Paper SA03

Preparing PK/PD ADaM Datasets and eCRT Data Package for FDA Submission

Junying Zhang, Gilead Sciences, Inc, Foster City, U.S.A.

ABSTRACT

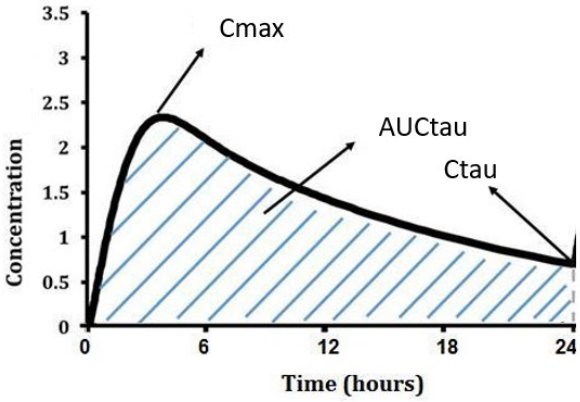
The CDISC Analysis Data Model (ADaM), which is required by FDA for clinical trial data submission, is designed to support analysis datasets. In this paper, we share our experience with creating the following ADaM datasets for pharmacokinetic & pharmacodynamic (PK/PD) analysis: ADSL and ADPKPD. The PK/PD ADaM electronic Case Report Tabulation (eCRT) data package has its own character. We will share our experience with creating the PK/PD ADaM eCRT data package and associated documentation for FDA electronic submission. It includes the ADSL and ADPKPD datasets in SAS$_{TM}$ Transport File Format (XPORT), the PK/PD analysis data reviewer's guide, define.xml and the SAS programs submitted in the eCTD module folder. This paper will only focus on the tasks of statistical programmers when preparing the PK/PD eCRT Data Package for FDA submission. The study and dataset information are masked to protect the privacy of the participants and confidential and proprietary information of Gilead Sciences, Inc. (Gilead).

INTRODUCTION

The PK/PD analysis studies the relationship between the drug exposure and the efficacy and safety of the drug. The PK/PD exposure refers to dose and various measures of integrated drug concentrations (e.g., AUCtau, Cmax, Ctau) in plasma and other biological fluids. Additionally, response refers to a direct measure of the pharmacologic effect of the drug, like the efficacy and safety effect of the drug. The derived PK/PD datasets (ADSL and ADPKPD) for analysis are complicated and contain population pharmacokinetics (Pop PK) parameters along with safety and efficacy data.

Figure 1 shows the meaning of three common Pop PK parameters that are used in the PK/PD analysis. AUCtau is the area under the plasma/serum concentration versus time curve over the dosing interval. Cmax is the maximum observed plasma/serum concentration of the drug. Ctau is the observed drug concentration at the end of the dosing interval. For drug dosed once daily, Ctau is the concentration level after dosing 24 hours.

Figure 1: Pop PK parameters AUCtau, Cmax, Ctau



SAMPLE STUDY123

In this paper, we share the experience of creating the PK/PD ADaM datasets for Study123 as well as preparing the eCRT data package for FDA regulatory submission. Study123 evaluates the efficacy and safety of Drug A in subjects with active ulcerative colitis disease. To assess the pharmacokinetic (PK) characteristics of Drug A is one of the secondary objectives of this study. Bioanalytical analysis for Drug A and its metabolite will be performed for samples collected from subjects in the active treatment group (Drug A high dose, Drug A low dose). The PK/PD electronic data package is created in conformance with CDISC standards.

PK/PD STATISTICAL ANALYSIS PLAN

For PK/PD analysis at Gilead, the PK/PD statistician prepares a PK/PD statistical analysis plan (SAP). The PK/PD SAP describes the statistical methods and data presentations to be used in the PK/PD analysis of Drug A in participants with active ulcerative colitis. The Pop PK models for Drug A and its metabolite using mixed effects modeling technique to characterize the PK of Drug A and metabolite are based on the data from Study123. The Pop PK parameters, AUCtau, Cmax, and Ctau for Drug A and its metabolite, are estimated based on the simulated concentration data using respective population PK models.

The PK/PD SAP contains detailed information on the PK data, PK/PD Analysis Sets, selected efficacy and safety endpoints, subject grouping and detailed statistical methodology to be used to analyze the Pop PK parameters with the PK/PD efficacy and safety endpoints. The PK/PD ADaM datasets and PK/PD tables, figures, and listings (TFLs) will be created according to the PK/PD SAP. PK parameters (AUCtau, Cmax, and Ctau for Drug A and metabolite) will be summarized by each dose level (Drug A high dose and Drug A low dose) separately and by pooled dose level respectively. Descriptive statistics (sample size, mean, standard deviation (SD), coefficient of variation [%CV], minimum, median, maximum, Q1, Q3, the geometric mean, its 95% confidence interval (CI), and the mean and SD of the natural-log transformed values) will be presented for the PK parameter data.

Only subjects with Pop PK exposures based on at least 1 measurable concentration will be included in the analysis. The PK/PD analysis includes PK exposure-efficacy analysis and PK exposure-safety analysis. Some selected efficacy and safety endpoints as examples are below.

Selected efficacy endpoints:
1) The proportion of subjects with endoscopy/bleeding/stool frequency (EBS) remission
2) The proportion of subjects with Mayo Clinical Score (MCS) remission

Selected safety endpoints:
1) The 6 most frequent treatment-emergent adverse events (TEAEs)
2) The 5 most frequent Grade 3 and above laboratory abnormalities.

SUBJECT GROUPING

All the PK/PD analyses were performed on the following 3 populations:

1) Population 1: All subjects who received Drug A;
2) Population 2: Only subjects who received Drug A high dose per day;
3) Population 3: Only subjects who received Drug A low dose per day;

For PK/PD analysis performed on Population 1 (i.e., pooled Drug A high and low dose), the determination of which octile subgroup a subject falls into is based on the rankings of PK exposure values (i.e., Drug A and metabolite AUCtau, Cmax and Ctau for PK exposure) from all the subjects, and with the number of observations approximately equally distributed within the 8 octile subgroups.

SAS procedure PROC UNIVARIATE is used to find percentiles of 12.5, 25 (i.e., Q1), 37.5, 50 (i.e., median), 62.5, 75 (i.e., Q3), and 87.5 values of AUCtau, etc. Values of AUCtau less than or equal to 12.5 percentile will be assigned to octile 1; values of AUCtau larger than 12.5 percentile and less than or equal to 25 percentiles will be assigned to octile 2; values of AUCtau larger than 25 percentile and less than or equal to 37.5 percentile will be assigned to octile 3; and so on. Values of AUCtau larger than 87.5 percentile will be assigned to octile 8.

For PK/PD analysis performed on Population 2 (i.e., Drug A high dose), or on Population 3 (i.e., Drug A low dose), the determination of which quartile subgroup a subject will fall into is based on the rankings of PK exposure values (i.e., Drug A and metabolite AUCtau, Cmax and Ctau for PK exposure) from the subjects who received Drug A high dose (for Population 2) or received Drug A low dose (for Population 3) respectively, and with the number of observations approximately equally distributed within the four quartile subgroups for each Population respectively.

Again, SAS procedure PROC UNIVARIATE will be used to find Q1, Q2 (i.e., median) and Q3 values of each Pop PK parameter (Drug A and metabolite AUCtau, Cmax and Ctau).

PK/PD TFLS

The PK/PD TFLs are created using the PK/PD ADAM datasets and are based on the PK/PD SAP. Analyses assessing the association between the AUCtau, Cmax and Ctau, and the efficacy endpoints will be performed. For each binary efficacy endpoint, the response rate (i.e., proportion of subjects with response) and the associated 95% confidence interval (CI) are summarized and plotted by the above specified AUCtau octile subgroups for Population 1, or by the above specified AUCtau quartile subgroups for Population 2 and 3, respectively.

In addition, descriptive statistics (sample size, mean, SD, %CV, minimum, median, maximum, Q1, Q3, geometric mean, geometric mean 95% CI, the mean and SD of the natural-log transformed values) will be presented for AUCtau, etc., by response value of the binary endpoint. Box plots of AUCtau, etc., will be displayed by the response value of the binary efficacy endpoint in the same way.

Some examples of the efficacy tables are below.

1) Summary Statistics of AUCtau by Octile Subgroups for Population 1 and Quartile Subgroups for Population 2 & 3.

2) Proportion of Subjects Achieving EBS Remission by AUCtau Octile Subgroup for Population 1 and Quartile Subgroup for Population 2 & 3.

3) Comparison of AUCtau by with or without Achieving EBS Remission

Analyses assessing the association between Drug A AUCtau, Cmax or Ctau (or metabolite AUCtau, Cmax or Ctau) and the safety endpoints are performed. For each binary safety endpoint, the proportion of subjects with the presence of the endpoint will be summarized and plotted by the above specified Drug A AUCtau, Cmax, or Ctau (or metabolite AUCtau, Cmax, or Ctau) octile (for Population 1) or quartile subgroups (for Population 2 and 3). In addition, subjects will be grouped into subgroups based on the presence and absence of each binary safety endpoint. The descriptive statistics (sample size, mean, SD, minimum, median, maximum, Q1, Q3) of Drug A AUCtau, Cmax or Ctau (or metabolite AUCtau, Cmax or Ctau) will be presented by present/absent of the selected binary endpoint. Box plot of the Drug A AUCtau (or metabolite AUCtau) by the present/absent of each binary safety endpoint will also be provided.

Some examples of the safety tables are below.

1) Proportion of Subjects with Selected TEAE by AUCtau or Cmax Octile Subgroups for Population 1

2) Proportion of Subjects with Selected TEAE by AUCtau or Cmax Quartile Subgroups for Population 2 & 3

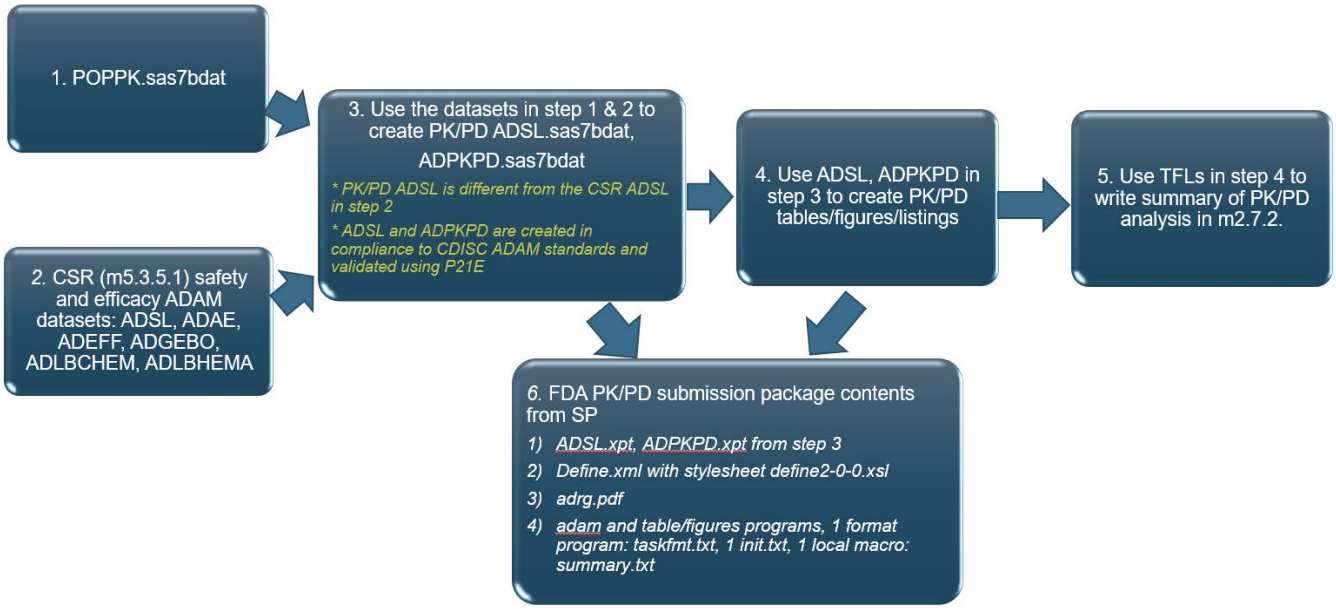
3) Proportion of Subjects with Selected Grade 3 or 4 TE Lab Abnormalities by AUCtau or Cmax Octile Subgroups for Population 1

4) Summary Statistics of AUCtau or Cmax by Selected Grade 3 or 4 TE Lab Abnormalities

PK/PD ADAM DATASETS

Figure 2 shows the flow chart of the PK/PD analysis data flow. It begins with using the POPPK.sas7bdat and clinical study report (CSR) safety and efficacy ADaM datasets to create the PK/PD ADSL, ADPKPD. Then the PK/PD ADSL, ADPKPD are used to generate the TFLs which are summarized in the Module 2.7.2 Clinical Pharmacology report.

Figure 2: Flowchart for PK/PD analysis data flow



PK/PD ADSL

As per the ADaM Implementation Guide (IG) and FDA Study Data Technical Conformance Guide, ADSL is a required analysis dataset for any standardized analysis data submission package. In addition to supporting all analyses specified in the SAP, ADSL contains variables to support baseline characteristics, PK/PD efficacy endpoints, PK/PD safety endpoints and Pop PK parameters. Demographic and baseline characteristics, population flags, treatment assignment, important treatment dates, etc. are copied from the CSR-ES.ADSL directly to PK/PD ADSL. CSR-ES is used to refer to the Clinical Study Report-Efficacy and Safety ADaM datasets prepared in eCTD module 5 in order to distinguish from the PK/PD ADSL. Variables for population flags (SAFFL, FASFL, PKPDFL), treatment assignment (TRT01P, TRT01A), and important dates related to treatment in study are copied into PK/PD ADSL. All subjects in the CSR-ES.ADSL dataset are included in PK/PD ADSL. Certain demographic and baseline characteristics variables in CSR-ES.ADSL, such as RFICDTC, RANDDTC, etc. that are not needed for the PK/PD TFL generation have been excluded from the PK/PD ADSL. All CDISC ADaM IG v1.1 ADSL Core=Req variables like STUDYID, USUBJID, SUBJID, SITEID are kept in PK/PD ADSL.

Pop PK parameter variables are copied from Pop PK dataset (POPPK.sas7bdat) to PK/PD ADSL. The Pop PK parameters of AUCtau, Cmax, and Ctau for Drug A and metabolite, will be estimated from the Pop PK modeling and provided by the PK scientist in the NONMEM.POPPK dataset. NONMEM refers to nonlinear mixed effects modelling and is the gold standard software in Population Pharmacokinetic and Pharmacokinetic-Pharmacodynamic modelling. The population PK models for Drug A and its primary metabolite using nonlinear mixed effects modeling technique to characterize the pharmacokinetics of Drug A and metabolite were based on the data from Study123. PK parameters (AUCtau, Cmax, and Ctau for Drug A and metabolite) were estimated based on the concentration data using respective population PK models. A masked NONMEM.POPPK dataset is shown in Figure 3 below.

Figure 3: Masked NONMEM.POPPK

	Study ID	Subject ID	Source	Dose	Drug A AUCtau (h*ng/mL)	Drug A Cmax (ng/mL)	Drug A Ctau (ng/mL)	Metabolite AUCtau (h*ng/mL)	Metabolite Cmax (ng/mL)	Metabolite Ctau (ng/mL)
1	Study123	10000	Drug A	low dose	1500	500	3			
2	Study123	10000	metabolite	low dose				36000	1900	1000
3	Study123	10001	Drug A	low dose	1750	370	2			
4	Study123	10001	metabolite	low dose				30000	1600	700
5	Study123	10002	Drug A	high dose	6000	2200	7			
6	Study123	10002	metabolite	high dose				65000	3200	1400
7	Study123	10003	Drug A	low dose	1950	360	5			
8	Study123	10003	metabolite	low dose				30000	1500	900

PK/PD efficacy and safety variables in study are created in PK/PD ADSL and copied to ADPKPD. PK/PD efficacy endpoints are derived from CSR-ES.ADEFF and ADGEBO. PK/PD safety endpoints are derived from CSR-ES.ADLB and CSR-ES.ADAE.

A masked PK/PD ADSL dataset is shown in Figure 4. Due to space limitations, we only show part of the PK/PD ADSL variables. Nasopharyngitis (Y/N) and Headache (Y/N) are two safety endpoints of adverse events with value “Y” and “N”. “Y” indicates the subject has experienced the adverse event and “N” means not. EBS Remission and MCS Remission are two efficacy endpoints with value “Y” and “N”. “Y” indicates the subject has achieved the efficacy endpoint and “N” means not.

Figure 4: Masked PK/PD ADSL dataset

	Study Identifier	Unique Subject Identifier	Age	Sex	Race	Safety Population Flag	Full Analysis Set Population Flag	Planned Treatment for Period 01	Actual Treatment for Period 01	PK/PD Analysis Set Flag
1	Study123	Study123-12345-10000	55	F	BLACK	Y	Y	Drug A low dose	Drug A low dose	Y
2	Study123	Study123-12345-10001	60	F	WHITE	Y	Y	Drug A low dose	Drug A low dose	Y
3	Study123	Study123-12345-10002	45	F	ASIAN	Y	Y	Drug A high dose	Drug A high dose	Y
4	Study123	Study123-12345-10003	30	M	WHITE	Y	Y	Drug A low dose	Drug A low dose	Y

Drug A AUCtau (h*ng/mL)	Drug A Cmax (ng/mL)	Drug A Ctau (ng/mL)	Metabolite AUCtau (h*ng/mL)	Metabolite Cmax (ng/mL)	Metabolite Ctau (ng/mL)	Nasopharyngitis (Y/N)	Headache (Y/N)	EBS Remission	MCS Remission
1500	500	3	36000	1900	1000	N	N	N	N
1750	370	2	30000	1600	700	N	N	N	N
6000	2200	7	65000	3200	1400	Y	N	N	N
1950	360	5	30000	1500	900	N	N	N	N

PK/PD ADPKPD

The PK/PD analysis dataset, ADPKPD, consists of randomized subjects who received at least one dose of Drug A and had at least one evaluable PK parameter of interest estimated from the Pop PK modelling (i.e. PKPDFL='Y'). The population indicator variables, treatment variables and variables used as covariates for statistical analyses (AGE, SEX, RACE, etc.) in PK/PD ADSL comprise the core variable list and are included in ADPKPD.

Since all the PK/PD analyses are performed on the population 1, 2, and 3, a record level DOSEGRP (with label "Dose Group") variable is created, to describe the record of the data used for Population 1, 2 or 3. The records for Population 1, 2, 3 are chosen by DOSEGRP = 'All Doses', 'Drug A high dose', or 'Drug A low dose'. For the PK/PD analyses performed on Population 1 (i.e., pooled Drug A high and low dose), subjects are categorized into 8 octiles based on the distribution of all subjects' Pop PK exposures. For analyses performed on Population 2 (i.e., Drug A high dose), subjects are categorized into 4 quartiles based on the distribution of the Pop PK exposures of the subjects who received Drug A high dose. For analyses performed on Population 3 (i.e., Drug A low dose), subjects are categorized into 4 quartiles based on the distribution of the Pop PK exposures of the subjects who received Drug A low dose.

Currently ADaM has three data structures: ADSL (Subject-Level Analysis Dataset), BDS (Basic Data Structure), and OCCDS (Occurrence Data Structure). These three structures correspond to the SUBJECT LEVEL ANALYSIS DATASET, BASIC DATA STRUCTURE, and OCCURRENCE DATA STRUCTURE classes of ADaM datasets. A BDS dataset contains one or more records per subject, per analysis parameter, per analysis timepoint. [1] Examples of BDS data include Laboratory Test Results, Vital Signs, Efficacy Analysis Dataset, etc. The OCCURRENCE DATA STRUCTURE is suitable for occurrence analysis which is the counting of subjects with a given record or term. Examples of occurrence data structure are Adverse Events, Concomitant Medications, and Medical History. Typically, findings data fit nicely into BDS, while events and most interventions fit nicely into OCCDS. [2] The ADPKPD follows the ADaM fundamental principles and other ADaM conventions, but does not follow one of the three defined structures (ADSL, BDS, OCCDS). So ADPKPD is classified into the class of ADAM OTHER.

A masked PK/PD ADPKPD dataset is shown in Figure 5. Due to space limitations, we only show part of the PK/PD ADPKPD variables. Dose Group=All Doses indicates population 1. The subgroup variables like Drug A AUCtau Subgroup, …, Metabolite Ctau Subgroup will have value in 1, 2, 3, 4, 5, 6, 7 and 8, indicating which octile the Pop PK parameters are in for the combined Drug A high and low dose group. Dose Group=Drug A high dose and Drug A low dose indicate population 2 and 3, respectively. The subgroup variables like Drug A AUCtau Subgroup, …, Metabolite Ctau Subgroup will have value in 1, 2, 3 and 4, indicating which quartile the Pop PK parameters are in for the Drug A high dose group or the Drug A low dose group.

Figure 5: Masked PK/PD ADPKPD dataset

	Study Identifier	Unique Subject Identifier	PK/PD Analysis Set Flag	Drug A AUCtau (h*ng/mL)	Drug A Cmax (ng/mL)	Drug A Ctau (ng/mL)	Metabolite AUCtau (h*ng/mL)	Metabolite Cmax (ng/mL)	Metabolite Ctau (ng/mL)	Nasopharyngitis (Y/N)
1	Study123	Study123-12345-10000	Y	1500	500	3	36000	1900	1000	N
2	Study123	Study123-12345-10000	Y	1500	500	3	36000	1900	1000	N
3	Study123	Study123-12345-10001	Y	1750	370	2	30000	1600	700	N
4	Study123	Study123-12345-10001	Y	1750	370	2	30000	1600	700	N
5	Study123	Study123-12345-10002	Y	6000	2200	7	65000	3200	1400	Y
6	Study123	Study123-12345-10002	Y	6000	2200	7	65000	3200	1400	Y
7	Study123	Study123-12345-10003	Y	1950	360	5	30000	1500	900	N
8	Study123	Study123-12345-10003	Y	1950	360	5	30000	1500	900	N

Headache (Y/N) is one safety endpoint of adverse event.

EBS Remission is one efficacy endpoint defined in SAP.

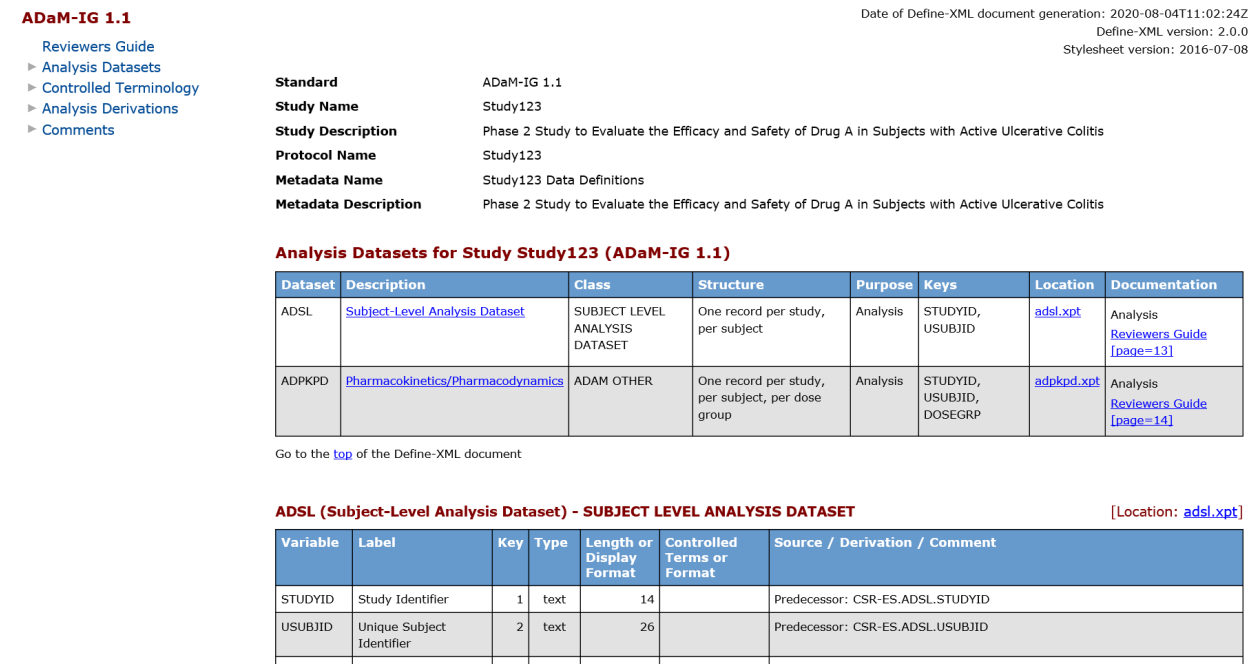
Drug A AUCtau of this subject is in the 4th octile of all PK population

Headache (Y/N)	EBS Remission	MCS Remission	Dose Group	Drug A AUCtau Subgroup	Drug A Cmax Subgroup	Drug A Ctau Subgroup	Metabolite AUCtau Subgroup	Metabolite Cmax Subgroup	Metabolite Ctau Subgroup
N	N	N	All Doses	4	4	3	5	5	4
N	N	N	Drug A low dose	1	1	1	1	1	1
N	N	N	All Doses	8	8	8	8	8	8
N	N	N	Drug A low dose	4	4	4	4	4	4
N	N	N	All Doses	4	5	4	4	4	4
N	N	N	Drug A high dose	1	1	1	1	1	1
N	N	N	All Doses	7	7	5	6	7	6
N	N	N	Drug A low dose	3	3	2	2	3	2

DEFINE.XML V2 FOR FDA

Define.xml version 2.0.0 is used to describe CDISC SDTM, SEND and ADaM datasets for the purpose of submissions to the FDA, as well as any proprietary (non-CDISC) dataset structure. The define.xml is implemented using extensions to the CDISC Operational Data Model (ODM) XML schema. [3] Figure 6 is define.xml created from the PK/PD ADaM specification. Parameter Value Level Metadata does not exist in PK/PD define.xml since none of the analysis datasets are Basic Data Structure datasets.

Figure 6: Masked PK/PD define.xml v2 for FDA submission



ADaM-IG 1.1

- Reviewers Guide
- Analysis Datasets
- Controlled Terminology
- Analysis Derivations
- Comments

Date of Define-XML document generation: 2020-08-04T11:02:24Z
Define-XML version: 2.0.0
Stylesheet version: 2016-07-08

Standard	ADaM-IG 1.1
Study Name	Study123
Study Description	Phase 2 Study to Evaluate the Efficacy and Safety of Drug A in Subjects with Active Ulcerative Colitis
Protocol Name	Study123
Metadata Name	Study123 Data Definitions
Metadata Description	Phase 2 Study to Evaluate the Efficacy and Safety of Drug A in Subjects with Active Ulcerative Colitis

Analysis Datasets for Study Study123 (ADaM-IG 1.1)

Dataset	Description	Class	Structure	Purpose	Keys	Location	Documentation
ADSL	Subject-Level Analysis Dataset	SUBJECT LEVEL ANALYSIS DATASET	One record per study, per subject	Analysis	STUDYID, USUBJID	adsl.xpt	Analysis Reviewers Guide [page=13]
ADPKPD	Pharmacokinetics/Pharmacodynamics	ADAM OTHER	One record per study, per subject, per dose group	Analysis	STUDYID, USUBJID, DOSEGRP	adpkpd.xpt	Analysis Reviewers Guide [page=14]

Go to the top of the Define-XML document

ADSL (Subject-Level Analysis Dataset) - SUBJECT LEVEL ANALYSIS DATASET [Location: adsl.xpt]

Variable	Label	Key	Type	Length or Display Format	Controlled Terms or Format	Source / Derivation / Comment
STUDYID	Study Identifier	1	text	14		Predecessor: CSR-ES.ADSL.STUDYID
USUBJID	Unique Subject Identifier	2	text	26		Predecessor: CSR-ES.ADSL.USUBJID

ANALYSIS DATA REVIEWER'S GUIDE

The PK/PD analysis data package is slightly different from the analysis data package we prepare for CSR-Efficacy and Safety. The PK/PD analysis data reviewer's guide (ADRG) is created to provide regulatory reviewers with a clear, concise set of instructions that facilitates the review process. ADRG is a required component of the ADaM electronic data package. The PK/PD ADRG is created following the PHUSE ADRG template. [4] Because the majority of the data in the PK/PD analysis datasets stem from the study CSR analysis datasets, ADRG for the PK/PD analysis database is initiated by copying the majority of the study CSR ADRG. Because of the different source datasets used to create the PK/PD ADaM datasets, ADRG Section 1.4 Source Data Used for Analysis Dataset Creation explains in detail how we create the PK/PD ADSL and ADPKPD from the source NONMEM.POPPK and CSR efficacy and safety ADaM datasets.

ADRG for FDA submission is saved as adrg.pdf and the file should satisfy certain pdf properties according to the Portable Document Format (PDF) Specifications from the FDA (September 2016 v4.1). [5] For example, File Options of Compatibility is set to 'Acrobat 5.0 (PDF 1.4)' and File Option of Optimize for fast web view is checked.

In PK/PD ADRG, Section 1.3 Study Data Standards and Dictionary Inventory, we use the same standard or dictionary as documented in the CSR.ES ADaM ADRG. Since the PK/PD ADaM electronic data package does not have the ADaM Structure for Occurrence Data (OCCDS) Version 1.0, this is deleted in the PK/PD analysis data reviewer's guide Section 1.3.

Table 1: ADRG Section 1.3 Study Data Standards and Dictionary Inventory

Standard or Dictionary	Versions Used
ADaM	2.1
ADaM IG	1.1
SDTM	1.4
SDTM IG	3.2
Controlled Terminology	NCISDTM 2018-12-21, NCIADaM 2018-12-21
Data Definitions	Define 2.0.0
Medications Dictionary	WHODrug BMAR19
Medical Events Dictionary	MedDRA 22.1
Toxicity Grading Dictionary	CTCAE v4.03

In ADRG Section 5.2, 'Analysis Dataset', Table 2 specifies the functional category or categories for each analysis dataset. It includes categories of Efficacy, Safety, Baseline or Other Subject Characteristics, PK/PD, and Primary Objective. Some datasets may belong to both Safety and Efficacy. The PK/PD ADSL dataset contains the demographic and baseline characteristics, efficacy and safety endpoints; therefore, the three categories are checked for ADSL. The PK/PD ADPKPD dataset carries over the efficacy and safety endpoints from PK/PD ADSL as well as creating PK/PD related analysis variables, so the three categories are checked for ADPKPD.

Table 2: ADRG Section 5.2 Analysis Datasets

Dataset Dataset Label	**Class**	**Efficacy**	**Safety**	**Baseline or other subject characteristics**	**PK/PD**	**Primary Objective**	**Structure**
ADSL Subject-Level Analysis Dataset	Subject Level Analysis Dataset	X	X	X			One record per study, per subject
ADPKPD PK/PD Analysis Dataset	ADaM Other	X	X		X		One record per study, per subject, per dose group

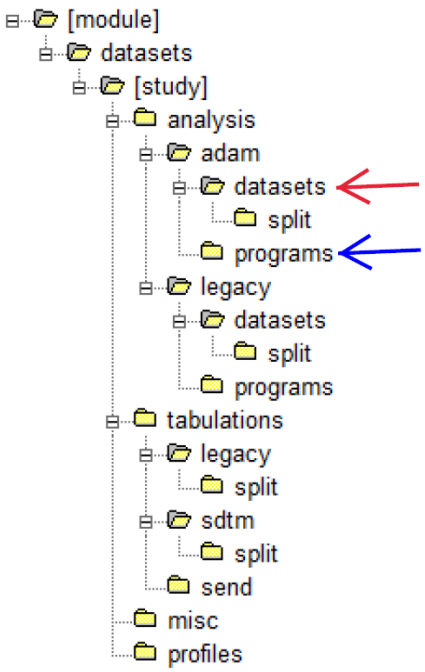
PK/PD DATA PACKAGE FOLDER STRUCTURE

The specification for organizing datasets and their associated files in folders within a submission is summarized in Figure 7, as noted in the FDA Study Data Technical Conformance Guide (TCG). [6] For ease of use with the define file and in the electronic Common Technical Document (eCTD) folder structure, all analysis datasets should be kept in one folder. From ADaMIG v1.1, if a set of analysis datasets includes an ADaM-compliant ADSL dataset (as required for a CDISC-conformant submission), then the whole set of analysis datasets should be placed into the /adam folder. If not, the whole set of analysis datasets should be

placed into the /legacy folder. Since we have created the ADaM-compliant ADSL dataset and ADPKPD datasets, the analysis datasets are placed in the /module/datasets/study/analysis/adam/datasets/ folder. The below figure is borrowed from TCG. [6]

The PK/PD data package is submitted as a separate study under the module 5/datasets eCTD folder. PK/PD ADaM XPT files, adrg.pdf, define.xml and the stylesheet define2-0-0.xsl are put in the /datasets folder as indicated in the red arrow in Figure 7. All SAS programs (in .txt format) created for generation of analysis datasets and safety and efficacy summary tables and figures are put in the /programs folder as indicated in the blue arrow in Figure 7.

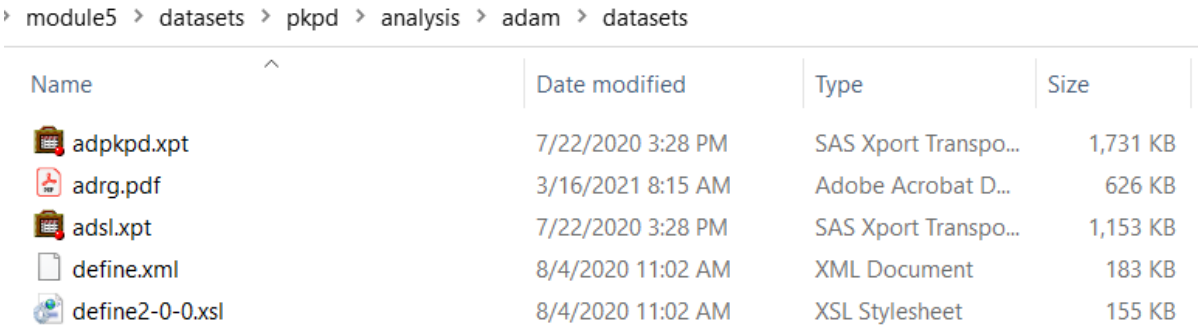
Figure 7: Folder Structure for Study Datasets



ANALYSIS DATA WITHIN THE ECTD STRUCTURE

Besides the PK/PD ADaM datasets in xpt format (adsl.xpt, adpkpd.xpt), the adrg.pdf, define.xml and the stylesheet define2-0-0.xsl are also included in the /module5/datasets/pkpd/analysis/adam/datasets folder. Any folders that are not populated in the study folder structure, such as the /split folder, the /legacy analysis folder, and the /tabulations folder are removed before the final submission to FDA. Figure 8 is an example of the files in the /datasets folder.

Figure 8: Files in datasets folder



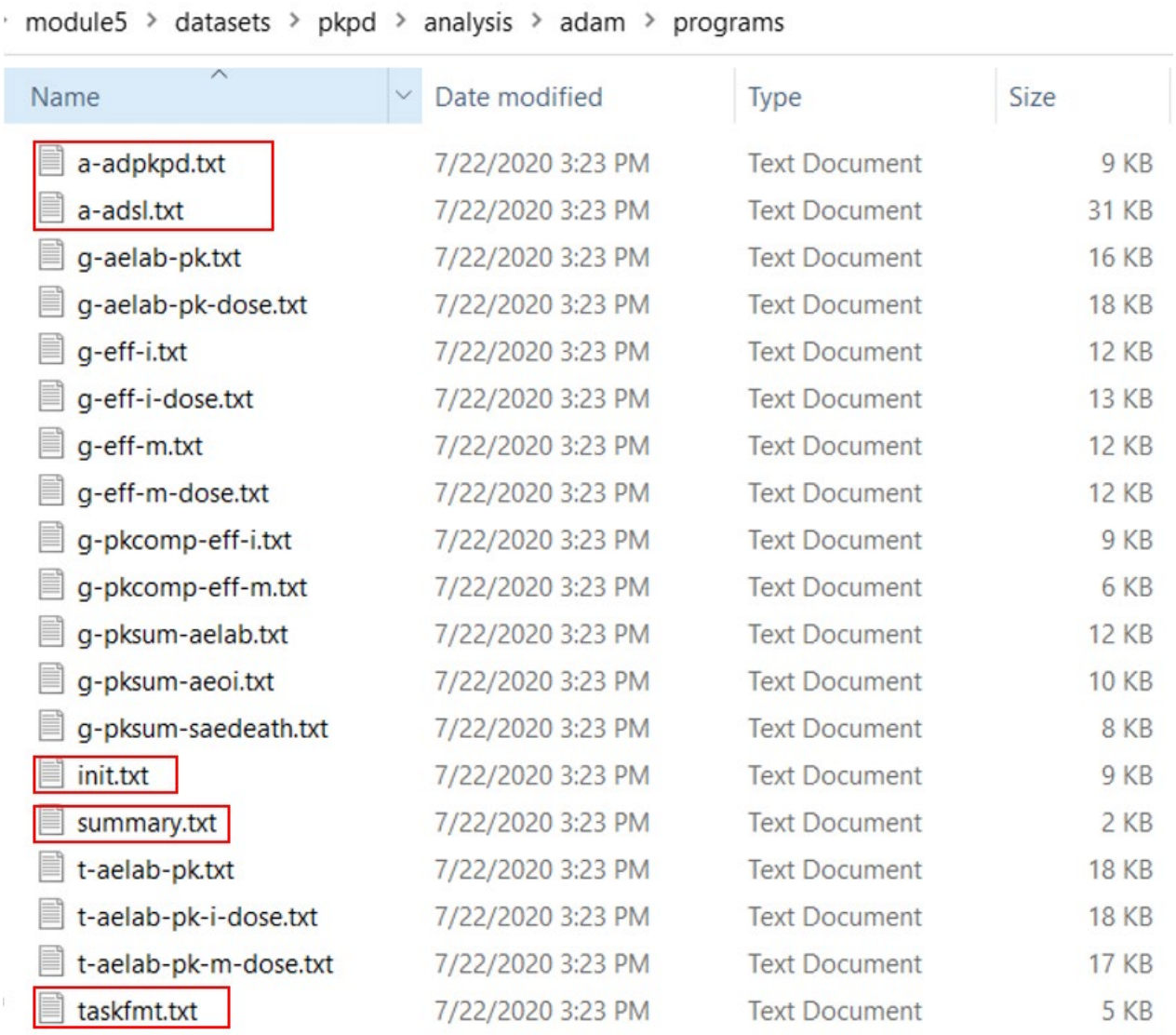
ANALYSIS PROGRAMS WITHIN THE ECTD FOLDER STRUCTURE

ADRG section 7 explains in detail the submission of SAS programs along with macros for analysis dataset generation and TFL creation. The SAS programs in txt formats are placed in the /module5/datasets/pkpd/analysis/adam/programs/ folder. All SAS

programs for analysis datasets and safety and efficacy summary tables and figures will be submitted. They are all created using SAS v9.4 and converted to .txt file. The submitted programs also include a format generating program that controls the formats used in the generation of the ADaM datasets and summary tables and figures, macro definition initialization programs that were included into all submitted programs, and macros that were used either in the generation of the ADaM datasets or of the summary tables and figures. Figure 9 is a screenshot of part of the programs in the programs/ folder. A detailed explanation of programs is as below.

- a-adsl.txt, a-adpkpd.txt, ADaM generation programs.
- taskfmt.txt contains user-defined SAS formats that are used in ADaM programs and tables, figures and listings programs.
- init.txt is task initialization file used to set options and macro variables used for all programs.
- summary.txt is a local macro program.
- All others are tables and figures programs.

Figure 9: Files in the programs folder



PINNACLE 21 ENTERPRISE VALIDATION

Pinnacle 21 Enterprise$_{TM}$ (P21E) is used to validate the PK/PD ADaM data package. Except for the Pop PK parameters, the variables derived in the PK/PD ADaM datasets use information from the CSR safety and efficacy ADaM datasets, so the Standard (ADaM-IG 1.1), SDTM CT, ADaM CT, MedDRA and validation engine (FDA 2010.1) are the same as the CSR.ES ADaMIG package. Figure 10 is the P21E ADaMIG data package which is used to validate the PK/PD ADaM electronic submission package for CDISC conformance.

Figure 10: PK/PD P21E ADaMIG data package

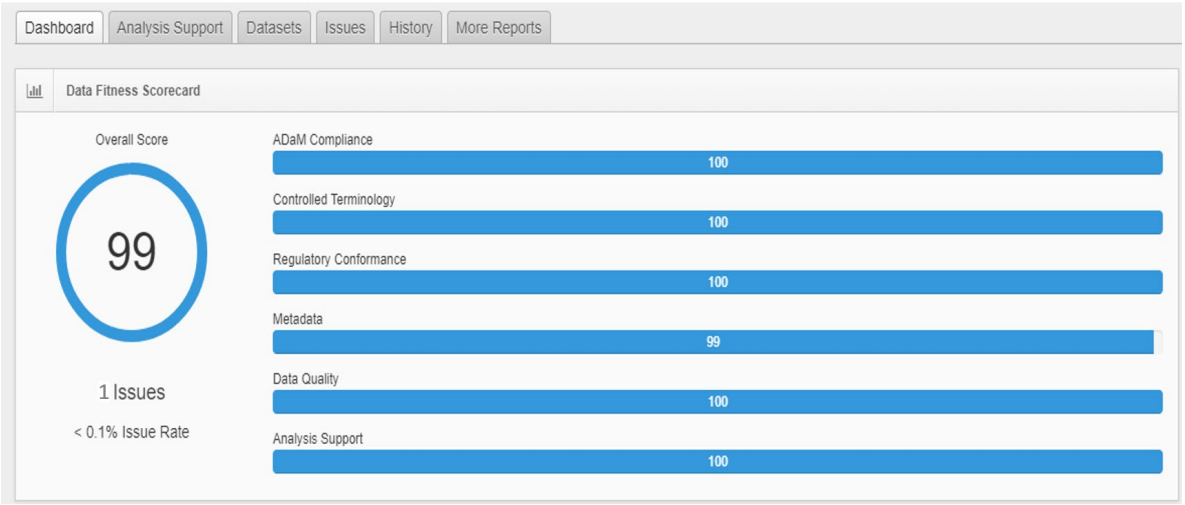
Name	Standard	SDTM CT	ADAM CT	SEND CT	MedDRA	WHODD	LOINC	SNOMED	Compatibility	Last Updated
ADaMIG	ADaM-IG 1.1	2018-12-21	2018-12-21		22.1				FDA 2010.1	2021-04-14 07:55:34

adsl.xpt, adpkpd.xpt, adrg.pdf, define.xml, and define2-0-0.xsl are put in a .zip file and P21E is used to validate the .zip data package. The P21E validation report shows one warning which could not be resolved and are documented in the ADRG section 6.2 as in Table 3. The P21E data fitness score is 99% as shown in Figure 11 and the electronic data package is of very good quality.

Table 3: PK/PD ADRG Section 6.2 Issues Summary

Dataset(s)	Diagnostic Message and/or Check ID	Severity	Count/ Issue Rate	Explanation
ADSL	CT2002: RACE value not found in 'Race' extensible codelist	Warning	75/3.7%	New terms were added to extensible codelist 'Race' per study protocol needs. Known limitation in Pinnacle 21 Enterprise.

Figure 11: PK/PD P21E validation data fitness scorecard



CONCLUSIONS

Creating the CDISC ADaM-compliant PK/PD ADSL and ADPKPD as well as the define.xml and ADRG will help the regulatory agencies to better understand the PK/PD data structure and how the analysis tables and figures are created using the ADaM datasets. This working process also involves close cooperation with clinical pharmacology, the PK/PD biostatistician and CSR safety and efficacy programming. Hope this paper is helpful for programmers who are going to work on the PK/PD analysis.

REFERENCES

[1] Analysis Data Model Implementation Guide Version 1.1
https://www.cdisc.org/system/files/members/standard/foundational/adam/ADaMIG_v1.1.pdf

[2] ADaM Structure for Occurrence Data (OCCDS) Version 1.0
https://www.cdisc.org/system/files/members/standard/foundational/adam/ADaM_OCCDS_v1.0.pdf

[3] CDISC Define-XML Specification Version 2.0
https://www.cdisc.org/standards/data-exchange/define-xml

[4] Analysis Data Reviewer's Guide (ADRG) Package
https://advance.phuse.global/display/WEL/Deliverables#:~:text=Analysis%20Data%20Reviewer%27s%20Guide%20(ADRG)%20Package

[5] Portable Document Format (Pdf) Specifications from the FDA (September 2016 v4.1)
https://www.fda.gov/files/drugs/published/Portable-Document-Format-Specifications.pdf

[6] STUDY DATA TECHNICAL CONFORMANCE GUIDE
https://www.fda.gov/media/152175/download

ACKNOWLEDGEMENTS

I would like to thank my colleague Shu-Min Chuang for inspiration from her many PK/PD studies at Gilead. I would like to acknowledge the following colleagues at Gilead for their knowledge-sharing, contribution, and review of this paper: Yuhui Sun, Aatiya Zaidi, Junming Yang, Amy Caron, Uta Meyer, Anna Kunina, and Santosh Kumar Lingala.

CONTACT INFORMATION

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

Junying Zhang
Gilead Sciences, Inc.
333 Lakeside Dr
Foster City, CA 94404
Work Phone: +1 (650) 522-5759
Email: junying.zhang@gilead.com
Web: https://www.gilead.com