PharmaSUG 2026 - Paper PO-386

Case Study: Integrating ADPPK CDISC Standards into Pharmacometric Programming and Analysis Workflows

Praseeda Rajan, Renuka Hegde, Neelima Thanneer, Yue Zhao and Erin Dombrowsky: Bristol Myers Squibb

ABSTRACT

With the release of the CDISC ADaM Population Pharmacokinetic (popPK) Implementation Guide (IG) and anticipated FDA requirements, pharmacometric programmers must adapt to new standards for creating ADPPK datasets. This case study provides a comprehensive guide for integrating ADPPK CDISC standards into pharmacometric programming and analysis workflows. Achieving compliance necessitated extensive workflow changes, from dataset preparation to submission, including accurate metadata alignment and updating internal tools. A key challenge was harmonizing and pooling multiple studies into a single CDISC-compliant ADPPK dataset, requiring mapping of historical variables and inclusion of standardized variables (e.g., PARAM/PARAMCD, AVAL/AVALC, RACE, SEX). Analysis-specific variables (e.g., ARACE, AGEGR1) were derived as recommended, and model files were updated for compatibility with new variable requirements. Permissible and conditional flag variables were leveraged for record exclusion within model files. The result was a single dataset suitable for both modeling and regulatory submission, supporting automation and tool development. Conformance to the ADPPK IG enhances data quality, consistency, and facilitates seamless interactions with health authorities. This case study serves as a practical guide for programmers and modelers to achieve robust compliance with emerging regulatory standards.

INTRODUCTION

With the release of the Clinical Data Interchange Standards Consortium (CDISC) Implementation Guide (CDISC. 2023) and the forthcoming FDA recommendation to submit datasets in ADPPK format, comprehensive changes are required to the dataset preparation-to-final submission workflow to comply with the ADPPK IG. Pharmacometric programmers face the challenge of ensuring compliance with the ADPPK IG without prior experience in programming ADaM datasets.

This paper illustrates a case study aimed at ensuring that dataset formats meet analysis software requirements while adhering to ADPPK IG standards. The goal is to bridge the gap between CDISC compliance requirements and the practical needs of pharmacometric modeling software such as Non Linear Mixed Effects Modelling (NONMEM).

Population PK (popPK) is a model-based representation of Pharmacokinetic (PK) processes with a statistical component, enabling identification of the sources of inter- and intra-individual variability. This approach is well suited for large heterogeneous PK datasets generated as part of standard multi-study clinical programs and is often used as a basis for simulations to inform dose selection and other milestones of drug development. The Analysis Data Model (ADaM) ADPPK standard is intended to provide a dataset template for model-based population pharmacokinetic analysis, utilized as an input into a software package specific to popPK analysis.

Key objectives include:

- Ensuring datasets conform to CDISC ADPPK IG variable requirements.
- Providing practical guidance on naming conventions, variable types, and imputation.
- Addressing common programming challenges encountered during integration.
- Demonstrating how industry-wide standards improve data quality and automation.

METHODS

ADPPK datasets should be created to be “analysis ready,” ensuring all necessary variables are included for its intended analytical use. The source of each variable is traced back to its immediate predecessor dataset, such as dosing or subject-level datasets, following CDISC guidelines. Multiple CDISC datasets can serve as input sources to populate the ADPPK dataset, with ADaM datasets being the most used, though other sources may also contribute as needed. Variables are incorporated as outlined in the ADaM IG to meet the CDISC standard. All variable metadata must be consistent with the ADaM IG.

Variables fall into three categories (Table 1):

Variable Type	Description
Required Variables	Minimum variables mandated for inclusion
Permissible Variables	Optional variables based on analysis needs
Conditional Variables	Variables included only if specific conditions are met

Table 1 . Variable Classification

Below are the key required variables (Table 2) and common standard Basic Data Structure variables in the CDISC ADPPK dataset that are available in other ADaM datasets (Table 3).The list is not all inclusive and many variables particularly covariates are study-specific.

Variable Name	**Variable Label**
STUDYID	Study Identifier
USUBJID	Unique Subject Identifier
USUBJIDN	Unique Subject Identifier (N)
AFRLT	Actual Rel Time from First Dose
EVID	Event ID
DV	Dependent Variable Result
MDV	Missing Dependent Variable Result
AMT	Actual Amount of Dose Received (unit)

Table 2 . ADPPK IG - Key Required Variables

Variable Name	**Variable Label**
AGE	Age
SEX	Sex
RACE	Race
AVISIT	Analysis Visit
DOSEP	Planned Dose
PARAM/PARAMCD	Param/Param Code
TRTA	Actual Treatment
ADY	Analysis Day

Table 3 . ADPPK IG - Common BDS Variables

Variables should be arranged in a logical, meaningful order that is relevant to their intended use (e.g., identification and event variables, time variables, covariates) and not simply alphabetic.

It is recommended to follow the general naming conventions and variable types as suggested in the ADaM IG. Suffixes BL, N, and I are appended to variables to indicate baseline, numeric, and imputed variables, respectively (Table 4). <COV> is an alias identifying a covariate (e.g. WT).

Naming Convention	Description	Example
<COV>BL	baseline covariate	WTBL, BMIBL
<COV>N	numeric version of categorical covariate; one-to-one relationship with <COV>	SEXN, RACEN
<COV>I	covariate with imputed values	WTI, BMII
<COV>GRy	grouping covariates	AGEGR1

Table 4 . Variable Naming Convention

General Imputation Conventions are shown below:

- Time-Variant Covariates: Derive using related time-varying covariates at the same time point. If values are missing and imputation is required, follow the pharmacometric Analysis Plan (e.g., Last Observation Carried Forward (LOCF)).

- Missing Dates and Time Records: Avoid imputations where possible and ensure Case Report Forms (CRFs) collect all necessary data. If date/time imputations are required, follow the standard imputation procedures, document them (CDISC. 2023) and flag the imputed records.

- Infusion duration: When infusion start or stop times are missing, use the protocol-defined duration alongside available sample times (e.g., trough, End of Infusion (EOI), or pre-dose) to back-calculate the missing time. Always impute duration to the protocol-defined value if the recorded duration deviates by ±100% or more and flag all imputed records.

For pooled datasets with differing variable names in historic studies, variables should be renamed in compliance with ADPPK IG based on information in the specifications.

RESULTS

Presented below are several examples of frequently asked questions by pharmacometric programmers encountered during the development of ADPPK:

HOW SHOULD RACE BE REASSIGNED FOR ANALYSIS?

If RACE classification requires reassignment for analysis, such as modifications from source data (Study Data Tabulation Model (SDTM) or ADaM) it is recommended to create a distinct variable for this analysis (e.g., ARACE). Ethnicity can be recoded specific to the analysis using AETHNICJ/ETHNICJ.

To optimize dataset compactness, it is recommended to avoid including both numeric and character forms of the same covariate unless they are necessary for analytical or reporting purposes. Typically, RACE and RACEN are retained; if AVISITN is not needed, only AVISIT should be incorporated.

RACE	RACEN	ARACE	AETHNICJ
White	1	White	Non-Japanese
Japanese	3	Asian	Japanese

HOW CAN GROUPING VARIABLES BE NAMED?

Grouping variables can be created by using the source variable name and adding a grouping indicator. For example, the grouping variable for age ≥60 and age <60 is named AGEGR1:

AGEGR1	AGE
Age < 60	44
Age ≥ 60	66

HOW CAN PK, DOSE, AND PD RECORDS BE REPRESENTED IN PARAM AND PARAMCD WHEN THERE ARE MULTIPLE PK ANALYTES OR PD PARAMETERS?

Simplify PARAM/PARAMCD as PK for pharmacokinetics, DOSE for dose, and PD for pharmacodynamics. For multi-analyte datasets, retain PARAM/PARAMCD and use DVID/DVIDN to distinguish analytes for analysis software.

Multiple Pharmacodynamic parameters can be coded as PD1, PD2, etc., or named per ADPPK IG and sponsor specifications.

Example:

PARAM	PARAMCD	EVID	DVID	DVIDN
Pharmacokinetics	PK	0	ABC123	1
Pharmacokinetics	PK	0	ABC345	2
Dose	DOSE	1		
Diastolic Blood Pressure	DIABP	2	DIABP	4
Systolic Blood Pressure	SYSBP	2	SYSBP	5

HOW IS AVAL ASSIGNED FOR PK, DOSE, AND PD RECORDS?

AVAL denotes the numerical value attributed to the PARAM or PARAMCD variable within the ADPPK dataset.

PARAMCD	EVID	AVAL	DV	AMT
PK	0	0.1	0.1	
DOSE	1	50		50
DIABP	2	82	82	

HOW TO HANDLE VARIABLES WHILE POOLING STUDIES?

Rename variables and flags from different sources per ADPPK IG in the pooled dataset. Baseline Weight from different studies is renamed as WTBL (Figure 1). Alternatively, imputations can also be done during modeling.

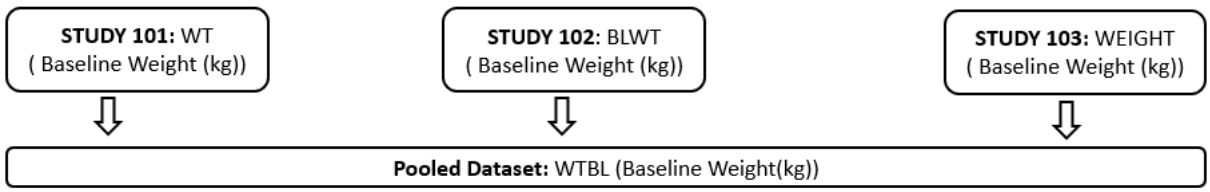

Figure 1. Example for Renaming Variables

For example, FLG from Study 101 and EXCL from Study 102 are renamed as FLGREAS (Figure 2) in the pooled dataset:

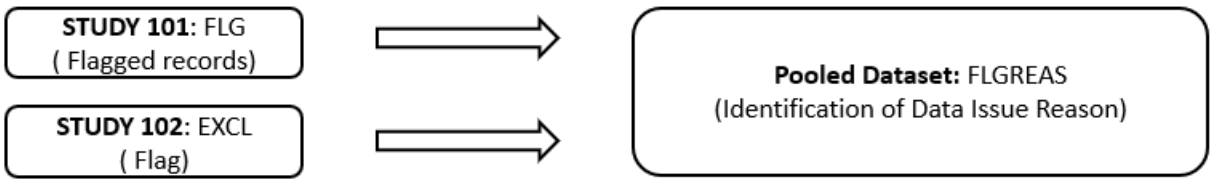

Figure 2 . Example for Renaming Flag Variable

NONMEM ONLY ALLOWS NUMERIC VARIABLES — HOW CAN THIS BE HANDLED?

Since NONMEM operates exclusively with numeric variables, it is recommended to arrange all numeric variables at the start of the dataset, followed by character variables. Character variables may be dropped or ignored in the NONMEM input statement.

The DROP statement can also be used in NONMEM:

```
$INPUT C STUDYNO ID VPCID RACEN STUDY=DROP
```

HOW TO HANDLE IMPUTATION OF VALUES WHEN ADPPK IG RECOMMENDS KEEPING THE VARIABLE FROM SOURCE UNCHANGED?

A new imputed variable is created and indicated by name to clearly distinguish between source and imputed variables. Imputation is typically performed as part of the modeling process..

Example :

USUBJID	VISIT	WT (Weight in KG)	WTI (Imputed Weight in KG)
1	Baseline	60	60
1	Cycle 1 Day 1	59	59
1	Cycle 2 Day 1		59
1	Cycle 3 Day 1	60	60

CONCLUSION

Establishing an industry-wide standard for popPK datasets that meet CDISC and modeling requirements results in key benefits:

- **Streamlined automation and tool development:** Standardized variable names and structures reduce the time spent on custom programming for each study, enabling reusable macros and automation pipelines.

- **Enhanced data quality through compliance with the ADPPK IG:** Consistent variable naming, imputation documentation, and metadata alignment reduce errors and improve traceability across submission-ready datasets.
- **Standard Identify/Exclusion record flags for analysis:** Clear, predefined flags for record inclusion/exclusion improve reproducibility, reduce reviewer burden, and ensure transparent audit trails during regulatory submissions.

Adopting the ADPPK CDISC standards requires an upfront investment in programmer training and workflow redesign; however, it positions pharmacometric teams to more efficiently meet FDA submission requirements and supports cross-study analyses that rely on harmonized data structures.

REFERENCES

CDISC. 2021. “Analysis Data Model (ADaM) Implementation Guide Version 1.3.” https://www.cdisc.org/standards/foundational/adam/adam-implementation-guide-v1-3

CDISC. 2023. "Basic Data Structure for ADaM popPK Implementation Guide v1.0." https://www.cdisc.org/standards/foundational/adam/basic-data-structure-adam-poppk-implementation-guide-v1-0

CONTACT INFORMATION

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

Praseeda Rajan
Data Science Lead, Data Science, Clinical Pharmacology and Pharmacometrics
Bristol Myers Squibb, Princeton, NJ, USA
Praseeda.Rajan@bms.com

Renuka Hegde
Associate Director, Data Science, Clinical Pharmacology and Pharmacometrics
Bristol Myers Squibb, Princeton, NJ, USA
Renuka.Hegde@bms.com

Neelima Thanneer
Executive Director, Data Science & CPAR, Clinical Pharmacology and Pharmacometrics
Bristol Myers Squibb, Princeton, NJ, USA
neelima.thanneer@bms.com

Yue Zhao
Director, PMX, Clinical Pharmacology and Pharmacometrics
Bristol Myers Squibb, Princeton, NJ, USA
Yue.Zhao@bms.com

Erin Dombrowsky
Director, Data Science, Clinical Pharmacology and Pharmacometrics
Bristol Myers Squibb, Tampa, FL, USA
Erin.Dombrowsky@bms.com