/nesug/nesug07/po/po07.pdf

ACKNOWLEDGMENTS

The authors would like to thank Jerry Salyers for his input in and review of this paper.

CONTACT INFORMATION

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

Kaleigh Ragan
Crinetics Pharmaceuticals
kragan@crinetics.com

Richann Watson
DataRich Consulting
richann.watson@datarichconsulting.com

Any brand and product names are trademarks of their respective companies.