PharmaSUG 2019 - Paper SS-126

Considerations in Effectively Generating PK Analysis Input Datasets

Jianli Ping, Gilead Sciences Inc.

ABSTRACT

Generating high quality pharmacokinetic (PK) analysis input or PK merge data set in a timely manner is critical for PK parameter generation and downstream programming on CDISC compliant PC/PP data sets. While there is guidance for SDTM PC/PP and ADaM ADPC/ADPP detailed in the CDISC implementation guides, there is a lack of general standards for creating PK analysis input data sets. The different formats of source data as captured from CRF and PK lab can result in further challenges during PK merge data set creation. In this paper, general approaches for PK merge data set creation are proposed that can facilitate PK analysis and generation of PK CDISC compliant data sets and also as a source data for NonMEM programming. This paper also discusses the core PK merge data set variables with commonly selected matrices that should be included and their derivations through case studies and excerpts of SAS programs. Data set checks after PK merge data set generation are recommended, which can help identify possible data/programming issues.

INTRODUCTION

One of the major tasks of the statistical programming team is to provide pharmacokinetic (PK) analysis input, also called PKmerge, data sets for PK parameter analysis. In most cases, the PK scientists use the WinNonlin® software or NonMEM® software, which has special requirements for key variables that need to be included with specified formats. Some studies may need additional covariates to be obtained from multiple sources such as biomarker or laboratory data sets. While there is guidance for SDTM PC/PP and ADaM ADPC/ADPP detailed in the CDISC implementation guides (CDISC, 2013), there is a lack of general standards for creating PK analysis input data sets. The different patterns of workflow regarding PK data analysis from source data inputs to the generation of tables, figures, and listings (TFLs) have been discussed by many professional teams and may have their advantages and disadvantages (Cherukuri etc., 2015; Li etc., 2018). Each pharmaceutical company may have its own standard operating procedures (SOP) considering their indications of test drugs and turnaround time for their early stage decision making. Schaefer (2014) provided an overview of procedures to handle PK data in CDISC standards regarding what key variables should be included through the connection between PC and PP data sets.

From the standards point of view of a pharmaceutical company, it is essential to have input data sets able to fulfill the analysis requirements for PK parameters, efficiently facilitate the generation of SDTM/ADaM in a timely manner, and be flexible enough to include sufficient variables as a direct source for population PK modeling and simulation. Therefore, generating high quality pharmacokinetic analysis input or PKmerge data sets in a timely manner is critical for PK parameter generation and downstream programming on CDISC compliant PC/PP data sets.

In this paper we will discuss the core PKmerge data set variables with commonly selected matrices that should be included and their derivations to facilitate PK analysis and generation of PK CDISC compliant data sets and also as a source data for NonMEM programming. The procedures for PKmerge data set generation can improve work efficiency and consequent programming on data sets and TFLs. Some practices on quality control for PKmerge data sets will also be proposed.

OVERALL PK DATA WORKFLOW

The overall PK data workflow is summarized in Figure 1, in which PKmerge data set creation is a key component for overall PK data processing since PKmerge data set will be the source data for PK parameter analysis and for PC and ADPC generation. Additionally, PKmerge data set can also be the source data set for PopPK data set creation. Once PK parameter analysis is completed, the resultant parameter file in spreadsheet format will be used as input data for generating SDTM PP, following SDTM

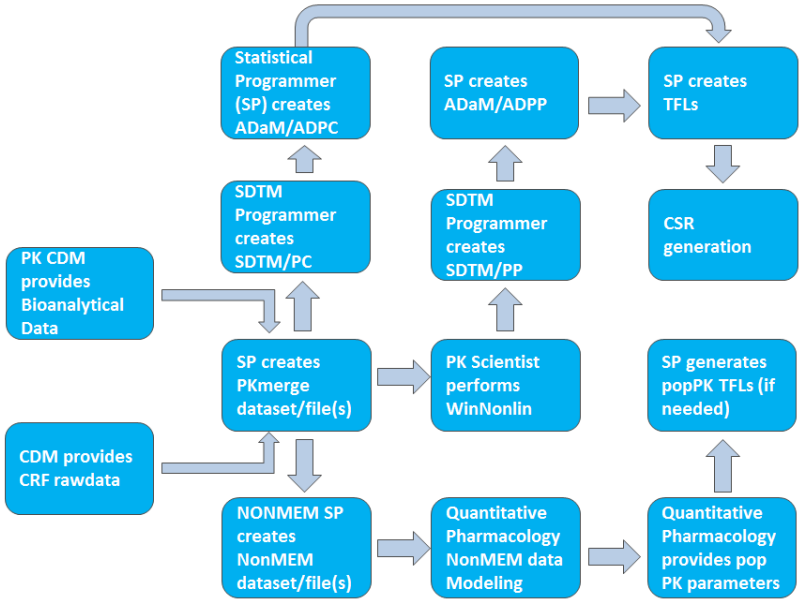


Figure 1 PK Data Workflow from Source Data Input through CSR Creation

implementation guide (CDISC, 2013). New variables for PK TFLs creation can be derived in ADPC/ADPP. Together with other ADaM data sets and TLFs, CSR can be completed.

The PKmerge data set being generated prior to PC creation in this PK data workflow is mainly due to four considerations: flexible PKmerge data structure with variables needed for parameter generation to be presented in horizontal manner; CDISC compliant data sets PC/ADPC with variables to be presented in

TPT	TPTNUM	SAMSTTIM	SAMENTIM	PCORRES	WNLCONCN	DEV	ENDEV	URVOL	BIOMATX
Predose	0	-0.25	-0.25	BLQ	0	-15	-15	174	Urine
0-6.0 h postdose	1000006	0	6	322	322	0	0	345	Urine
6.0-12.0 h postdose	1006012	6	12	200	200	0	0	600	Urine
12.0-24.0 h postdose	1012024	12	24	108	108	0	0	1325	Urine
24.0-48.0 h postdose	1024048	24	48	51.7	51.7	0	0	1798	Urine
48.0-72.0 h postdose	1048072	48	72	21.7	21.7	0	0	1512	Urine
72.0-96.0 h postdose	1072096	72	96	BLQ		0	0	2296	Urine
96.0-120.0 h postdose	1096120	96	120	BLQ		0	0	2534	Urine

TPT	PCDTC	PCENDTC	PCDOSDTC	SCHDTM	SCHENDTM
Predose	2017-02-16T08:12:00	2017-02-16T08:12:00	2017-02-16T08:27:00	2017-02-16T08:27:00	2017-02-16T08:27:00
0-6.0 h postdose	2017-02-16T08:27:00	2017-02-16T14:27:00	2017-02-16T08:27:00	2017-02-16T08:27:00	2017-02-16T14:27:00
6.0-12.0 h postdose	2017-02-16T14:27:00	2017-02-16T20:27:00	2017-02-16T08:27:00	2017-02-16T14:27:00	2017-02-16T20:27:00
12.0-24.0 h postdose	2017-02-16T20:27:00	2017-02-17T08:27:00	2017-02-16T08:27:00	2017-02-16T20:27:00	2017-02-17T08:27:00
24.0-48.0 h postdose	2017-02-17T08:27:00	2017-02-18T08:27:00	2017-02-16T08:27:00	2017-02-17T08:27:00	2017-02-18T08:27:00
48.0-72.0 h postdose	2017-02-18T08:27:00	2017-02-19T08:27:00	2017-02-16T08:27:00	2017-02-18T08:27:00	2017-02-19T08:27:00
72.0-96.0 h postdose	2017-02-19T08:27:00	2017-02-20T08:27:00	2017-02-16T08:27:00	2017-02-19T08:27:00	2017-02-20T08:27:00
96.0-120.0 h postdose	2017-02-20T08:27:00	2017-02-21T08:27:00	2017-02-16T08:27:00	2017-02-20T08:27:00	2017-02-21T08:27:00

Row	STUDYID	DOMAIN	USUBJID	PCSEQ	PCGRPID	PCREFID	PCTESTCD	PCTEST	PCCAT	PCSPEC	PCORRES
1	ABC-123	PC	123-0001	1	Day 1	A554134-10	DRGA_MET	Drug A Metabolite	ANALYTE	PLASMA	<0.1
2	ABC-123	PC	123-0001	2	Day 1	A554134-10	DRGA_PAR	Drug A Parent	ANALYTE	PLASMA	<0.1
3	ABC-123	PC	123-0001	3	Day 1	A554134-11	DRGA_MET	Drug A Metabolite	ANALYTE	URINE	<2
4	ABC-123	PC	123-0001	4	Day 1	A554134-11	DRGA_PAR	Drug A Parent	ANALYTE	URINE	<2
5	ABC-123	PC	123-0001	5	Day 1	A554134-11	VOLUME	Volume	SPECIMEN PROPERTY	URINE	3500
6	ABC-123	PC	123-0001	6	Day 1	A554134-11	PH	PH	SPECIMEN PROPERTY	URINE	5.5
7	ABC-123	PC	123-0001	7	Day 1	A554134-12	DRGA_MET	Drug A Metabolite	ANALYTE	PLASMA	5.4
8	ABC-123	PC	123-0001	8	Day 1	A554134-12	DRGA_PAR	Drug A Parent	ANALYTE	PLASMA	4.74
9	ABC-123	PC	123-0001	9	Day 1	A554134-13	DRGA_MET	Drug A Metabolite	ANALYTE	PLASMA	5.44

Figure 2 Examples of urine test as presented in PKmerge data set (top and middle) and PC Domain Model following SDTM Implementation Guide (version 3.2, bottom)

vertical manner, typically as one record per sample characteristic or time-point concentration per reference time point or per analyte per subject; efficient time management for PK data set and parameter generation in parallel; and elimination of possible data redundancy of covariates across different data sets. The structure differences between PKmerge and PC can be observed in Figure 2, where PKmerge has urine concentration, volume, start/end collection time, start/end time deviation, scheduled start/end collection time as column variables (top) while SDTM PC (bottom) copied from SDTM implementation guide (CDISC, 2013) has urine volume presented vertically in a similar manner to PK concentration records collected either at a time point or an interval. Since PKmerge data is not required for submission, some of the variables used for PK parameter analysis can be included into designated domains to avoid such data to be redundant. For example, ADA (Anti-drug Antibody) can be included in XD domain (Figure 3), weight and height included in VS domain and lab covariates included in LB domain, when all these data can be included in PKmerge.

VISIT	VISITN	TPT	TPTNUM	ADASCRN	ADACONF	ADATITR
Day 1	1	Predose	0			
Day 1	1	1.0 h postdose	1			
Day 1	1	2.0 h postdose	2			
Day 1	1	4.0 h postdose	4			
Day 1	1	8.0 h postdose	8			
Day 1	1	24.0 h postdose	24			
Day 1	1	48.0 h postdose	48			
Day 8	8	Single Anytime	168			
Day 15	15	Predose	0	Negative		
Day 22	22	Single Anytime	528			
Day 29	29	Predose	0	Negative		
Day 36	36	Single Anytime	864			
Day 43	43	Single Anytime	1032	Negative		
Day 57	57	Single Anytime	1368	Positive	Positive	1
Day 71	71	Single Anytime	1656	Positive	Positive	8
Day 85	85	Single Anytime	2040	Positive	Positive	8
Day 113	113	Single Anytime	2688.5333333	Positive	Positive	16

DOMAIN	XDSEQ	XDTESTCD	XDTEST	XDORRES	XDSTRESC	XDSTRESN	XDSPEC	XDBLFL	VISIT	XDDTC
XD	1	ADACONF	ADA Confirmed	Positive	POSITIVE	.	SERUM	Y	Day -1	2018-01-14T10:00:00
XD	2	ADACONF	ADA Confirmed	Positive	POSITIVE	.	SERUM		Day 15	2018-01-29T07:55:00
XD	3	ADACONF	ADA Confirmed	Positive	POSITIVE	.	SERUM		Day 29	2018-02-12T07:55:00
XD	12	ADASCRN	ADA Screening	Positive	POSITIVE	.	SERUM	Y	Day -1	2018-01-14T10:00:00
XD	13	ADASCRN	ADA Screening	Positive	POSITIVE	.	SERUM		Day 15	2018-01-29T07:55:00
XD	14	ADASCRN	ADA Screening	Positive	POSITIVE	.	SERUM		Day 29	2018-02-12T07:55:00
XD	23	ADATITR	ADA Titer	4	4	4	SERUM	Y	Day -1	2018-01-14T10:00:00
XD	24	ADATITR	ADA Titer	64	64	64	SERUM		Day 15	2018-01-29T07:55:00
XD	25	ADATITR	ADA Titer	16	16	16	SERUM		Day 29	2018-02-12T07:55:00

Figure 3 ADA variables included in PKmerge data set as covariates in PKmerge (top) and XD domain with detailed test results and sampling information (bottom)

There are increasing discussions and proposals regarding the PK process in recent years. A white paper by the PhUSE CSS Development team outlined a workflow from clinical and bioanalysis data sets to PC/PP TFLs generation, in which SDTM PC is the source data set for non-compartmental analysis (NCA) (Vandenhende et al., 2014). An NCA PK analysis CDISC ADaM standard-ADNCA has been under development, which would support NCA, PK and PD data review and generation of PK/PD TLFs (Lucius et al., 2018). The ADNCA would comply with ADaM standards and be flexible to work with different software packages. There is a comparison on four PK data processes and their Pros/Cons (Li et al. (2018). The workflow we follow is very similar to the Process 2 among those 4 processes. To improve work efficiency standard macros can be used to generate some common PKmerge variables. Since

PKmerge data contains all required variables for PK parameter analysis, those variables not needed for PK outputs can be removed from PC/ADPC while complying with general IG for SDTM and ADaM. Additionally, some time-related variables in PKmerge can also be directly populated into the data sets for PopPK to eliminate the reprogramming on some of those variables.

STEPS FOR PKMERGE DATA SET CREATION

Overall procedures for creating PKmerge data sets/files can be outlined into following eight steps.

1: A standard PKMerge specification template with all the core variables should be created with input from Biostatistician and PK scientist;

2: Convert PK concentration excel file to SAS data set if needed;

3: Programming on derived variables in SAS program;

4: Run pkmerge program to create pkmerge data set;

5: Run pkmerge_UAT;

6: Export pkmerge data set into csv file;

7: Self check and release for validation;

8: Deliver pkmerge.csv file and pkmerge_UAT output lists to PK scientist. In this paper, we focus on the steps that can help standardize programming procedures to improve the work efficiency.

Construct PKmerge Mapping Specification File

The standard PKMerge mapping specification document can be updated to incorporate study specific variables as required by PK scientists, as showed in Figure 4. This mapping specification also defines mapping rule and implemented SAS codes, which can ease programming work for variables like SAMTIME, ASSAYDTC by reusing pre-defined SAS codes and a macro call separately.

Variable Order	Domain	CDISC Variable Name	CDISC Variable Label	Type	Length	Mapping Action	Raw Variables	Implemented SAS Code	Mapping Rule
1	PKMERGE	STUDYID	Study Identifier	Char	15	ASSIGN	raw.dm.project	strip(project)	STUDYID
2	PKMERGE	INVID	Investigator Identifier	Char	5	ASSIGN	raw.dm.invid	invid	
3	PKMERGE	SUBJID	Subject Identifier for the Study	Char	20	ASSIGN	raw.pkconc.subjid raw.dm.subjid	subjid	The last 4 or 5 digit from USUBJID.
4	PKMERGE	TREAT	Treatment	Char	100	ASSIGN	raw.pkconc.treatment_id	strip(treatment_id)	For open-label or after data finalization, If treatment sequence is included in the study, TREAT is from pkconc.treatment_id . For other studies TREAT will be populated using the same SDTM methodology for the DM domain as ACTARM. For blinded studies, the variable is assigned real or dummy treatment based on TA and study need.

18	PKMERGE	SAMTIME	Actual Hours	Num	8	CODE		%m_sdtm_isodate(srcdt=pkdat_raw.pktim, dstdt=_pcdtc, srcfmt=raw);
								%m_sdtm_isodate(srcdt=exstdat_raw.exsttim ,dstdt=_pcdosdtc);
								
								if _pcdtc ne ' ' then _pcdtm=input(_pcdtc, e8601dt.);
								if _pcdosdtc ne '' then _pcdosdtm=input(_pcdosdtc,e8601dt.);
								
								if n(_pcdtm,_pcdosdtm)=2 then samtime=(_pcdtm-_pcdosdtm)/3600;
									Difference between Sample collection date and time and dosing date and time (for the PK sampling) in hours.
									
									This variable will be populated only for PLASMA, SERUM, WHOLE BLOOD and RADIOACTIIVTY (PLASMA and WHOLE BLOOD) or sample with one timepoint (not interval) collection.

76	PKMERGE	ASSAYDTC	Date of Assay Analysis	Char	20	MACRO	raw.pkconc.assay_date	%m_sdtm_isodate(srcdt=assay_date ,dstdt=assaydtc);	Populate from PKCONC file and keep in SUPPPC.

Figure 4 PKmerge mapping specification with selected variables illustrating how they are defined and determined with SAS codes and a macro call

Core Variables in PKmerge Data Set

The PKmerge data set primarily contains dosing information, timing variables in referencing to their dose administration, PK concentration, and covariates from demographics, lab, anti-drug antibody and vital signs. The variables in PKmerge data set can vary for different studies but the variables mentioned below as core PKmerge variables should be included for all studies and can be standardized with pre-defined codes and macros call.

EXDOSE: Populate dose per administration from CRF dosing data.

PCTESTCD: Assign pharmacokinetic test short name and will be carried to PC/ADPC.

TPT: Populate nominal time point from CRF.

TPTNUM: Derive planned time point number from TPT.

PCDTC: Populate PK sample collection date and time from CRF data.

BIOMATX: Populate biological matrix from PK concentration data.

SCHDTM: Derive PK sampling scheduled date and time based on dosing date and time of the PK sample and nominal time point.

SAMTIME: Derive elapsed time (SAMTIME) as the difference between sample collection date and time (PCDTM) and numeric dosing date and time (PCDOSDTM) (for the PK sampling) in hours.

SAMTMFD: Derive elapsed time from first dose (SAMTMFD) as the difference between sample collection date and time (PCDTM) and convert char (FDOSDTC) to numeric first dosing date and time (FDOSDTM) (for the PK sampling) in hours.

DEV: Derive deviation (DEV) in minutes as difference between the PK sampling date and time (PCDTM) and the scheduled date and time (SCHDTM) for each timepoint (PKTPT).

SAMPAGE: Derive sample age (SAMPAGE) in days as difference between ASSAYDT and the PK sampling date (PCDT).

PCORRES: Populate result or finding in original units, i.e. the concentration value.

WNLCONCN: ASSIGN WinNonlin concentration (WNLCONCN) to 0 when concentration is BLQ till the first quantifiable concentration and ASSIGN to missing after first quantifiable concentration.

WNLACTM: Assign first WinNonlin actual time (WNLACTM) to 0 if the timepoint is predose and the actual hour (SAMTIME) is a negative value.

PCNAM: Assign vendor name for PK sample testing.

Generate PKConc Data Set

With PK concentration data provided by vendors in spreadsheet or CSV format, there is a need to convert it into SAS data format. This step is performed by using a macro call after file name and output name being defined, as showed below.

```
%macro readpk(dsin=, dsout=);
   proc import datafile="../source/pkconc/qa/&dsin..xls"
     out=&dsout dbms=xls replace;
     guessingrows=4000;
     getnames=yes;
   run;
%mend readpk;
%readpk(dsin=%str(BA_PK_GS-US-xxx-xxxx_PK_CONC_Unblinded_QA_FRONTAGE_
02Jul18), dsout=pkconc0);
```

Time Points Conversion

Determining numeric values of time points, particularly intensive and interval samples, can be challenging since the same time points will be carried over into PC/ADPC to comply with the 1:1 mapping between TPT and TPTNUM. To provide unique timepoint for interval samples, we proposed seven digits values for numeric time points that combines the start and end timepoint interval. A macro can be used to convert character time points into numeric time points, and start and end time for an interval sample such as urine or feces, as showed in Figure 5.

```
%macro tptconv(indsn=, tpt=, tptchk=N);
data _&indsn. %if &tptchk=Y %then %do; &indsn._chk; %end;;
  set &indsn.;
  if index(upcase(&tpt), 'PREDOSE') then _tptnum=0;
  else if index(upcase(&tpt), 'MIN ') then _tptnum=round(input(scan(&tpt., 1, ' '), best.)/60, 0.001);
  else if index(upcase(&tpt), 'ANYTIME') then _tptnum=7000;
  else if index(upcase(&tpt), 'TROUGH') then _tptnum=8000;
  else if index(upcase(&tpt), 'TERMINATION') then _tptnum=9000;
  else if index(upcase(&tpt), 'UNSCH') then _tptnum=9999;
  else if index(upcase(&tpt), 'TO ') then do;
       _sttptn=input(scan(&tpt, 1, ' '), best.);
       _entptn=input(scan(&tpt, 3, ' '), best.);
       _tptnum=input('1'||put(_sttptn, z3.)||put(_entptn, z3.), best.);
  end;
  else if index(&tpt, '-') then do;
       _sttptn=input(scan(&tpt, 1, '-'), best.);
       _entptn=input(scan(scan(&tpt, 2, '-'), 1, ' '), best.);
       _tptnum=input('1'||put(_sttptn, z3.)||put(_entptn, z3.), best.);
  end;
  else if not (index(upcase(&tpt), 'TO ') and index(&tpt, '-')) and index(upcase(&tpt), 'H ')
       then _tptnum=input(scan(&tpt., 1, ' '), best.);
  else if not (index(upcase(&tpt), 'TO ') and index(&tpt, '-')) and index(upcase(&tpt), 'HR ')
       then _tptnum=input(scan(&tpt., 1, ' '), best.);
  %if &tptchk=Y %then %do;
    if _tptnum>. then output _&indsn.;
    else output &indsn._chk;
  %end;
run;
%mend tptconv;
%tptconv(indsn=tpttest, tpt=tpt, tptchk=Y);
%tptconv(indsn=tpttest, tpt=tpt, tptchk=N);
```

TPT	_TPTNUM	_STTPTN	_ENTPTN
Predose	0	.	.
5 min postdose	0.083	.	.
10 min postdose	0.167	.	.
0.25 h postdose	0.25	.	.
0.5 h postdose	0.5	.	.
1 h postdose	1	.	.
1.5 h postdose	1.5	.	.
2.5 hr postdose	2.5	.	.
288 h postdose	288	.	.
0-6.0 h postdose	1000006	0	6
0-48 h postdose	1000048	0	48
0 to 48 h postdose	1000048	0	48
24 to 48 h postdose	1024048	24	48
48 to 72 h postdose	1048072	48	72
480 to 504 h postdose	1480504	480	504
0 to 504 h postdose	1000504	0	504
Single Anytime	7000	.	.
Trough	8000	.	.
Early Termination	9000	.	.

Figure 5 The macro converts character into numeric time points with the SAS codes (top) and output example (bottom)

The resultant time points take the unit of hour as this is the typical unit required for NCA. When urine or feces sample is collected, the digits from 2nd through 4th contain the starting collection hour and digits 5th to 7th contain the ending collection hour, so a sample collected as '24 to 48 h postdose' would have an TPTNUM=1024048. If it is an accumulative period such as from 0-48, its TPTNUM would be 1000048. When it is sparse sample, a fixed value would be assigned such as 7000 for 'Single Anytime', 8000 for 'Trough' and 9000 for 'Early Termination'.

Pre-Defined SAS Code and/or Macro Call

The PKmerge mapping specification has pre-defined SAS codes for most of the variables. If the variables need input from multiple raw data sets, these variables are pre-programmed and included in the mapping specification file. A standard macro call reads the mapping specification file with the included programs and creates a SAS program and PKmerge data set. This automated process improves quality, reduce programming time and possible errors.

Setting Proper Number of Significant Digits

The PK concentration values are generally provided with 3 significant digits. However if there is a need to convert the values with meaningful number of significant digits the below macro m_sigdits() can be used for such purpose, as showed in Figure 6.

```
%macro m_sigdits(indata=,  /* the dataset with variable to be converted,
                              for example pkmerge */
          var=,  /* the variable to be converted, for example wnlconcn */
          digt=3); /* user defined number of significant digits, default as 3 */

%m_sigdits(indata=comp, var=wnlconcn);

data comp;
   set comp;
   if wnlconcn>. then pcorres=strip(wnlconcn_c);
run;
```

Figure 6 Macro m_sigdits converting a record with user-defined number of significant digits

To perform this conversion, a user would need to define the source data set, variable for conversion, and desired number of significant digits with 3 as default number. This macro can also be called if there is a need of conversion for PK parameter values as well.

PKmerge for Blinded Study

Sometime PKmerge data set needs to be generated for a blinded study before formal study unblinding for modeling purpose. In this case, the unblinded programmer external to the study can generate the unblinded PK Merge data set, masks the real subject IDs with the dummy subject IDs and provide it to the modeling team. The unblinded programmer can also generate the blinded PKmerge data set and provides it to the study SP to setup SDTM, ADaM data sets and TFLs.

PKmerge Output Control

Even though the PKmerge data set is initially created in SAS format, the input for WinNonLin and NONMEM software is csv format. When converting data from SAS to CSV, partial date/time in SAS data format may be converted into their date/time in csv format improperly such as filling ‘00:00’ to hours when no time provided. To reserve the partial data/time in their original format, a macro is developed that can apply quotation marks to records in the output file with the SAS ODS csv options when there are records with partial date/time (Figure 7).

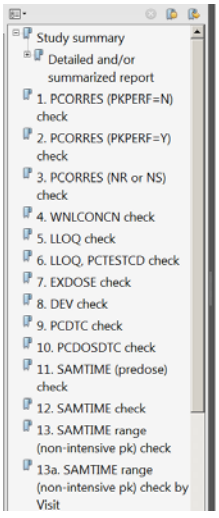
```
%macro m_csvexpt(dsin=pkmerge, /* pkmerge dataset name in rawdata folder */
                       csvver=, /* define csv file version */
                     textyn=N); /* define if text format is to be selected */

  %if &textyn^=Y %then %do;
    proc export data=rawdata.&dsin.
      outfile="../source/pkmerge/&study._pk_merge&csvver..csv" dbms=csv replace;
    run;
  %end;

  %if &textyn=Y %then %do;
    ods csv file="../source/pkmerge/&study._pk_merge&csvver..csv"
                  options(prepend_equals="yes" quote_by_type="yes");
    proc print data=rawdata.&dsin. noobs;
    run;
    ods csv close;
  %end;
%mend m_csvexpt;
```

```
%m_csvexpt(dsin=pkmerge, csvver=_final, textyn=N);
%m_csvexpt(dsin=pkmerge, csvver=_final, textyn=Y);
```

Figure 7 A macro to output PKmerge with quote to reserve the partial date/time in their original format

Quality Control for PKmerge Data Sets Creation

To ensure quality of PKmerge data files, the validations performed include independent programming of data set, self-check on key PKmerge variables, and applying Pkmerge_UAT utility. The PKmerge _UAT performs cross check on common PKmerge data variables and outputs a pdf file that contain a detailed output of summary check results (Figure 8). The output can be checked by production programmer, validation programmer and PK scientists with their focuses of functions.



Dataset Summary

Category	Summary
Study Number	GS-US-XXX-XXXX
Total Number of Subject	100
Site ID: Subject ID Range (Num)	01069: 30001-30025 (25)--01069: 10001-10025 (25)--01069: 20001-20025 (25)--01069: 40001-40025 (25)
BMI (min - max)	19.11 - 29.24
BSA (min - max)	1.48 - 2.34
First and Last Dose Date/Time	2017-06-20T08:20:00 - 2017-08-17T09:57:00
First and Last PK Sample Collection Date	2017-06-20 - 2017-08-18
Pharmacokinetic Test Name	GS-xxxx:Pharmacokinetic Concentrations
Treatment Sequence	Single dose of GS-xxxx 1 mg in AM - CIPRO 500 mg (BID on Day 7) - CIPRO 500 mg (BID on Day 8) + Single dose of GS-xxxx 1 mg in AM (Cohort 3);Single dose of GS-xxxx 1 mg in AM - PBC 500 mg (BID on Day 7) - PBC 500 mg (BID on Day 8) + Single dose of GS-xxxx 1 mg in AM (Cohort 2);Single dose of GS-xxxx 1 mg in AM - VORI 200 mg (BID on Day 7) - VORI 200 mg (BID on Day 8) + Single dose of GS-xxxx 1 mg in AM (Cohort 1);Single dose of GS-xxxx 2 mg - RIF 600 mg once a day (QD) in PM (Day 2-7) - RIF 600 mg in PM (12 hours after GS-xxxx) on Day 8 + Single dose of GS-xxxx 2 mg in AM (Cohort 4)
Dosing Period	<blank>;Period 1;Period 2
Treatment	<blank>;1 mg GS-xxxx;2 mg GS-xxxx
Visit	Day 1;Day 8
Intensive PK Samplings Count by Specimen and Analyte	Plasma: GS-xxxx:Day 1 (13); Day 8 (13)
Non-intensive PK Samplings by Specimen and Analyte	<blank>: Pharmacokinetic Concentrations:Day 8
LLOQ (unit) by Specimen and Analyte	Plasma, GS-xxxx: 0.5 (pg/mL)

Figure 8 A pdf output from running PKmerge_UAT utility that performs cross check showed on the left panel and detailed check results in the main body

ADVANTAGES AND DISADVANTAGES OF APPLIED PK DATA WORKFLOW

Advantages

It can be more efficient for PC/ADPC generation since all the variables from PKmerge data set can be directly used or easily determined with Pkmerge data are presented in horizontal structure.

Pkmerge data set can be very flexible in terms of variables that need to be included to facilitate the request on covariates from PK scientists and to provide source data set for PopPK programming.

This workflow can ease the PK data unblinding control with PKmerge being created in a separate folder where unblinded data access is easily granted to a small group who are directly involved in PK related data processing and analysis.

PKmerge programming can start as soon as CRF data and PK concentration files are available.

Disadvantages

This PK data work flow and discussed PKmerge generation procedures may need additional resources for programming Pkmerge data sets since PKmerge is generated with a separated programming procedure.

Additional time should be added to have exclusion flags and analysis comments added back to PKmerge data set after review. However, this may also happen in a similar manner even when PC and ADPC are used for NCA.

The PKPD programming has been described as inevitably challenging and numerous efforts have been attempted across FDA, universities, pharmaceutical industries, and CRO (Lucius et al, 2018; Su and Kang, 2018; Vandenhende et al., 2014; Schaefer 2014). Su and Kang (2018) outlined the strategies to improve efficiency and accuracy by means of standardization, quality improvement, and automation. The workflow and programming procedures discussed in this paper provided a set of practices that we are following in improving the PK programming efficiency, quality of PKmerge data sets, and utilization of PKmerge data set for both PK parameter analysis and source data for NonMEM programing. We are working on more automation of source data conversion, macro calling, and exporting of PKmerge data sets/files.

CONCLUSION

With NCA commonly used for PK parameter analysis in clinical trials, it is critical to create PKmerge input data in a timely manner and with high quality. This data file can be generated prior to or after PC/ADPC generation. Since data structures for both SDTM/ADaM are more strictly defined, there are advantages in generating PKmerge separately and prior to PC generation. Constructing a PKmerge mapping specification with key variables pre-defined or being able to call existing macros, and developing isolated functional macros can help improve the PKmerge programming efficiency.

REFERENCES

Cherukuri K, 2015. Programming Pharmacokinetic (PK) Timing and Dosing Variables in Oncology Studies: Demystified. PharmaSUG 2015. Available at https://www.pharmasug.org/proceedings/2015/PO/PharmaSUG-2015-PO06.pdf

Clinical Data Interchange Standards Consortium, Inc., 2013. Study Data Tabulation Model Implementation Guide: Human Clinical Trials. Version 3.2

Li X, Mehta S, and Elam E. 2018. Compare Pharmacokinetic Data Submission Processes in CDISC Environment. PharmaSUG 2018. Available at https://www.pharmasug.org/proceedings/2018/AD/PharmaSUG-2018-AD22.pdf

Vandenhende F, Burton I, Ahrweiler S, and Buchheit V. 2014. “Analyses and Displays Associated to Non-Compartmental Pharmacokinetics – With a Focus on Clinical Trials” A White Paper by the PhUSE CSS Development of Standard Scripts for Analysis and Programming Working Group. Available at http://www.phusewiki.org/wiki/images/e/ed/PhUSE_CSS_WhitePaper_PK_final_25March2014.pdf

Reinbolt L and MacDonald L. 2018. Coming soon: ADNCA and the PK submission. PharmaSUG 2018. Available at https://www.pharmasug.org/proceedings/2018/DS/PharmaSUG-2018-DS02.pdf

Schaefer P, 2014. Overview of Handling of PK Data in CDISC Standards. PK Webinar 2014. Available at https://www.cdisc.org/system/files/all/Education/PK_Webinar_Dec2014_CDISC_Published.pdf

Su J and Kang J. 2018. Challenges and Strategies in PKPD Programming. PharmaSUG 2018. Available at https://www.pharmasug.org/proceedings/2018/AA/PharmaSUG-2018-AA13.pdf

ACKNOWLEDGMENTS

I would like to thank Krishna Sivakumar, Director, Statistical Programming for her encouragement and guidance and providing constructive suggestions throughout the paper development. I would also like to thank Phase 1 group for the team support.

CONTACT INFORMATION

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

Jianli Ping
Jianli.Ping@gilead.com