Pharmacokinetic Data Submission in the CDISC Environment

Y. Xie, P. Chai, X. Li, N. Wang

Clinical Pharmacology Sciences, Celerion

OVERVIEW

Using SDTM datasets as a starting point, this poster shares an approach for generating ADaM datasets to support Pharmacokinetic (PK) analysis that integrates production of define.xml data documentation at each step in an auditable process and shows how it can streamline data preparation and reconciliation, table and figure generation, and submission of PK results to regulatory agencies in a controlled, regulated environment.

INTRODUCTION

FDA has encouraged the use of sponsor-independent Clinical Data Interchange Standards Consortium (CDISC) standards to present traceable, transparent, and comprehensive clinical trial data. The PK analysis processes incorporating CDISC standards are presented.

METHODS

The key goal is to generate electronic submission-ready analysis data (supported by data documentation) in a standardized format. These data are then used as the input for table and figure generation. Study Data Tabulation Model (SDTM) standards are designed to support individual subject listings with a complete set of study data. Analysis Data Model (ADaM) standards are designed to complement the SDTM data by providing "analysis ready" datasets that facilitate table and figure generation and review. Both SDTM and ADaM data must be supported by data documentation. The machine-readable define.xml supported by the Case Report Tabulation Data Definition Specification (CRT-DDS) is the best way to document SDTM and ADaM datasets, and define.xml generation can be integrated with the data mapping process. Core processes include generating input metadata based on CDISC standards to support the data mapping and define.xml generation, inclusion of treatment and common variables, facilitating exclusion of selected data points with flags, incorporating non-compartmental PK parameters from WinNonlin and other derived content, and reviewing the data for substantive and structural accuracy.

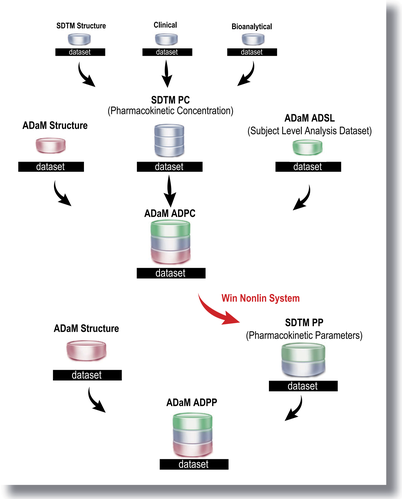
Creating Structure datasets for SDTM and ADaM

Celerion standard structure datasets for SDTM and ADaM include four parts: general study information, dataset metadata, variable metadata, and results metadata. General study information contains protocol number, study number, study description, SAS version, SDTM/ADaM version. Dataset metadata contain all possible dataset names and dataset labels. Variable metadata contain possible variable names, variable labels, variable origin, and etc. Results metadata contain result formats, result length, and etc. The screenshot of the variables in structure datasets are as follow (structure datasets for SDTM and ADaM are similar):

Structure dataset

Process to Create SDTM and ADaM Datasets

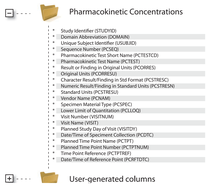
Celerion, Inc. standard procedure to create ADaM datasets for PK analysis is clear, stable, and efficient. First, SDTM PC dataset is created based on SDTM structure dataset, clinical datasets, and bioanalytical dataset. Second, based on the ADaM structure dataset, SDTM PC dataset and ADaM ADSL(Subject Level Analysis Dataset) are merged to create ADaM ADPC dataset. ADaM ADPC dataset supports PK parameters calculation. It also provides information to create PK concentration tables and figures. Third, using WinNonlin, PKist calculates the PK parameters from ADaM ADPC. WinNonlin creates a dataset including all these calculated PK parameter information, and SDTM PP dataset is created based on this dataset. Fourth, ADaM ADPP dataset is created based ADaM structure dataset and SDTM PP. ADPP dataset is the PK analysis dataset which is used for producing summary tables, statistical tables, and any other PK analysis.



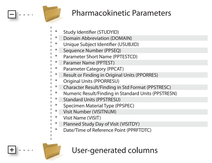
Creating SDTM datasets (PC and PP)

PC dataset is based on the original bioanalytical datasets, clinical datasets, and the standard structure dataset. Clinical datasets provide information for medication, PK sampling, etc. PP dataset is based on the PK parameter datasets from the WinNonlin software. The screenshot of the variables in those datasets are as follow:

PC Domain



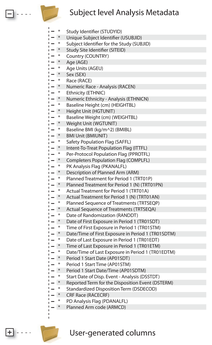
PP Domain



Creating ADaM datasets (ADPC and ADPP)

ADPC and ADPP datasets are based on the SDTM datasets (PC and PP) merged with Subject Level Analysis Data (ADSL) which has one record for each subject including the subject identifiers, demographic information, dosing information, and flags indicating whether the subject will be included in the analysis. Those datasets are "analysis ready" datasets that facilitate table and figure generation and review.

ADSL dataset



ADPC dataset (ADPP dataset is similar to ADPC dataset)

Two-step Data-driven Define.xml Creation Using Proc Template

Well formed data (SDTM and ADaM) contain all information needed to understand study results, but that information is typically not easily accessible to end users. Combining the data with xml format data documentation gives end-users a data and documentation package that is complete, user friendly, and ready to submit to the FDA. Define.xml data documentation has internal and external links that allow end users to quickly find the level of information they need, from source CRF references to variable lists to code lists and comments associated with key content.

A two-step data-driven process is used to create the define.xml document. The first step is the production of six SAS® format define data sets by mining the data for unique information. The six define data sets include six levels of study information which include study level, domain level, variable level, value level, supporting document and code list. The second step is integration of the SAS® define data into an xml formatted document through a robust approach driven by SAS® PROC TEMPLATE. Sorted define data sets are restructured and transformed into a group of tagsets that comprise the define document. Below are the screenshots of an example for SDTM define.xml.

Domain Level Define Date Created by the First Step

#	Dataset N...	Dataset L...	Observati...	Repeating...	Purpose o...	Is Referen...	Domain D...	Sort Key	Dataset lo...	Sort Key f...
1	DM	Demograp...	Special Pur...	No	Tabulation	No	One record...	STUDYID ...	dm.xpt	1
2	CO	Comments	Special Pur...	Yes	Tabulation	No	One record...	STUDYID ...	co.xpt	2
3	SE	Subject Ele...	Special Pur...	Yes	Tabulation	No	One record...	STUDYID ...	se.xpt	3
4	SV	Subject Vis...	Special Pur...	Yes	Tabulation	No	One record...	STUDYID ...	sv.xpt	4
5	CM	Concomita...	Interventions	Yes	Tabulation	No	One record...	STUDYID ...	cm.xpt	5
6	EX	Exposure	Interventions	Yes	Tabulation	No	One record...	STUDYID ...	ex.xpt	6
7	SU	Substance...	Interventions	No	Tabulation	No	One record...	STUDYID ...	su.xpt	7
8	AE	Adverse Ev...	Events	Yes	Tabulation	No	One record...	STUDYID ...	ae.xpt	8
9	DS	Disposition	Events	Yes	Tabulation	No	One record...	STUDYID ...	ds.xpt	9
10	MH	Medical Hi...	Events	Yes	Tabulation	No	One record...	STUDYID ...	mh.xpt	10
11	DV	Protocol D...	Findings	Yes	Tabulation	No	One record...	STUDYID ...	dv.xpt	11
12	EG	ECG Test ...	Findings	Yes	Tabulation	No	One record...	STUDYID ...	eg.xpt	12
13	IE	IncusionEx...	Findings	Yes	Tabulation	No	One record...	STUDYID ...	ie.xpt	13
14	LB	Laboratory...	Findings	Yes	Tabulation	No	One record...	STUDYID ...	lb.xpt	14
15	PE	Physical Ex...	Findings	Yes	Tabulation	No	One record...	STUDYID ...	pe.xpt	15
16	QS	Questionn...	Findings	Yes	Tabulation	No	One record...	STUDYID ...	qs.xpt	16
17	VS	Vital Signs	Findings	Yes	Tabulation	No	One record...	STUDYID ...	vs.xpt	17
18	PC	Pharmacok...	Findings	Yes	Tabulation	No	One record...	STUDYID ...	pc.xpt	18
19	PP	Pharmacok...	Findings	Yes	Tabulation	No	One record...	STUDYID ...	pp.xpt	19
20	TA	Trial Arms	Trial Design	No	Tabulation	Yes	One record...	STUDYID A...	ta.xpt	20

Domain Level Define.xml Created by the Second Step

Dataset	Description	Class	Structure	Purpose	Keys	Location
DM	Demographics	Special Purpose	One record per subject	Tabulation	STUDYID USUBJID	dm.xpt
CO	Comments	Special Purpose	One record per comment per subject	Tabulation	STUDYID USUBJID IDVAR IDVARVAL	co.xpt
SE	Subject Elements	Special Purpose	One record per actual Element per subject	Tabulation	STUDYID USUBJID SESTDTC	se.xpt
SV	Subject Visits	Special Purpose	One record per actual visit per subject	Tabulation	STUDYID USUBJID VISITNUM	sv.xpt
CM	Concomitant Medications	Interventions	One record per recorded medication occurrence or constant-dosing interval per subject	Tabulation	STUDYID USUBJID CMSTDTC CMTRT	cm.xpt
EX	Exposure	Interventions	One record per constant dosing interval per subject	Tabulation	STUDYID USUBJID EXSTDTC	ex.xpt
SU	Substance Use	Interventions	One record per substance type per reported occurrence per subject	Tabulation	STUDYID USUBJID SUSTDTC	su.xpt
AE	Adverse Events	Events	One record per adverse event per subject	Tabulation	STUDYID USUBJID AESTDTC AETERM	ae.xpt
DS	Disposition	Events	One record per disposition status or protocol milestone per subject	Tabulation	STUDYID USUBJID DSSTDTC	ds.xpt

RESULTS

The current process overview emphasizes how the use of input metadata files (used to automatically set domain and variable names and labels and to reference source information) and integrated review tools allows scientists to more efficiently direct and review ADaM data generation. The use of SAS driver macros and modular standard programs further enhances automation while allowing flexibility for study design.

CONCLUSIONS

Building robust upstream data processes around CDISC dataset specifications and standards can accelerate PK analysis and reporting for efficient agency review.

ACKNOWLEDGEMENTS

The authors thank Celerion for the support received for this work and Steven Kirby for his expert leadership on SDTM, ADaM, and Define.xml.

www.celerion.com