PharmaSUG2014 - Paper DS03

Considerations in Creating SDTM Trial Design Datasets

Jerry Salyers, Accenture Life Sciences, Berwyn, PA
Richard Lewis, Accenture Life Sciences, Berwyn, PA
Kim Minkalis, Accenture Life Sciences, Berwyn, PA
Fred Wood, Accenture Life Sciences, Berwyn, PA

ABSTRACT

Many sponsors are now submitting clinical trials data to the FDA in the format of the CDISC SDTM. The Trial Design Model (TDM) datasets can be especially challenging because, in most cases, they are being created retrospectively from the protocol, and cannot be created from electronic data. From the most recent Common Data Standards Issues document, it is clear that FDA is placing a greater emphasis on incorporating the trial design datasets into any SDTM based submission.

This paper will discuss some of the considerations and challenges in creating the TDM datasets, using case studies of both relatively simple and more complex trials. We will highlight a number of the pitfalls and misconceptions that are commonly seen when sponsors and their vendors attempt to create the TDM datasets for the first time. Included will be practical advice on which datasets should be created first, which datasets drive the creation of others, and how the trial-level and subject-level datasets relate to each other. The presentation will conclude with a list of resources for TDM dataset creation.

INTRODUCTION

Since the time of one of the author's first paper on the creation of the trial design datasets in 2011 (Wood and Lenzen), these datasets have increasingly become a part of most SDTM-based submissions. As a general rule, most of the creation of the standard "T" tables continues to be "retrospective", performed after the supporting data is already collected. This retrospective creation can lead to difficult, or even incorrect, presentations of the planned conduct of the study. We are, however, starting to see more occasions where sponsors and their vendors are developing the "T" tables earlier in the data specification process and thus are seeing the benefit of this prospective approach. Although this is definitely a step in the right direction, there continues to be a lack of basic understanding of many facets of trial design. This paper will provide an overview of the Trial Design datasets, and present a number of case studies that illustrate some of the challenges we have seen.

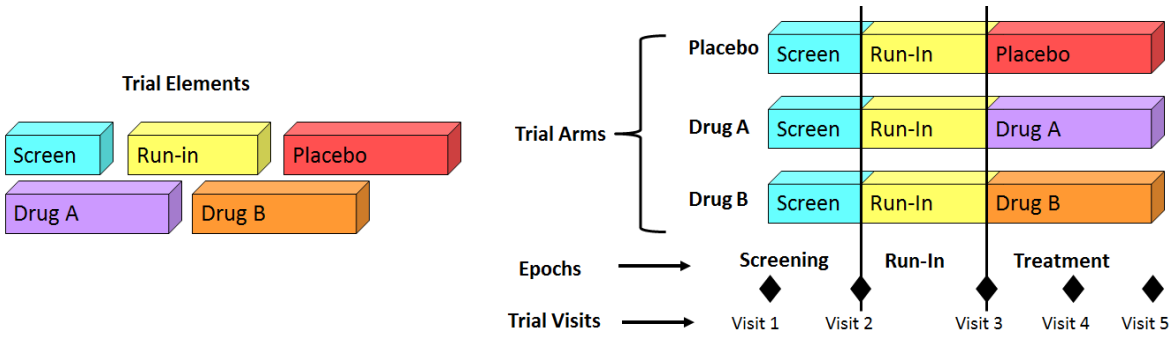
TRIAL DESIGN OVERVIEW

The Trial Design Model, described in the SDTM and the SDTM Implementation Guide (SDTMIG), provides a standardized way to represent the basic aspects of the "planned" conduct of a clinical trial. The purpose of standardizing the plan from the protocol is to allow reviewers (and others) to do the following:

- Clearly and quickly grasp the design of a clinical trial
- Understand the rules for how subjects transition through various periods or phases of a trial
- Compare the designs of different trials
- As Trial Design creates a machine-readable version of the protocol, it allows reviewers to search a data warehouse for clinical trials with similar features
- Compare planned and actual treatments for subjects in a clinical trial.

The Trial Design Model (TDM) is built on the concepts and identification of Elements, Arms, Epochs, and Visits. These are illustrated in Figure 1 below.

Figure 1. Trial Design Basic Concepts



Trial Elements represent the "building blocks" of the clinical trial. In the Trial Elements (TE) dataset, each Element is represented by an Element Code (ETCD) and an Element description (ELEMENT). The TE dataset also contains variables providing "rules" for each Element that describe how a subject transitions into (TESTRL) and out of the Element (TEENRL). There is never a time within the trial that a subject is not in one of the trial Elements, so the rules have to be created in such a way that a "gap" (or an "overlap") is not created. Put another way, while the Element "End Rule" may indicate when an element should end, the Start Rule of any Element drives the transition from the previous Element. TE also contains a "duration" variable (TEDUR) that can be used either in lieu of, or in addition to, the Element End Rule, to indicate a protocol-defined duration for the Element. For the simple trial design shown in the diagram above, the TE dataset might look like this:

STUDYID	DOMAIN	ETCD	ELEMENT	TESTRL	TEENRL	TEDUR
1999001	TE	SCREEN	Screen	Date of Informed Consent	First dose of run-in medication	
1999001	TE	RUNIN	Run-in	Frist dose of run-in medication	First dose of blinded trial medication	P1W
1999001	TE	PLAC	Placebo	First dose of blinded trial medication when randomized to placebo	Last dose of trial medication	P2W
1999001	TE	DRUGA	Drug A	First dose of blinded trial medication when randomized to Drug A	Last dose of trial medication	P2W
1999001	TE	DRUGB	Drug B	First dose of blinded trial medication when randomized to Drug B	Last dose of trial medication	P2W

The TA dataset shows the planned order of Elements (TAETORD) within each of the Arms along with identifying "decision points" where subjects may "branch" (TABRANCH) into an Element that is unique for their assigned Arm, or "rules" by which a subject may "skip" (TATRANS, not shown) the next Element(s) in their assigned sequence. Epoch is defined in TA as a vertical slice of time that is independent of Arm and provides a way to describe what is happening across Elements while a trial is blinded. "Decision points" or "branches" always occur between Epochs. The Arm (and ARMCD) values defined in TA must be the same as those used in the Demographics (DM) dataset. For one of the Arms in the trial above, the dataset may look like:

STUDYID	DOMAIN	ARMCD	ARM	TAETORD	ETCD	TABRANCH	EPOCH
1999001	TA	DRUGA	Drug A	1	SCREEN	Randomized to Drug A	SCREENING
1999001	TA	DRUGA	Drug A	2	RUNIN		RUN-IN
1999001	TA	DRUGA	Drug A	3	DRUGA		TREATMENT

The TV table represents the planned visits, or "clinical encounters" that are defined with the protocol's time and event schedule. To facilitate data collection and to satisfy constraints inherent in some data management systems, data such as AEs or concomitant medications may be collected in a type of "Log" visit. However, these "visits" should not be accounted for in TV. Just as with TE, TV contains a "start rule" (TVSTRL) that is used to indicate when, according to the protocol, a visit is scheduled to occur. Most visits are assumed to last only a single day and thus a visit's "end rule" (TVENRL) is often not populated. While ARMCD is an "Expected" variable, ARM and ARMCD are only populated if the visit schedule varies by Arm. As the SDTM IG explains, "If the timing of Visits…does not depend on which Arm a subject is in, ARMCD should be null. The TV table for our example study might look something like this:

STUDYID	DOMAIN	VISITNUM	VISIT	ARMCD	TVSTRL	TVENRL
1999001	TV	1	Visit 1		Informed Consent date; Start of Visit 1 assessments	End of Visit 1 assessments
1999001	TV	2	Visit 2		Subject dispensed one week supply of run-in medication and takes first dose	
1999001	TV	3	Visit 3		Subject dispensed 2 week supply of study medication and takes first dose	
1999001	TV	4	Visit 4		Visit 4 assessments, 1 week after starting study medication +/- 1 day	
1999001	TV	5	Visit 5		Visit 5 assessments, 2 weeks after starting study medication +/- 1 day	At Trial Exit

Additionally, there are the Trial Inclusion (TI) and Trial Summary (TS) datasets. TI represents the inclusion/exclusion criteria as detailed either in the protocol or in the CRF and is trial level data. If the study data only collects an over-arching question regarding a subject's eligibility, then TI will come from the protocol and/or any amendments that impact subject inclusion/exclusion. It is imperative that the text in IETEST within the TI domain match the IETEST in

the subject-level IE domain which records any criteria that a subject may have failed. If the protocol has amendments that impact study inclusion/exclusion, complete sets of the inclusion/exclusion criteria must be included for each version of the protocol.

The Trial Summary (TS) dataset represents key information about the trial as a whole. This dataset, if constructed first, will often help to establish the treatment specific trial elements as well as define the number of Arms for the study. The TS domain has been substantially changed with the release of the SDTM IG v3.1.3. Prior to this release, there was no real concept of "Required" trial summary parameters. If a parameter such as AGEMAX (Maximum Age of Subjects) had no protocol-defined value, the parameter was simply omitted. With the advent of the "null-flavor" variable in TS, the AGEMAX parameter has become "Required" whether or not the maximum age is defined in the protocol. In cases where the maximum age is not defined, the parameter is retained, but with the Null Flavor value of 'PINF' (positive infinity). Later in this paper, we will look more closely at the TS dataset along with several specific parameters. It should be noted that there is an extensible CDISC Controlled Terminology (CT) codelist for Trial Summary parameters as well as CT for a number of parameter values or choices.

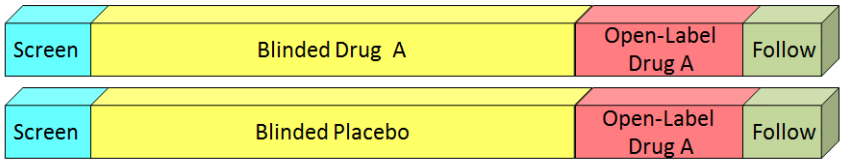
TRIAL DESIGN CASE STUDIES

The design of different types of clinical trials that may not be as routine or as simple as the trial in our illustration above. As the design of clinical trials can change dramatically depending on the "Phase" of the study or the indication, how the trial is modeled within the trial design datasets also changes dramatically. In the following examples, we will examine some challenges sponsors may face in the creation of Trial Design datasets for more-complicated studies.

Example Trial 1

This is a study involving "Drug A" which has a placebo-controlled 12-week double-blind period followed by a 4-week open-label extension where all subjects take active medication. A diagram of this trial is shown in Figure 2.

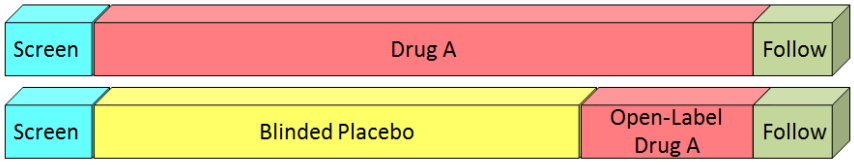
Figure 2. Double Blind into Open Label – Design Based on Protocol



The sponsor's initial attempt at developing the TA dataset is shown in the table below (for ease of display, we've left out the STUDYID column), which would have resulted in the design shown in Figure 3:

DOMAIN	ARMCD	ARM	TAETORD	ETCD	ELEMENT	TABRANCH	EPOCH
TA	A	Drug A 50 mg	1	SCRN	Screen	Randomized to Drug A	SCREENING
TA	A	Drug A 50 mg	2	DRGA50	Drug A 50 mg		TREATMENT
TA	A	Drug A 50 mg	3	FU	Follow Up		FOLLOW-UP
TA	B	Placebo	1	SCRN	Screen	Randomized to Placebo	SCREENING
TA	B	Placebo	2	PLAC	Placebo		TREATMENT
TA	B	Placebo	3	DRGA50OL	Drug A 50 mg OL		OPEN LABEL TREATMENT
TA	B	Placebo	4	FU	Follow Up		FOLLOW-UP

Figure 3. Double Blind into Open Label – Design Based on Sponsor TA Dataset



With this (mis)representation, subjects who were initially randomized to active treatment have only a single TREATMENT Epoch while subjects who were initially randomized to placebo have a TREATMENT Epoch and an

OPEN LABEL TREATMENT Epoch. However, based on the protocol, as well as the collected data, both treatment groups should have the same study Epochs, even though, for the subjects initially randomized to 12 weeks of active drug, their medication doesn't change over the course of the 16-week dosing period. Our modified TA dataset (which reflects the design in Figure 2, shows the following rows:

DOMAIN	ARMCD	ARM	TAETORD	ETCD	ELEMENT	TABRANCH	EPOCH
TA	A	Drug A 50 mg	1	SCRN	Screen	Randomized to Drug A	SCREENING
TA	A	Drug A 50 mg	2	DRGA50	Drug A 50 mg		TREATMENT
TA	A	Drug A 50 mg	3	DRGA50OL	Drug A 50 mg OL		OPEN LABEL TREATMENT
TA	A	Drug A 50 mg	4	FU	Follow Up		FOLLOW-UP
TA	B	Placebo	1	SCRN	Screen	Randomized to Placebo	SCREENING
TA	B	Placebo	2	PLAC	Placebo		TREATMENT
TA	B	Placebo	3	DRGA50OL	Drug A 50 mg OL		OPEN LABEL TREATMENT
TA	B	Placebo	4	FU	Follow Up		FOLLOW-UP

Example Trial 2

This study is a three-period crossover trial where subjects receive a single dose of Drug B at 2 mg, 6 mg, or placebo in a crossover fashion. The following is an excerpt (STUDYID, ARM, and TABRANCH left out) of the initial TA dataset, highlighting two of the sequences where P = Placebo, A = 2 mg, and B = 6 mg:

DOMAIN	ARMCD	TAETORD	ETCD	ELEMENT	EPOCH
TA	P-A-B	1	SCREEN	Screen	SCREENING
TA	P-A-B	2	S1P1	Seq 1, Period 1, Placebo	TREATMENT 1
TA	P-A-B	3	WASHOUT	Washout	WASHOUT
TA	P-A-B	4	S1P2	Seq 1, Period 2, 2 mg	TREATMENT 2
TA	P-A-B	5	WASHOUT	Washout	WASHOUT
TA	P-A-B	6	S1P3	Seq 1, Period 3, 6 mg	TREATMENT 3
TA	A-P-B	1	SCREEN	Screen	SCREENING
TA	A-P-B	2	S2P1	Seq 2, Period 1, 2 mg	TREATMENT 1
TA	A-P-B	3	WASHOUT	Washout	WASHOUT
TA	A-P-B	4	S2P2	Seq 2, Period 2, Placebo	TREATMENT 2
TA	A-P-B	5	WASHOUT	Washout	WASHOUT
TA	A-P-B	6	S2P3	Seq 2, Period 3, 6 mg	TREATMENT 3

In this table, individual Elements have been assigned based on where they occur in the dosing sequence, instead of there simply being one placebo Element, one 2 mg Element, and one 6 mg Element. The remaining Arms or sequences were similarly modeled. Thus the TE dataset contains many more rows than are necessary. Our feedback to the sponsor was to create a single Element for each possible dose, regardless of when it occurs in the subject's assigned sequence:

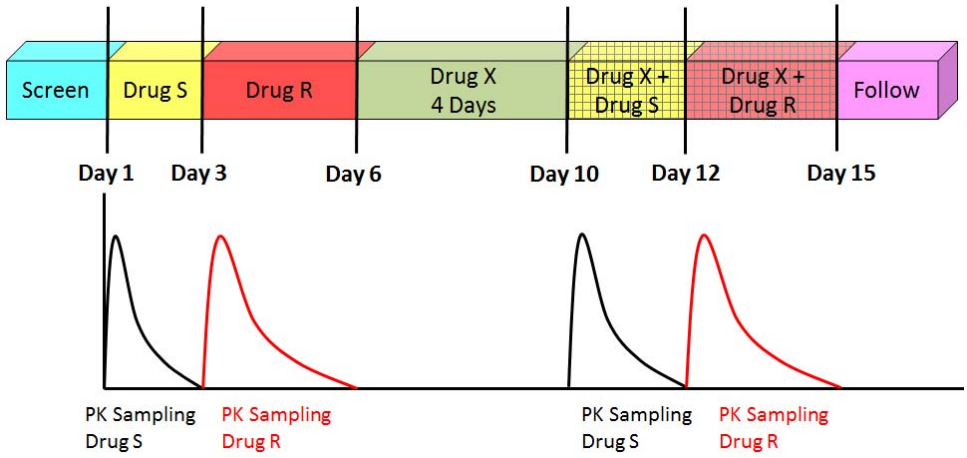
DOMAIN	ARMCD	TAETORD	ETCD	ELEMENT	EPOCH
TA	P-A-B	1	SCREEN	Screen	SCREENING
TA	P-A-B	2	PLACEBO	Placebo	TREATMENT 1
TA	P-A-B	3	WASHOUT	Washout	WASHOUT
TA	P-A-B	4	DRUGB2MG	Drug B 2 mg	TREATMENT 2
TA	P-A-B	5	WASHOUT	Washout	WASHOUT
TA	P-A-B	6	DRUGB6MG	Drug B 6 mg	TREATMENT 3
TA	A-P-B	1	SCREEN	Screen	SCREENING
TA	A-P-B	2	DRUGB2MG	Drug B 2 mg	TREATMENT 1
TA	A-P-B	3	WASHOUT	Washout	WASHOUT
TA	A-P-B	4	PLACEBO	Placebo	TREATMENT 2
TA	A-P-B	5	WASHOUT	Washout	WASHOUT
TA	A-P-B	6	DRUGB6MG	Drug B 6 mg	TREATMENT 3

Example Trial 3

Consider this more complicated "single-sequence" crossover study (See Figure 4 below for a graphic representation of the study design). In this trial, subjects receive the following:

- On Day 1, a single dose of "Drug S, followed by PK sample collections for 48 hours
- On Day 3, a single dose of "Drug R" followed by PK evaluations for 72 hours
- On Day 6, BID dosing of "Drug X' for a total of 9 days
- On Day 10, a single dose of Drug S in conjunction with the continued BID dosing of Drug X, followed by PK evaluations for 48 hours (analogous to the single dose of Drug S given at Day 1)
- On Day 12, a single dose of Drug R in conjunction with their continued BID dosing of Drug X, followed by PK evaluations for 72 hours (analogous to the single dose of Drug R given at Day 3)

Figure 4. "Single sequence" crossover design



The following represents two possible TA tables:

DOMAIN	ARMCD	TAETORD	ETCD	ELEMENT	EPOCH
TA	A	1	SCREEN	Screen	SCREENING
TA	A	2	DRUGS	Drug S, Single 10 mg Dose	TREATMENT
TA	A	3	DRUGR	Drug R, single 10 mg Dose	TREATMENT
TA	A	4	DRUGX1	Drug X, 400 mg BID, Days 6-9	TREATMENT
TA	A	5	DRUGXS	Drug S 10 mg in combination with Drug X	TREATMENT
TA	A	6	DRUGX2	Drug X, 400 mg BID, Day 11	TREATMENT
TA	A	7	DRUGXR	Drug R 10 mg in combination with Drug X	TREATMENT
TA	A	8	DRUGX3	Drug X, 400 mg BID, Days 13-14	TREATMENT
TA	A	9	FOLLOWUP	Follow-Up	FOLLOW-UP

OR

DOMAIN	ARMCD	TAETORD	ETCD	ELEMENT	EPOCH
TA	A	1	SCREEN	Screen	SCREENING
TA	A	2	DRUGS	Drug S, Single 10 mg Dose	TREATMENT
TA	A	3	DRUGR	Drug R, single 10 mg Dose	TREATMENT
TA	A	4	DRUGX	Drug X, 400 mg BID, Days 6-9	TREATMENT
TA	A	5	DRUGXS	Drug S 10 mg in combination with Drug X	TREATMENT
TA	A	6	DRUGXR	Drug R 10 mg in combination with Drug X	TREATMENT
TA	A	7	FOLLOWUP	Follow-Up	FOLLOW-UP

Note the two extra Elements in the first version of the table where the continued BID dosing on Day 11 and Days 13 and 14 are broken out into separate Elements.

We consider the second version of TA to be the more correct version for the above study. The former table was the design originally favored by the sponsor. We provided the latter TA table back to them along with the TE table below.

DOMAIN	ETCD	ELEMENT	TESTRL	TEENRL
TE	SCRN	Screen	Date/time of Informed Consent	Date/time of first single dose of Drug S
TE	DRUGS	Drug S, Single 10 mg Dose	Date/time of first single dose of Drug S	Date/time of first single dose of Drug R
TE	DRUGR	Drug R, single 10 mg Dose	Date/time of first single dose of Drug R	Date/time of first BID dose of Drug X

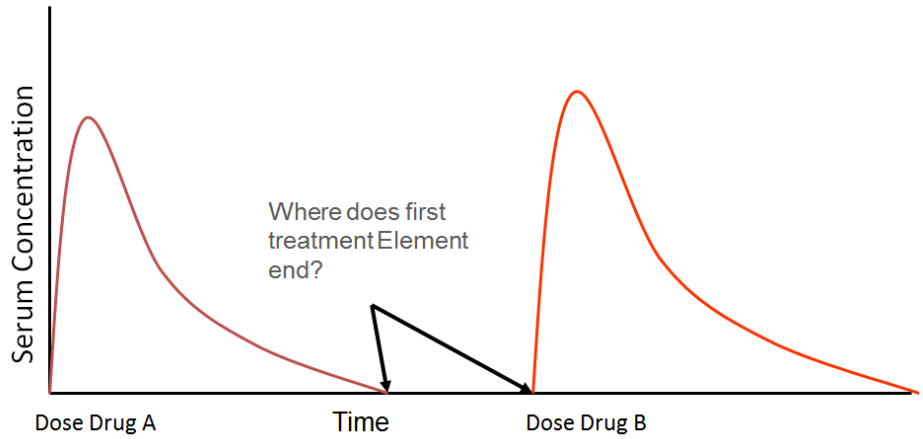
TE	DRUGX	Drug X, 400 mg BID, Days 6-9	Date/time of first BID dose of Drug X	Date/time of second single dose of Drug S while subject is receiving Drug X
TE	DRUGXS	Drug S 10 mg in combination with Drug X	Date/time of second single dose of Drug S while subject is receiving Drug X	Date/time of second single dose of Drug R while subject is receiving Drug X
TE	DRUGXR	Drug R 10 mg in combination with Drug X	Date/time of second single dose of Drug R while subject is receiving Drug X	Final PK sample collected at Day 15
TE	FOLLOWUP	Follow-Up	Discharge from unit after final PK sample collection	At Trial Exit

As we see, the sponsor had originally wanted to split out into separate Elements the continued BID dosing of Drug X following the second single dose of Drug S and Drug R, respectively. Our explanation was basically that trial design is less concerned with these specific "protocol-defined" days than it is with defining a "treatment strategy". We also advised them that we felt it was very important that the Element "start rules" (and by default, the "end rules") be consistent between the initial single dose of either Drug R or Drug S and the second single dose of the two drugs. The evaluation period for each should remain the same. When subjects continue to receive study drug on Study Day 11 and on Study Day 13-14, these should not represent additional Elements in the study. Subjects are continuing to take the study drug during this time simply to remain at steady state while the pharmacokinetic profiles of Drug S and Drug R are being evaluated in the presence of Drug X.

COMMON ISSUES

As our previous trial design paper (Wood and Lenzen, 2011) pointed out, often constructing the T tables is more a matter of art than it is pure science. To be sure, as shown in the above examples, there are instances where some representations are "more correct". There are also occasions, however where there may be more than one "correct" design. Consider information in Figure 4: does the Element reflecting the dose of Drug A last all the way to the dose of Drug B? Or is there an Element in between? The decision will be based upon how one wants to analyze the in-between period. Is this a continuation of the Drug A dosing period that will contain evaluations related to Drug A, or is it considered to be a washout period with no assessments related to Drug A? Often there is no single correct answer that's appropriate for all trials. Again, it comes down to the strategy and the analysis.

Figure 5. Determining the Best Representation of Trial Elements



Let's take a moment and look at each trial design domain and highlight just a few of the more important issues and problems that we often see when reviewing some sponsors attempts at trial design.

TA:

- Branching statements should be included on the line of the Element where, at its conclusion, it leads into the Element that makes the Arm unique.
- Transition "rules" should be present when the protocol indicates that subjects may qualify for a "shortened path" within their assigned Arm. As with branching, statements should be included on the line of the Element where, at its conclusion, leads to the transition.
- All values for ETCD and ELEMENT should be present in TE.
- The values of ARMCD and ARM should be those used in DM.

TE

- The Element Start Rules and End Rules should point to events for which there is collected data. An example of an issue is when the protocol states that an Element ends when the subject leaves the clinic, but the CRF did not collect the date this occurred. As a result, one may have to define a surrogate Element End Rule that's different from what the sponsor intended in the protocol. The earlier trial design is created, the more likely that the required data points will be part of the data collection.
- A Start Rule should be specific enough to ensure there is no confusion as in:
    - If a subject participates in two Elements of the same drug with different dose levels, the dose level needs to be included in the Start Rule.
    - When there are actually multiple Arms with one or more unique treatments, an Element of "Treatment" alone is not sufficient.
- TEDUR should not be populated for Elements of variable lengths. An example of such a misuse would be for a Screening Element with TEDUR = 'P28D' where the Screening period is defined in the protocol as simply "within 28 days of randomization".

TV

- An End Rule should be present for visits that are likely to last more than one day. An example would be for an oncology trial where a visit equates to a cycle of treatment.
- Unscheduled visits should not be included in TV, since TV is only for planned visits. Unscheduled visits would only be present in Subject Visits.
- When study visits vary, even slightly, between Arms of the study, a full set of the Visit structure should be supplied for each Arm with ARMCD populated in each instance.

TI

- Versions of TI should be included when an amendment has affected any of the inclusion/exclusion criteria.
- There should be a 1:1 relationship between IETESTCD and IETEST. Sometimes sponsors fail to update the IETEST after updating IETESTCD as the result of an amendment.

TS

The TS domain underwent some significant changes in SDTM-IG v3.1.3 compared to previous versions. To illustrate these changes as well as provide an example of a TS dataset, we'll use the following study description:

> *This open label Phase II study will randomize approximately 138 subjects (18 years of age and older) with metastatic bladder or urethra cancer into 3 treatment ARMs: Docetaxel (75 mg/m2), Docetaxel + Drug A (10 mg/kg), and Docetaxel + Drug B (12 mg/kg).*

Obviously, there's much more to the study than this, such as specifying the primary and secondary study objectives, where the study is to be conducted, etc., but just this short description is enough to start with.

STUDYID	DOMAIN	TSSEQ	TSGRPID	TSPARMCD	TSPARM	TSVAL	TSVALNF	TSVALCD	TSVCDREF
2014015	TS	1		ACTSUB	Actual Number of Subjects	130			
2014015	TS	1		ADAPT	Adaptive Design	N		C49487	CDISC
2014015	TS	1		ADDON	Added onto Existing Treatments	Y		C49488	CDISC
2014015	TS	1		AGEMIN	Planned Minimum Age	P18Y			ISO 8601
2014015	TS	1		AGEMAX	Planned Maximum Age		PINF		ISO 21090
2014015	TS	1	DOC	CURTRT	Current Therapy or Treatment	Docetaxel		15H5577CQD	UNII
2014015	TS	1	DOC	DOSE	Dose per Administration	75			
2014015	TS	2	DRUGA	DOSE	Dose per Administration	10			
2014015	TS	3	DRUGB	DOSE	Dose per Administration	12			
2014015	TS	1	DOC	DOSU	Dose Units	mg/m2		C67402	CDISC
2014015	TS	2	DRUGA	DOSU	Dose Units	mg/kg		C67401	CDISC
2014015	TS	3	DRUGB	DOSU	Dose Units	mg/kg		C67401	CDISC
2014015	TS	1		INDIC	Trial Indication	Bladder Cancer		406024002	SNOMED CT
2014015	TS	1		NARMS	Planned	3			

					Number of Arms				
2014015	TS	1		PLANSUB	Planned Number of Subjects	138			
2014015	TS	1		RANDQT	Randomization Quotient	0.67			
2014015	TS	1		TBLIND	Trial Blinding Schema	OPEN LABEL		C49659	CDISC
2014015	TS	1		TPHASE	Trial Phase Classification	Phase II Trial		C15601	CDISC
2014015	TS	1	DRUGA	TRT	Investigational Therapy or Treatment	Drug A			
2014015	TS	1	DRUGB	TRT	Investigational Therapy or Treatment	Drug B			

Just a few points regarding the above table and TS in general:

- Even if it's a SDTM IG 3.1.2 TS dataset (and thus no "Required" parameters), as many parameters as possible should be used to adequately describe the study.
- When employing both ACTSUB and PLANSUB parameters, the numbers for both should refer to the same degree of subject participation (i.e., both indicating number of "randomized" subjects (planned and actual)
- TSSEQ – is a sequence number within a parameter allowing multiple records with the same TSPARMCD, such as DOSE in the table above
- TSGRPID – can be used to tie together any block of records in the dataset. In this case, the DOSE and DOSU records were tied to the treatments administered.
- TSPARMCD and TSPARM come from CDISC Controlled Terminology (CT) as do many of the responses in the TSVAL column. For those responses, the case shown in TSVAL matches the case in CT.
- In this three Arm study, the drug Docetaxel acts as a CURTRT (Current Therapy or Treatment) for both investigational products, Drug A and Drug B.
- The parameter INDIC (Trial Indication) is linked to the NCI Term Browser located at http://ncit.nci.nih.gov/ncitbrowser. This parameter represents the overall disease indication for the study medication. For Phase I studies that primarily collect PK data on a "healthy subject" population, this parameter should be omitted.
- The value shown for RANDQT represents the percentage (the decimal-based fraction of 1) of subjects that are planned to be exposed to investigational treatment. If all subjects are exposed at one time or another during the course of the trial, this would be 1.
- For Docetaxel, the value in TSVALCD comes from the FDA Substance Registration System's "Unique Ingredient Identifier" (UNII) available at http://fdasis.nlm.nih.gov/srs.
- For those values that come from CDISC CT, NCI EVS "Concept Code" appears in the TSVALCD. This comes from Column A of the controlled terminology spreadsheet available via the NCI website at http://www.cancer.gov/cancertopics/cancerlibrary/terminologyresources/cdisc.

CONCLUSIONS

Creating SDTM trial design datasets can often be challenging, depending on the complexity of the study. Even after gaining experience across different therapeutic areas, two people still may not model a study in exactly the same way, but both may be defensible and, for the most, correct. Our general philosophy is to try to keep the datasets as simple as possible, but still within the framework of the model and the intent of the study protocol. It takes a lot of practice and experience with trials of many types of designs before one can become adept at creating Trial Design datasets. We expect the relative importance of these datasets in an SDTM-based submission to be increasing as more and more sponsors submit them, and reviewers get more experience in seeing how valuable they are.

RESOURCES

- Wood, F. and Lenzen, M. (2011) Trials and Tribulations of Trial Design. PharmaSUG Proceedings, May 2011.
- Study Data Tabulation Model. Clinical Data Interchange Standards Consortium (CDISC) Submission Data Standards (SDS) Team, Version 1.4, December 2013.
- Study Data Tabulation Model Implementation Guide: Human Clinical Trials. Clinical Data Interchange Standards Consortium (CDISC) Submission Data Standards (SDS) Team, Version 3.2, December 2013
- There are also several case studies and Trial Design examples posted in the "Member's Only" area on the CDISC website (http://www.cdisc.org).

ACKNOWLEDGMENTS

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

CONTACT INFORMATION

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

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

Richard Lewis
Senior Consultant, Data Standards Consulting
Accenture Life Sciences
1160 W. Swedesford Road Bldg One
Berwyn, PA 19312
262.716.4167
richard.m.lewis@accenture.com

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

Kim Minkalis
Consultant, Data Standards Consulting
Accenture Life Sciences
1160 W. Swedesford Road Bldg One
Berwyn, PA 19312
484.467.4988
kimberly.a.minkalis@accenture.com