

CDISC ADaM Phases, Periods, and Subperiods: A Case Study

Jennifer Fulton, Westat

ABSTRACT

Many typical clinical studies, especially Phase 1 or 2, are comprised simply of Screening, Treatment, and Follow-up time periods. Occasionally the Treatment portion of the study needs to be divided further into analysis periods, and even sub-levels within one or more of the analysis periods. Two examples are crossover studies, and studies that include a break or "wash out" period between doses of study medication. If it is important, for example, to know what the study circumstances were at the time when a particular adverse event started, then the CDISC ADaM permissible Phase, Period, and Subperiod variables should be utilized. This paper will present a case study on an actual clinical trial that presented interesting challenges for the correct implementation of the ADaM guidelines for these variables. In this study the route of administration was being investigated. Rather than comparing study drugs, the route changed for the same subject from one set of visits to the next. Topics will include: key ADaM variables that come into play, ADSL variables and how they relate to other ADaM Basic Data Structure (BDS) domain variables, and SAS code to conquer some of the difficult or unusual obstacles. The paper assumes a basic knowledge of CDISC ADaM concepts.

INTRODUCTION

The ever growing and evolving CDISC Analysis Data Model, or ADaM, can pose new and interesting (or frustrating!) challenges for SAS programmers. This is especially true when a clinical trial does not fit the straightforward structure of simple Screening, Treatment, and Follow-up time periods. If the time periods are further subdivided, the number of variables in the ADSL domain can grow very quickly, for example, and then the task of linking the Basic Data Structure (BDS) and Occurrence Data Structure (OCCDS) data sets to ADSL becomes daunting. The ADaM Implementation Guide (IG) offers quite an assortment of timing variables that the programmers may rarely use.

This paper will describe the approaches and obstacles to programming ADaM data for a case study of a clinical trial in which the subjects were treated with the same study drug but different routes of administration during different time points in the trial, in an effort to determine the safest and most effective way of delivering the drug. Examples of routes of administration (not necessarily used in this trial) include "oral", "intravenous", and "subcutaneous injections". This study structure required that ADaM variables for Phases, Periods, and Subperiods all were utilized. The author will also share code from SAS macros that helped to derive the numerous variables in an efficient manner. This paper assumes the reader has a basic to intermediate level of experience with CDISC concepts including the Study Data Tabulation Model (SDTM), ADaM, and Controlled Terminology.

STUDY DESIGN OVERVIEW

The clinical trial presented as the case study for this paper is part of an ongoing program, and therefore the specifics will need to remain confidential. Generalized terms such as "Route 1", "Route 2" will be used to represent the actual routes of administration, the study days have been rounded, and the drug name and indication will not be revealed. These details are not critical to the explanation of the particular challenges and lessons learned from this study as it relates to ADaM.

Each subject received the same study drug throughout the trial. During study days 1-10, the subject received the drug via Route 1. During study days 11-20, the subject received the drug via three different but related routes which will be referred to as Route 2.1, Route 2.2, and Route 2.3. Route 3 was used for days 21-30, and days 31-40 were the follow up time period. The study was for a commercially available drug and subjects received the approved route of administration during screening, and returned to the approved route of administration during the follow-up period. But to simplify the examples in the paper, information about treatment during the screening and follow-up period will not be included.

Below in Figure 1 is a study schema which is helpful to provide a visual of the study design. The Biostatistician inserted a similar schema into the study Statistical Analysis Plan (SAP), and it was a helpful guidance for programming the ADaM domains. The schema also illustrates how to assign the ADaM Phases, Analysis Periods, and Analysis Subperiods for this trial. Per the ADaMIG v1.2, these are the key *permissible* BDS timing variables:

Analysis Phase is represented by the character variable APHASE, and is a categorization of timing within a study. It can be a higher-level categorization of the analysis periods, or an analysis epoch. Analysis Phase is independent of the treatment variables within ADSL, and may be populated for spans of time where a subject is not on treatment. APHASEN is the numeric variable that can accompany APHASE to provide a numeric representation for sorting. (Associated ADSL variables: APHASEw)

Analysis Period is represented by the numeric variable APERIOD. It is a record-level timing variable that represents the analysis period within the study associated with the record for analysis purposes. APERIODC is the character variable that can accompany APERIOD to provide a text description. (Associated ADSL variables: APxxSDT, APxxEDT)

Analysis Subperiod within Period is represented by the numeric variable ASPER. This is a numeric value characterizing a sublevel within APERIOD. The value of ASPER resets to 1 when the APERIOD value changes. ASPERC is the character variable that can accompany ASPER to provide a text description. (Associated ADSL variables: PxxSw)

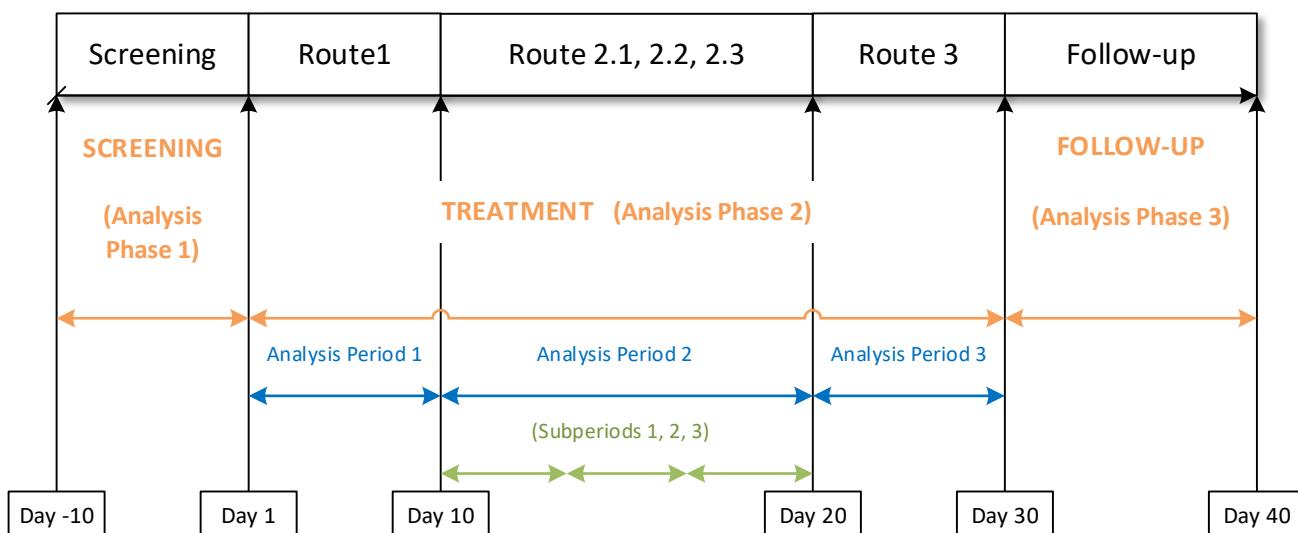


Figure 1. Study Schema

TREATMENT AND TIMING VARIABLES

There is an important relationship between the PERIOD and TREATMENT variables in the ADaM ADSL data. Planned treatment for Period xx is represented as TRTxP, and actual treatment for Period xx is represented as TRTxA. The "xx" must match up with the period number to give the treatment applied during the associated Analysis Period. Treatment start and end dates, Period start and end dates, and Subperiod start and end dates also must include the "xx" in the variable name to tie all of them together and make clear to the reviewer that all are associated with the same Analysis Period of the study. Display 1 illustrates some additional information that was presented in the case study SAP to guide programmers as to how to relate the study design to the treatment and timing variables. Note also that there may be a gap between the end of treatment in a period, and the end of the period itself. Also each period or phase ends one day before the start of the next phase or period to avoid overlap.

Analysis Phase 1 (APHASE="SCREENING", APASEN=1)

PH1STD = date of first screening visit

PH1EDT = date of Route 1 administration -1

Analysis Phase 2 (APHASE="TREATMENT", APASEN=2)

PH2STD = date of first Route 1 administration

PH2EDT = date of first day of follow-up period -1

Analysis Period1 (APERIOD=1, TRT01P="ROUTE 1")

AP01SDT = TR01SDT = date of first Route 1 administration

AP01EDT = date of first dose in the Route 2 period -1

TR01EDT = date of last dose in the Route 1 period

Analysis Period2 (APERIOD=2, TRT02P="ROUTE 2")

AP02SDT = TR02SDT = date of first dose in the Route 2 period

AP02EDT = date of first dose in the Route 3 period -1

TR02EDT = date of last dose in the Route 2 period

Subperiod 1 (ASPER=1, P02S1="ROUTE 2.1")

P02S1SDT = date of first Route 2.1 dose

P02S1EDT = date of first Route 2.2 dose -1

Subperiod 2 (ASPER=2, P02S2="ROUTE 2.2")

P02S2SDT = date of first Route 2.2 dose

P02S2EDT = date of first Route 2.3 dose -1

Subperiod 3 (ASPER=3, P02S3="ROUTE 2.3")

P02S3SDT = date of first Route 2.3 dose

P02S3EDT = date of first Route 3 dose -1

Analysis Period3 (APERIOD=3, TRT03P="ROUTE 3")

AP03SDT = TR03SDT = date of first dose in the Route 3 period

AP03EDT = date of first day of follow-up period -1

TR03EDT = date of last dose of Route 3 period

Analysis Phase 3 (APHASE="FOLLOWUP", APASEN=3)

PH3STD = date of first day of follow-up period/phase

PH3EDT = date of last visit in follow-up period/phase

Display 1. Variable Descriptions in the SAP to Guide Programmers

EXAMPLE DATA

When the study is broken into Phases, Periods, and Subperiods, there are two distinct steps:

1. Derive the timing variables in ADSL which define when each Phase/Period/Subperiod starts and ends.
2. Apply the timing variables from ADSL to the visit or event dates in the BDS and OCCDS data sets to assign each record to the appropriate Phase, Period, and Subperiod.

When these two steps are completed correctly, the data can be analyzed by Phase, Period, and Subperiod as desired by the Biostatisticians and Sponsors.

Tables 1a through e contain some key ADSL variables from the case study. Clearly, the number of variables to keep track of in ADSL grows when Phases, Periods, and Subperiods are required. There are many more variables that could be necessary to include in ADSL, such as dose unit and treatment duration for each period.

ADSL

ROW	USUBJID	TRT01P	TRT01PN	TRT02P	TRT02PN	TRT03P	TRT03PN	DOSE01P	DOSE02P	DOSE03P
1	xxx-yy-201	ROUTE 1	1	ROUTE 2	2	ROUTE 3	3	XXX	XXX	XXX
2	xxx-yy-204	ROUTE 1	1	ROUTE 2	2	ROUTE 3	3	XXX	XXX	XXX
3	xxx-yy-301	SCREEN FAIL	99	SCREEN FAIL	99	SCREEN FAIL	99			

Table 1a. Planned Treatment and Dose for Each Period

ROW	TR01SDTM	TR01EDTM	TR02SDTM	TR02EDTM	TR03SDTM	TR03EDTM
1	13MAY19:11:25:00	20MAY19:11:35:00	22MAY19:11:20:00	30MAY19:10:21:00	01JUN19:10:50:00	03AUG19:10:36:00
2	09JUL19:10:45:00	17JUL19:11:05:00	19JUL19:10:15:00	26JUL19:11:01:00	29JUL19:11:00:00	07AUG19:10:24:00
3						

Table 1b. Treatment Start and End Date/Time for Each Period

ROW	APHASE1	APHASE2	APHASE3	PH1SDT	PH1EDT	PH2SDT	PH2EDT	PH3SDT	PH3EDT
1	SCREENING	TREATMENT	FOLLOWUP	29APR2019	12MAY2019	13MAY2019	02JUL2019	03JUL2019	13JUL2019
2	SCREENING	TREATMENT	FOLLOWUP	25JUN2019	08JUL2019	09JUL2019	08AUG2019	09AUG2019	19AUG2019
3	SCREENING			05MAR2020	05MAR2020				

Table 1c. Phases with Start and End Dates

ROW	AP01SDT	AP01EDT	AP02SDT	AP02EDT	AP03SDT	AP03EDT
1	13MAY2019	21MAY2019	22MAY2019	31MAY2019	01JUN2019	02JUL2019
2	09JUL2019	18JUL2019	19JUL2019	28JUL2019	29JUL2019	08AUG2019
3						

Table 1d. Period Start and End Dates

ROW	P02S1SDT	P02S1EDT	P02S2SDT	P02S2EDT	P02S3SDT	P02S3EDT
1	22MAY2019	23MAY2019	24MAY2019	25MAY2019	26MAY2019	31MAY2019
2	19JUL2019	20JUL2019	21JUL2019	22JUL2019	23JUL2019	28JUL2019
3						

Table 1e. Subperiod Start and End Dates

Table 2 presents an example of some of the key variables in the associated ADaM Vital Signs (ADVS) data set.

ADVS

USUBID	PARAM	AVISIT	ADTM	TRTP	APHASE	APERIOD	APERIODC	ASPER	ASPERC
xxx-yy-201	Systolic Blood Pressure (mmHg)	SCREENING	29APR19:00:00:00	PRE-TREAT	SCREENING				
xxx-yy-201	Systolic Blood Pressure (mmHg)	DAY 1	13MAY19:11:41:00	ROUTE 1	TREATMENT	1	ROUTE 1		
xxx-yy-201	Systolic Blood Pressure (mmHg)	DAY 10	22MAY19:11:31:00	ROUTE 2.1	TREATMENT	2	ROUTE 2	1	ROUTE 2.1
xxx-yy-201	Systolic Blood Pressure (mmHg)	DAY 12	24MAY19:11:51:00	ROUTE 2.2	TREATMENT	2	ROUTE 2	2	ROUTE 2.2
xxx-yy-201	Systolic Blood Pressure (mmHg)	DAY 14	26MAY19:10:23:00	ROUTE 2.3	TREATMENT	2	ROUTE 2	3	ROUTE 2.3
xxx-yy-201	Systolic Blood Pressure (mmHg)	DAY 20	01JUN19:10:52:00	ROUTE 3	TREATMENT	3	ROUTE 3		
xxx-yy-204	Systolic Blood Pressure (mmHg)	SCREENING	25JUN19:00:00:00	PRE-TREAT	SCREENING				
xxx-yy-204	Systolic Blood Pressure (mmHg)	DAY 1	09JUL19:11:01:00	ROUTE 1	TREATMENT	1	ROUTE 1		
xxx-yy-204	Systolic Blood Pressure (mmHg)	DAY 10	19JUL19:10:26:00	ROUTE 2.1	TREATMENT	2	ROUTE 2	1	ROUTE 2.1
xxx-yy-204	Systolic Blood Pressure (mmHg)	DAY 12	21JUL19:10:58:00	ROUTE 2.2	TREATMENT	2	ROUTE 2	2	ROUTE 2.2
xxx-yy-204	Systolic Blood Pressure (mmHg)	DAY 14	23JUL19:11:02:00	ROUTE 2.3	TREATMENT	2	ROUTE 2	3	ROUTE 2.3
xxx-yy-204	Systolic Blood Pressure (mmHg)	DAY 20	29JUL19:11:03:00	ROUTE 3	TREATMENT	3	ROUTE 3		
xxx-yy-301	Systolic Blood Pressure (mmHg)	SCREENING	05MAR20:11:03:00	PRE-TREAT	SCREENING	1	SCREEN FAIL	1	

Table 2. Vital Signs

SOME ADDITIONAL OBSERVATIONS

The Analysis Periods do not include the Screening Phase of the study, and therefore the variables relating to Period and Subperiod are blank for the Screening records.

ASPER and ASPERC are blank for Periods 1 and 3, as they are not applicable.

The start and end dates for Phases, Periods, and Subperiods cannot be based on visit names/numbers (in case these vary from subject to subject). They have to be based on visit dates relative to planned events (ex. Period 2 is not assigned to start on visit Day 10, but rather on the first dose date using the Route 2 administration method).

The Screen Failure subject had data only during the Screening Phase of the study so values for variables related to subsequent parts of the study are all blank.

The value of TRTP in ADVS comes from the corresponding value of TRTxP that matches with the Period of the visit when the vital sign was recorded. The exception is the records where TRTP = "ROUTE 2.1", "ROUTE 2.2", and "ROUTE 2.3". The ADaMIG v1.2 does not include TRTx variables for Subperiods. TRTP can also be blank for the Screening records, but the author chose to derive it as "PRE-TREAT" for clarity in the data listings.

In this study, visit time was not collected for all assessments. This lead to challenges in using time as part of assigning Phase, Period, and Subperiod start and end dates. Therefore only visit dates were used, and the end of each period was defined as 1 day prior to the start of the next period.

SAS MACROS FOR DERIVING PHASE/PERIOD/SUBPERIOD VARIABLES

Below is some sample SAS macro code for one approach to derive Period start and end dates in ADSL, calling the macro once for each of the 3 Periods. Similar code can be used to loop over the Phases and Subperiods:

```
*--the MAKEDATE macro converts the input variable into the new name;
*--and format as needed for Period or Treatment start and end dates;
%macro makedate(_newvar=, _datevar=);
  &_newvar.dt = mdy(substr(&_datevar,6,2),substr(&_datevar,9,2),substr(&_datevar,1,4));
  &_newvar.tm = input(scan(&_datevar,2,'T'),time5.);
  &_newvar.dtm = dhms(&_newvar.dt,0,0,&_newvar.tm);
  format &_newvar.dt mmddyy8. &_newvar.dtm datetime20.;
%mend datetime;
```

```

--The periodloop macro assigns start and stop of each treatment and period,
based on the route of administration in SDTM.EX.;
%macro periodloop(per=, _route=);

proc sort data=ex out=ex&per;
  where exroute = "&_route";
  by usubjid exstdtc;
run;

--Select the first dose within the time periods for the selected route of;
--administration for each subject. This is the treatment start date and;
--period start date. ;
data per&per._start(keep=usubjid tr0&per.sdtm ap0&per.sdt);
  set ex&per;
  by usubjid;
  if first.usubjid;

  %makedate(_newvar=TR0&per.S, _datevar=exstdtc)
  %makedate(_newvar=AP0&per.S, _datevar=exstdtc)
run;

--Select the last dose within the time periods for the selected route of;
--administration for each subject. This is the treatment end date. ;
--Period end date is derived later outside of the macro;
data per&per._endtrt(keep=usubjid tr0&per.edtm);
  set ex&per;
  by usubjid exstdtc;
  if last.usubjid;

  %makedate(_newvar=TR0&per.E, _datevar=exendtc)
run;

--Merge the period/treatment start and end date information;
data per&per.(keep=usubjid tr0&per.sdtm tr0&per.edtm ap0&per.sdt);
  merge per&per._start &per._endtrt;
  by usubjid;
run;

%mend periodloop;

%periodloop(per=1, _route=%str(ROUTE 1));
%periodloop(per=2, _route=%str(ROUTE 2));
%periodloop(per=3, _route=%str(ROUTE 3));

```

A separate section of code at the end of the program derives Period end dates as one day prior to the subsequent Period start date, and calculates treatment duration for each treatment period. The end date of the last Period is the subject's end of study date. Or if a subject discontinues in the middle of Period 1 or Period 2, then the Period 1 or 2 end date is the subject's end of study date, and the start and end date for Period 3 will be missing.

Here is some sample SAS code for assigning visit dates within a BDS data program to the appropriate Phase/Period/Subperiod. This macro code is called within each applicable ADaM data program:

```
/*phases*/
if (.z < ph1sdt<=&_visitdt.) and (&_visitdt<=ph2sdt or ph1edt=.) then
aphase = aphase1;

else if (.z < ph2sdt<=&_visitdt) and (&_visitdt<=ph3sdt or ph2edt=.) then
aphase = aphase2;

else if (.z < ph3sdt<=&_visitdt) and (&_visitdt<=ph3edt or ph3edt=.) then
aphase = aphase3;

/*periods*/
if (.z < &_visitdt <= ap01sdt) then do;
aperiod = .;
aperiodc = ' ';
end;
else if (&_visitdt > ap01sdt) and (&_visitdt<= ap02sdt or ap01edt=.) then do;
aperiod = 1;
aperiodc = trt01p;
end;
else if (&_visitdt > ap02sdt) and (&_visitdt<= ap03sdt or ap02edt=.) then do;
aperiod = 2;
aperiodc = trt02p;
end;
else if (&_visitdt > ap03sdt) and (&_visitdt<= ap04sdt or ap03edt=.) then do;
aperiod = 3;
aperiodc = trt03p;
end;

/*subperiods*/
if (&_visitdt >= p02s1sdt) and (&_visitdt<= p02s2sdt or p02s1edt=.) then do;
asper = 1;
asperc = p02s1;
end;
else if (&_visitdt >= p02s2sdt) and (&_visitdt<= p02s3sdt or p02s2edt=.) then do;
asper = 2;
asperc = p02s2;
end;
else if ((&_visitdt >= p02s3sdt) and (&_visitdt<= p02s3edt or p02s3edt=..)) or
(&_visitdt. = ap03sdt) then do;
asper = 3;
asperc = p02s3;
end;
```

PINNACLE 21 WARNINGS

The SAS programmers encountered the following warnings when the data for this study was first analyzed by Pinnacle 21 software, and these provide useful tips for readers to keep in mind while working with this type of multiple-period data:

Inconsistent value for APERIODC/APERIOD: Within a study, All APERIODC values must be the same for each unique value of APERIOD when primary variable APERIOD is populated.

In this case study there were actually two different groups of subjects with different routes of administration explored. So at first there were two different values of APERIODC both corresponding to the same value of APERIOD=2. The solution was to name Period 2 with a more generalized label such as "Experimental", rather than to use the specific route being explored. Since TRTA/TRTAN and TRTP/TRTPN were assigned based on the values of APERIODC/APERIOD, Pinnacle21 reported a similar warning for these variables.

ADSL.TRTxxP is present but TRxxSDT/TRxxEDT is not present for periods greater than 01: If Planned Treatment for Period xx (TRTxP) is present, First/Last Treatment date for period xx (TRxxSDT/TRxxEDT) must be present if there is more than one treatment period. When there is only 1 treatment period 01, TRTSDT/TRTEDT is sufficient.

The ADSL for the case study already included TRxxSDTM and TRxxEDTM so TRxxSDT and TRxxEDT were easily added, but this is a good reminder of the difference in the treatment timing variables that are required when the study includes multiple periods.

BDS.APERIOD xx does not have a corresponding ADSL.TRxxSDT/TRxxEDT variable: The analysis period (APERIOD) will be a number like 01. There must be a corresponding first treatment date variable for this period (TR01SDT).

This warning is related to the warning above – the data included datetime variables TRxxSDTM and TRxxEDTM, but Pinnacle 21 was looking for the date version of these variables. However this points out to the reader the importance of linking the various treatment timing variables to the corresponding Period variables exactly.

CONCLUSION

The idea of dividing a study into Phases, Periods, and Subperiods may initially appear conceptually simple. But as with so many things in the worlds of clinical trial data collection and CDISC compliance, the application can quickly become complex. The case study in this paper presented further challenges in the study design including three groups of subjects with different visit schedules, and the groups did not all receive the same set of routes of administration. The concepts and tools presented were applied separately for each group, and combined where applicable and most efficient. The assignment of treatments and labeling of Phases, Periods, and Subperiods was revisited numerous times as the data was explored, and the different SDTM domains that were incorporated into the ADaM data had differences for which to adjust. The author recommends that the reader begins with a study schema that helps the study team visualize the way the study will be divided, and discuss with the Sponsor and Biostatistician information that is meaningful in dividing events and findings into the various time periods within the study. Then derive the start and end dates for each key time period carefully in ADSL, as this will be the reference into which the visit dates for the remaining ADaM data sets will be “dropped” in order to assign them to the proper time periods. It is a challenging but also interesting and ultimately satisfying puzzle to complete.

ACKNOWLEDGMENTS

Thank you to my supervisor and colleagues, editorial staff, and Westat senior leadership for your assistance and support in the writing of this paper.

DISCLAIMER: The contents of this paper are the work of the author and do not necessarily represent the opinions, recommendations, or practices of Westat.

RECOMMENDED READING

- *Analysis Data Model (ADaM) Version 2.1*
- *Analysis Data Model Implementation Guide Version 1.2*

CONTACT INFORMATION

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

Jennifer Fulton, Biostatistician
Westat
JenniferFulton@westat.com
<https://www.westat.com/>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

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