

NONMEM® – A Programmer point of view

Raghu Kishor Koukuntla, Cytel, Hyderabad, India

ABSTRACT:

Non Linear Mixed Effects Modeling software called as NONMEM® is a leading tool used for population PK/PD analysis. It is widely used to predict the effect of drug on the target patient population. The NONMEM® software requires data in a very specific pre-defined format and includes a mixture of data components like dosing records, PK and/or PD observations, time independent covariates, time dependent covariates, lab covariates and may include covariates such as concomitant therapies.

This paper will provide clear understanding of a NONMEM input dataset and helps as a reference document for NONMEM programming concepts.

1. INTRODUCTION

NONMEM is a software package. It is a computer program, written in FORTRAN 90/95, designed for analyzing population Pharmacokinetic/Pharmacodynamic (PK/PD) data in particular. It was developed at the University of California at San Francisco by two professors, Prof. Lewis Sheiner and Prof. Stuart Beal. NONMEM 7 is the current version of the software and is the property of ICON Development Solutions.

NONMEM is developed for analyzing population pharmacokinetic data in particular. Such data is typically collected from clinical studies of active pharmaceutical ingredients, involving the administration of a drug to individuals and the subsequent observation of drug levels (most often in the blood plasma). Proper modeling of these data involves accounting for both unexplainable inter-subject and intra-subject effects (random effects), as well as measured concomitant effects (fixed effects). NONMEM allows this mixed effect modeling. Such modeling is especially useful when there are only a few pharmacokinetic measurements from each individual sampled in the population, or when the data collection design varies considerably between these individuals. However, NONMEM is a general program which can be used to fit models to a wide variety of data.

To perform population PK/PD analysis using NONMEM software, it requires an input dataset in a specific pre-defined format. This input dataset includes variety of variables (Covariates) and observations, which will be discussed later in this paper.

2. BASIC DEFINITIONS - PK, PD AND POPULATION PK/PD

Understanding the NONMEM data structure needs understanding of some basic definitions used in this context.

2.1 Pharmacokinetics (PK)

PK is a relationship between the dose administered and the changes in the drug concentration in the body with time. It's measured at 4 different phases by drug concentration in blood, plasma, tissue. Four Phases of PK are: Absorption, Distribution, Metabolism, and Excretion (ADME).

2.2 Pharmacodynamics (PD)

The study of the body's pharmacological response to a drug (measured in terms of AEs & Efficacy).

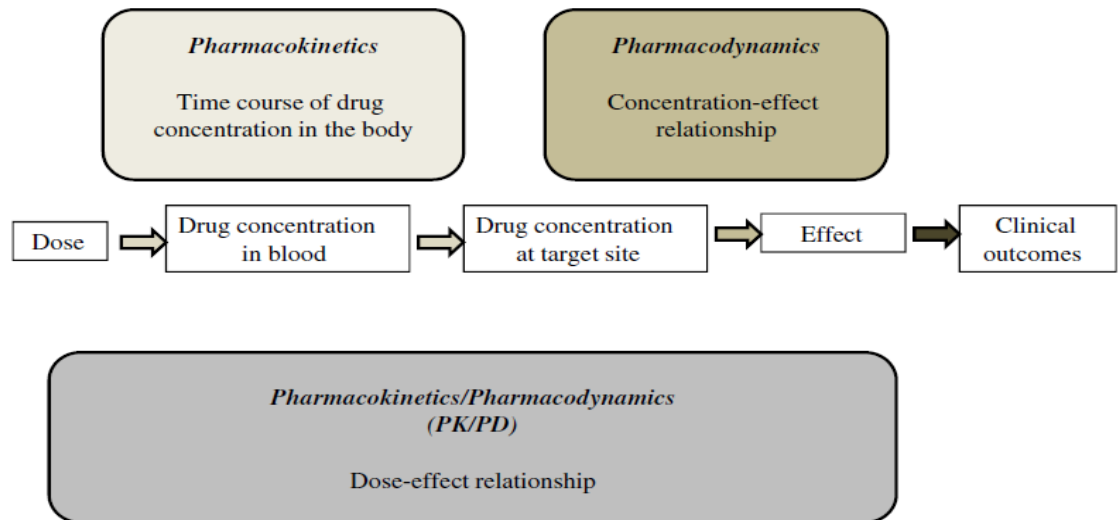


Fig. 1. The relationship between pharmacokinetics (PK) and pharmacodynamics (PD).

2.3 Population PK/PD analysis

Mathematical models describe the variability of concentration and response between individuals considered at the same time.

2.4 Individual PK vs. Population PK

Pharmacokinetic analysis is often broken into two areas "PK" and "population PK". Sometimes the first term "PK" is also called "individual PK". Individual PK and population PK are simply two different analysis techniques that we use in pharmacokinetic data analysis.

Individual PK usually has the following features:

Individual PK –

- Use graphical techniques
- Used for Phase I studies
- Not useful for predictive work
- Requires intensive sampling
- Analysis time is short
- Can determine
 - Maximum exposure (C_{max})
 - Total exposure (Area under the curve or AUC)
 - Clearance
 - Bioavailability
 - Volume of distribution
 - Terminal half-life ($t_{1/2}$)

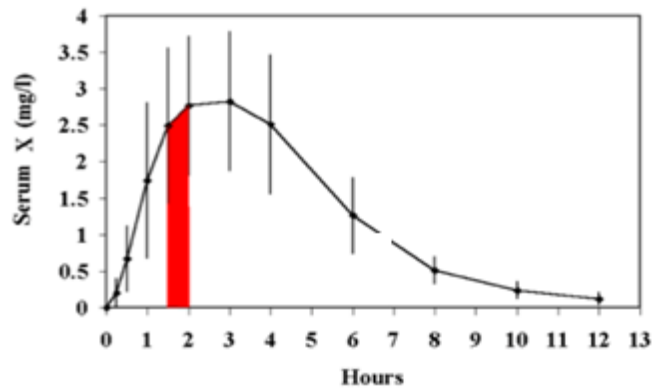


Fig. 2

However to understand how individuals from a population differ from one another, we need to perform population PK analysis. In this type of analysis, all data from all individuals is considered at the same time in a unified model.

Population PK analysis has the following features:

Population PK -

- Evaluates entire population
- Can be used for predictions and simulations
- Computationally intensive
- Intensive and sparse sampling
- Analysis time is longer
- PK/PD modeling (relationship between drug levels and drug effects)
- Can determine
 - Clearance
 - Volume of distribution
 - Effect of covariates (e.g. age, weight, sex, kidney function)

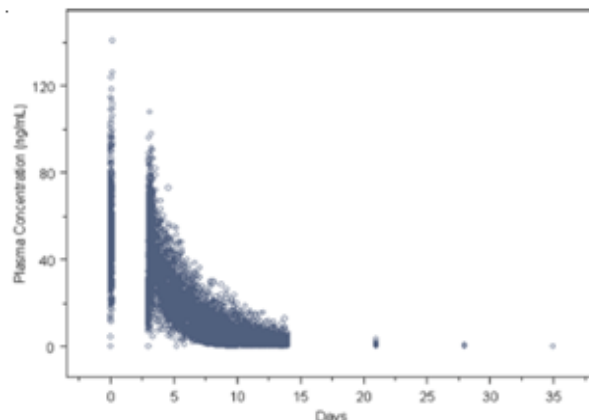
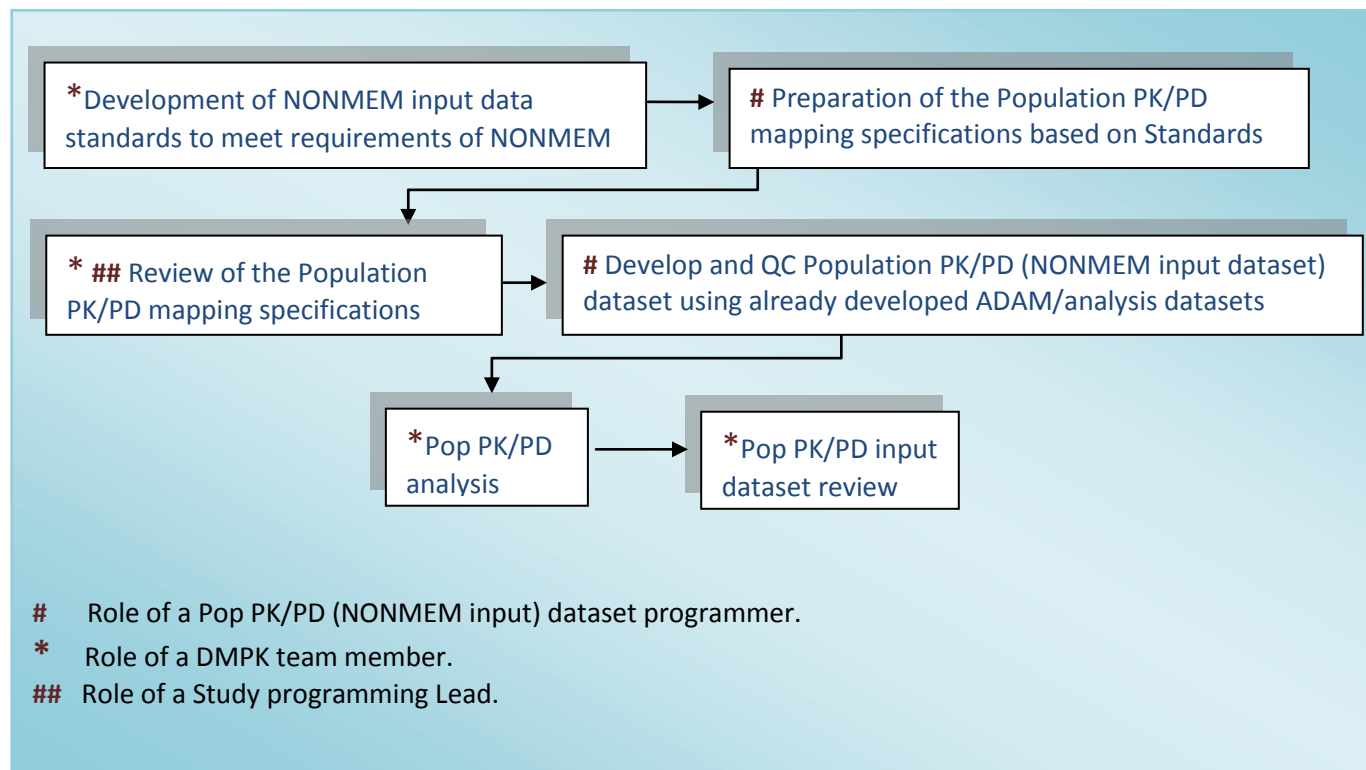


Fig. 3

3. NONMEM INPUT DATASET DEVELOPMENT WORK FLOW

Clinical Study team Lead, Drug Metabolism and Pharmacokinetics group lead and Lead Statistician together makes a decision whether to perform Population PK/PD analysis or not, if yes, it triggers the following work flow.

The general work flow in developing the input dataset to NONMEM



PhUSE 2014
NONMEM Software Input dataset (Pop PK/PD) development flowchart

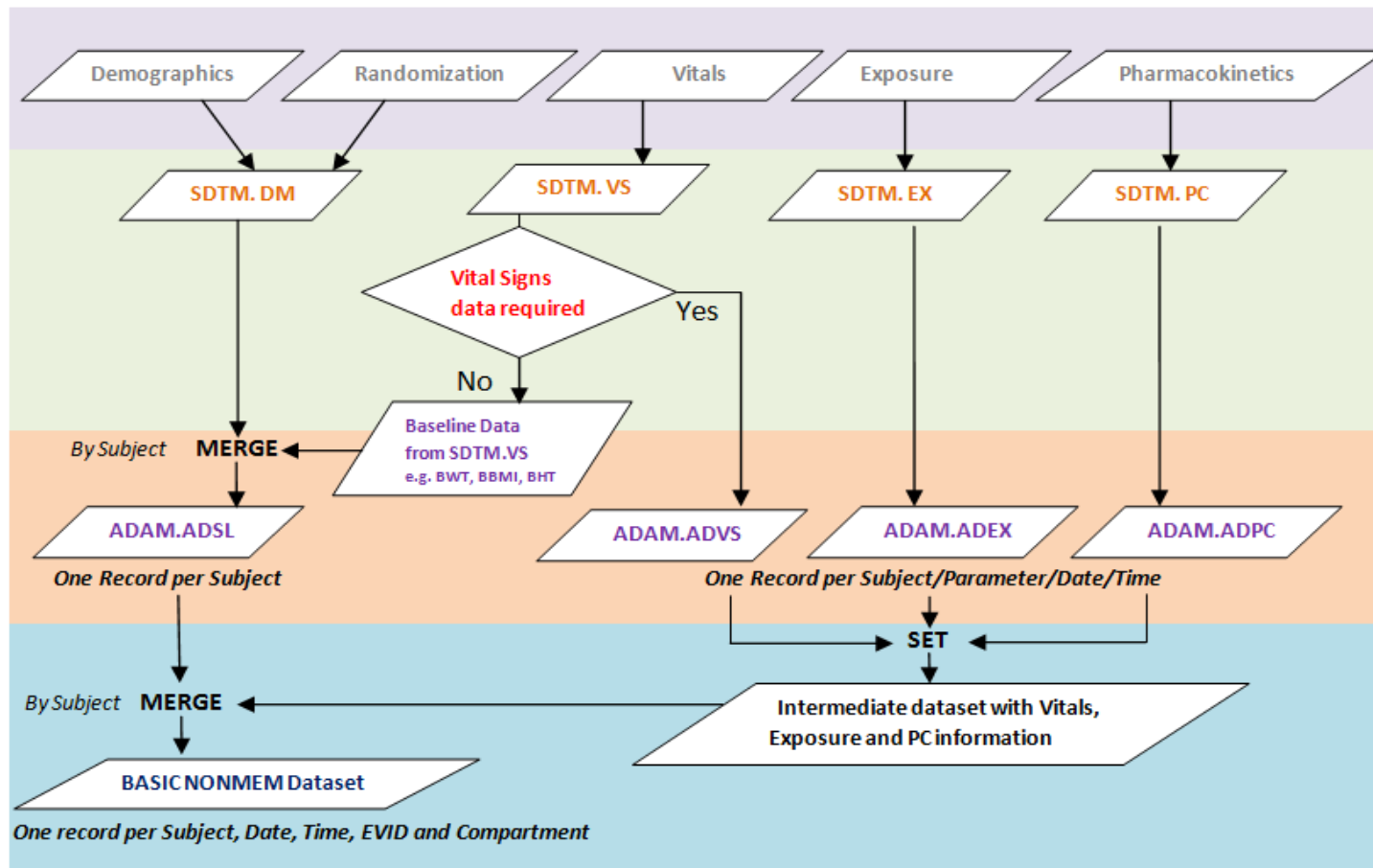


Fig. 4. A flow chart for Basic POP PK/PD dataset development

Based on Study design and analysis, PD baseline variables /Covariates (e.g. Lab baseline data, Biomarker baseline data) and CONMED variables may be added to Basic NONMEM dataset.

The program flow for adding additional required covariates and data records is shown in next page.

PhUSE 2014

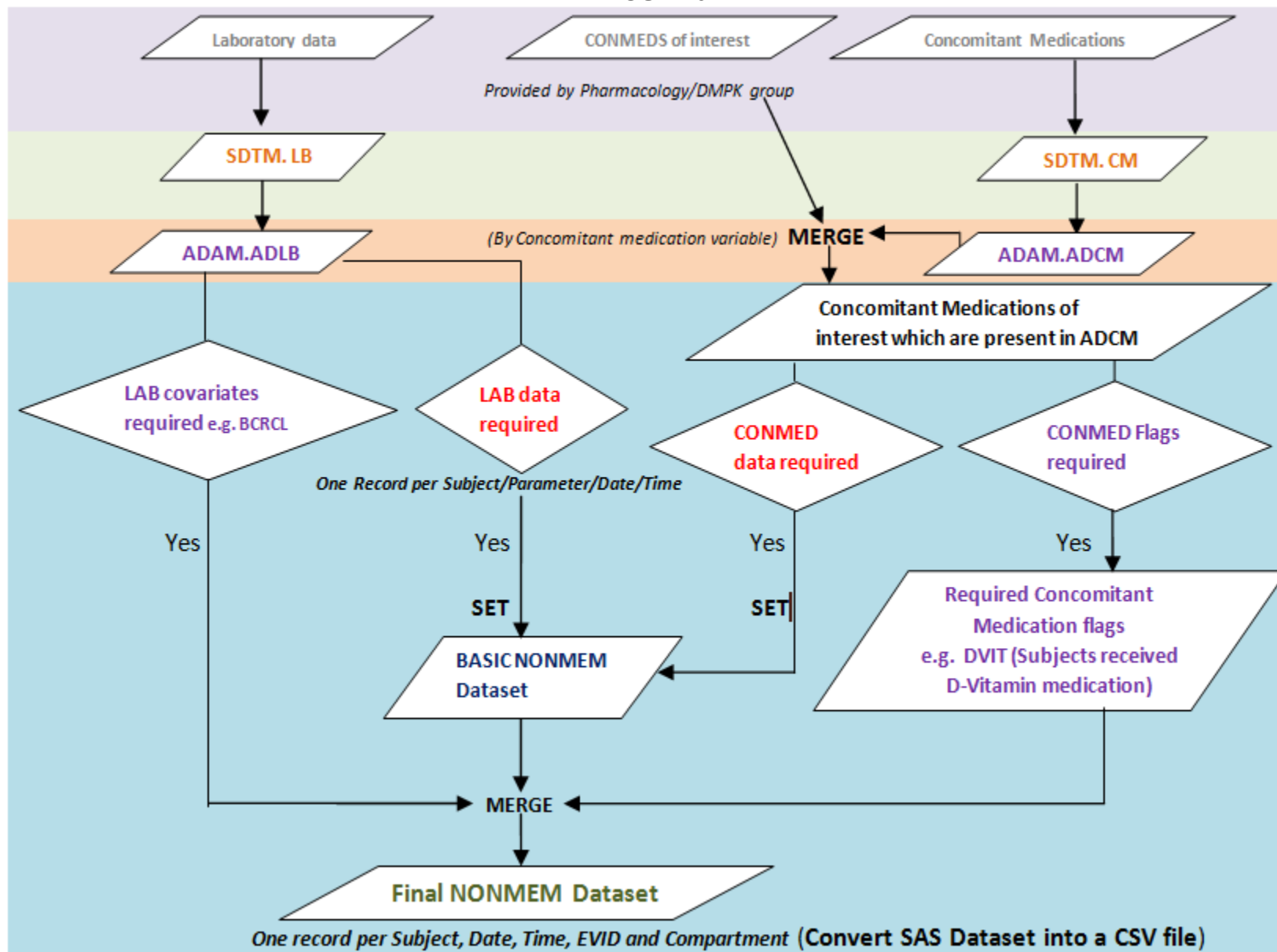


Fig. 5. A flow chart for POP PK/PD dataset development with additional covariates and data records

4. NONMEM INPUT DATASET STRUCTURE

The NONMEM® input dataset (here after referred as Pop PK/PD dataset) is an ASCII file typically has one row (i.e. record) per study, subject, date, time, event ID, and compartment flag. It consists of three basic components; the dosing records, observations i.e. the PK/PD results, and covariates (PD baseline and/or Demographics covariates). All data must be correct and consistent in order to create a good quality Pop PK/PD dataset.

The Pop PK/PD dataset requirements may differ based on the analysis

4.1 Population PK data requirements

PK data requirements will typically include the following:

- Dosing information including planned and actual dose with actual and nominal times of dosing and route of administration
- PK data: serum or plasma Investigational Product concentrations, metabolite concentrations (if appropriate) and target concentrations (if available) with actual and nominal times of measurement and an indicator if measures are BQL
- Baseline demographic covariates including body weight, BMI, age, race, gender, and disease duration
- Baseline PD measures (biomarker and/or efficacy) may also be appropriate

4.2 Population PK/PD data requirements

PK/PD data requirements will include the above as well as:

- Baseline PD measures (biomarker or efficacy)
- PD (biomarker or efficacy) measures with actual and nominal times of measurement and an indicator if measures are BQL

5. DIFFERENT CATEGORIES OF VARIABLES

5.1 Variable type categorization

- Common
- TA-Specific
- Study Specific

5.2 Need based categorization

- Required/Core
- Optional

Variable type categorization must be used together with Requirement based categorization to determine whether a variable is required (Core) or optional in a pop PK/PD data.

All Study Specific Variables are considered as required variables.

PhUSE 2014

Overall Pop PK/PD input data variables are categorized into the following five categories

Common	Required (Core)
	Optional
Therapeutic Area Specific	Required (Core)
	Optional
Study Specific	Required (Core)

Overall Categorization	Description	Examples
Common and core	Variables in this category should be included in all pop PK/PD input data files.	Subject ID, study ID, date, time, and actual elapsed time from first dose of treatment.
Common and optional	Variables in this category may be included in any pop PK/PD input data file depending on analysis needs.	Dose administration, fasted condition, planned visit code, Baseline Height, Week, BMI
TA-specific (e.g. Asthma, Oncology, Nephrology) and core	Variables in this category should be included in all pop PK/PD input data files for a particular TA	Diabetes specific: Compartment flag, stable Metformin dose, and diabetes duration Nephrology specific: Time of dialysis Oncology specific: Progression free survival in days
TA-specific and optional	Variables in this category may be included in any pop PK/PD input data file of a particular TA. Their inclusion depends on analysis needs.	Oncology Specific: Time varying weight, time varying creatinine, baseline GOG performance score.
Study-Specific	These variables are included in pop PK/PD input data which are Study specific (Particular analysis in a Protocol).	Reason for Dose Change

NONMEM input (Population PK/PD) dataset examples are shown in next page.

PhUSE 2014

Basic NONMEM Dataset Example

Data Type	PTID	DATE	TIME	DV	MDV	UNIT	AMT	EVID	CMT	VIS	BAGE	AGE	BWGT	WGT	BCRCL	BBMI	BMI
LAB PD	1001	10/25/2011	15:21	6.833	0	pmol/L		0	6	SCREENING	53.1	53.1	84.8	83.8	79.718	33.1	32.7
LAB PD	1001	11/8/2011	6:10	9.2	0	mg/dL		0	10	DAY -1	53.1	53.1	84.8	84.8	79.718	33.1	33.1
PK	1001	11/9/2011	5:10	50	0	ng/mL		0	1	DAY 1 - D1PRE	53.1	53.1	84.8	84.8	79.718	33.1	33.1
LAB PD	1001	11/9/2011	5:10	2.844	0	pmol/L		0	6	DAY 1 PREDOSE	53.1	53.1	84.8	84.8	79.718	33.1	33.1
LAB PD	1001	11/9/2011	5:10	9.6	0	mg/dL		0	10	DAY 1 PREDOSE	53.1	53.1	84.8	84.8	79.718	33.1	33.1
Dosing	1001	11/9/2011	6:09		1	.	210	1	3	DAY 1	53.1	53.1	84.8	84.8	79.718	33.1	33.1
PK	1001	11/10/2011	6:09	5420	0	ng/mL		0	1	DAY 2 - POST 24 HOURS	53.1	53.1	84.8	84.8	79.718	33.1	33.1
PK	1001	11/11/2011	6:09	8740	0	ng/mL		0	1	DAY 3 - POST 48 HOURS	53.1	53.1	84.8	84.8	79.718	33.1	33.1
LAB PD	1001	11/14/2011	6:09	4.987	0	pmol/L		0	6	DAY 6	53.1	53.1	84.8	84.8	79.718	33.1	33.1
LAB PD	1001	11/14/2011	6:09	9	0	mg/dL		0	10	DAY 6	53.1	53.1	84.8	84.8	79.718	33.1	33.1
LAB PD	1002	10/25/2011	15:24	8.9	0	mg/dL		0	10	SCREENING	58.3	58.2	79.4	78.8	60.199	26.2	26
LAB PD	1002	11/8/2011	6:05	3.3	0	mg/dL		0	13	DAY -1	58.3	58.3	79.4	79.4	60.199	26.2	26.2
PK	1002	11/9/2011	5:06	50	0	ng/mL		0	1	DAY 1 - D1PRE	58.3	58.3	79.4	79.4	60.199	26.2	26.2
LAB PD	1002	11/9/2011	5:06	8.7	0	mg/dL		0	10	DAY 1 PREDOSE	58.3	58.3	79.4	79.4	60.199	26.2	26.2
LAB PD	1002	11/9/2011	5:06	3.6		mg/dL		0	13	DAY 1 PREDOSE	58.3	58.3	79.4	79.4	60.199	26.2	26.2
Dosing	1002	11/9/2011	6:05		1	.	210	1	3	DAY 1	58.3	58.3	79.4	79.4	60.199	26.2	26.2
PK	1002	11/10/2011	6:05	7450	0	ng/mL		0	1	DAY 2 - POST 24 HOURS	58.3	58.3	79.4	79.4	60.199	26.2	26.2
PK	1002	11/11/2011	6:05	17300	0	ng/mL		0	1	DAY 3 - POST 48 HOURS	58.3	58.3	79.4	79.4	60.199	26.2	26.2
LAB PD	1002	11/14/2011	6:05	8.4	0	mg/dL		0	10	DAY 6	58.3	58.3	79.4	79.4	60.199	26.2	26.2
LAB PD	1002	11/14/2011	6:05	2.9	0	mg/dL		0	13	DAY 6	58.3	58.3	79.4	79.4	60.199	26.2	26.2

Example of NONMEM dataset with multiple covariates and additional observations

Time Independent Covariates Time dependent covariates Conmed variables																																	
Data Type	C	NMID	PTID	DATE	TIME	TAFD	STAFD	TALD	EVID	CMT	AMT	DV	MDV	BQL	TRTP	TRTA	VISIT	NTAFD	AGE	SEX	RACE	BWGT	BHGT	BBUN	BBMI	BBSA	BSBP	BDBP	AMTU	DVU	DVIT	PHOSBND	CONMED
CONMED	.	5	1001	2/11/2010	.	.	.	.	1	26	.	.	1	.	2	2	.	.	50	1	5	83.2	178	29.274	33.328	1.911	120	75	.	.	1	1	CALCIOTRIOL
CONMED	1	5	1001	2/10/2014	.	.	.	.	1	27	.	.	1	.	2	2	.	.	50	1	5	83.2	178	29.274	33.328	1.862	120	75	.	.	1	1	CALCIUM GLUCONATE
LAB PD	.	5	1001	8/28/2013	8:00	-341.5	36552.5	-341.5	0	5	.	86.681	0	0	2	2	SCR	-1344	50	1	5	83.2	178	29.274	33.328	1.911	120	75	.	pmol/L	1	1	.
VITAL	.	5	1001	8/28/2013	8:00	-341.5	36552.5	-341.5	0	19	.	130	0	0	2	2	SCR	-1344	50	1	5	83.2	178	29.274	33.328	1.911	120	75	.	mmHg	1	1	.
LAB PD	.	5	1001	9/2/2013	7:30	-222	36672	-222	0	6	.	2.075	0	0	2	2	SCR	-1344	50	1	5	83.2	178	29.274	33.328	1.911	120	75	.	mmol/L	1	1	.
PK	.	5	1001	9/11/2013	7:52	-5.6333	36888.3667	-5.6333	0	1	.	.	1	1	2	2	D1	-4	50	1	5	83.2	178	29.274	33.328	1.911	120	75	.	ng/mL	1	1	.
LAB PD	.	5	1001	9/11/2013	7:57	-5.55	36888.45	-5.55	0	5	.	88.772	0	0	2	2	D1	-4	50	1	5	83.2	178	29.274	33.328	1.911	120	75	.	pmol/L	1	1	.
VITAL	.	5	1001	9/11/2013	8:00	-5.5	36888.5	-5.5	0	20	.	75	0	0	2	2	D1	-4	50	1	5	83.2	178	29.274	33.328	1.911	120	75	.	mmHg	1	1	.
DOSING	.	5	1001	9/11/2013	13:30	0	36894	0	1	0	5	.	1	.	2	2	W1	-4	50	1	5	83.2	178	29.274	33.328	1.911	120	75	MG	.	1	1	.
PK	.	5	1001	10/2/2013	7:40	498.1667	37392.1667	42.3333	0	1	.	30.8	0	0	2	2	W4	500	50	1	5	83.2	178	29.274	33.328	1.915	120	75	.	ng/mL	1	1	.
LAB PD	.	5	1001	10/2/2013	7:40	498.1667	37392.1667	42.3333	0	5	.	72.441	0	0	2	2	W4	500	50	1	5	83.2	178	29.274	33.328	1.915	120	75	.	pmol/L	1	1	.
VITAL	.	5	1001	10/2/2013	8:00	498.5	37392.5	42.6667	0	19	.	109	0	0	2	2	W4	500	50	1	5	83.2	178	29.274	33.328	1.915	120	75	.	mmHg	1	1	.
DOSING	.	5	1001	10/2/2013	13:10	503.6667	37397.6667	0	1	0	5	.	1	.	2	2	W4	500	50	1	5	83.2	178	29.274	33.328	1.915	120	75	MG	.	1	1	.

6. NONMEM SPECIFIC REQUIRED VARIABLES

6.1 ID - Identifier Data Item

When the data is from a population, NONMEM expects the Identification data item, labeled ID, and expects the data to be organized into two or more "individual records". An individual record is a group of contiguous data records having the same value for the ID data item and presumably containing data from the same individual. ID data item values need not be consecutive, increasing, unique, nor begin with 1. E.g., 3, 5, 6, 10, 3, etc. is a possible sequence of ID values. Note the two instances of 3 as ID data item values. As long as these two instances are separated by different ID data item values (e.g. 5, 6, 10), they represent different individuals.

6.2 EVID - Event ID

EVID: EVID is a NONMEM specific variable. Every record in a NONMEM data set must contain an "Event Identification Data Item" identifying the kind of event described by the record, and labeled as EVID. Two categories of EVID are available: Observation (EVID=0) and Dosing (EVID=1).

6.3 DV-Dependent Variable

This is a value of an observation. There can be only one DV data item per data record. The position of the DV data item is not important. However, its position must be the same on all records.

Dosing Records: Enter "."

Missing or BQL Observation Records: Enter "." when a result is missing or the value that was reported was below the assay limit of quantitation

Other Observation Records: Enter numeric result of the sample or assessment. A result is reported in the units specified for a given observation record in CMT field.

6.4 MDV-Missing Dependent Variable

If there are records in an input data set which do not contain values of observations, then NONMEM needs to be informed of this fact. This is done using the Missing Dependent Variable data item labeled MDV.

The values of MDV are:

0 - The DV data item of the data record contains a value of an observation.

The record is referred to as an observation record.

1 - The DV data item of the data record does not contain a value of an observation.

E.g., Dosing Records: Enter "1"

Missing or BQL Observation Records: Enter "1"

6.5 AMT- Amount Dosing Records:

The Amount data item (AMT) gives the amount of a bolus dose or of an infusion of finite duration. This amount should be a positive number.

Enter actual total dose amount administered to a subject in units specified for a given dosing record in CMT field.

Other Records: Enter "."

6.6 RATE - RATE data item

The Rate data item (RATE) gives the rate of an infusion. This rate should be a positive number.

PhUSE 2014

6.7 CMT-Compartment data item

CMT is used to report the record contents that appear in DV and AMT for Observation and Dosing records, respectively. Different compartment value is assigned for different PD/Biomarker records.

For observation records: CMT specifies the compartment from which the predicted value of the observation is obtained.

For dosing records: CMT specifies the compartment into which the dose is introduced.

6.8 ADDL – Additional Dosing indicator

For Multiple Dosing studies, we use this variable to indicate number of additional doses. For example, a QD for 2 weeks would be 13 since there would be once a day dosing for 14 days.

6.9 II – Interdose Interval data item

The Interdose Interval data item labeled II gives the time between implied doses. For a steady-state infusion, it should be 0. For other steady state doses, it should be a positive number whose units are the same as the TIME data item. This item is required if ADDL is used to define additional implied doses. For example, the number of hours between dosing- 24 for QD, 12 for BiD, 168 for once weekly

6.10 SS - Steady State Indicator

SS is a NONMEM specific variable which represents steady state data item. Use of this data items implies that the dose and regimen combination is at steady-state, and is thus only appropriate when those conditions are satisfied.

SS is set to 1 when indicated, otherwise 0.

7. TIME RELATED NONMEM SPECIFIC VARIABLES

7.1 DATE - Date data item

The Date data item (DATE) is used to record calendar dates in month-day-year format. (e.g., 9 /1/2014).

7.2 TIME - Time data item

Time data item provides the clock time (usually in hours) as hh:mm. (e.g., the time 2:00 PM is recorded as 14:00).

7.3 TAFD - Actual Elapsed Time from 1st Dose

It represents nominal time after first dose for a subject. Observation records for a subject prior to the first dose (e.g. screening records) are reported as negative numbers. For plotting purposes, this behavior is adequate. However, software tools for regression (e.g. NONMEM) may not accept negative time values. STAFD is used to address this issue.

7.4 STAFD - Shifted Act Elapsed Time from 1st Dose

STAFD variable records the shifted actual elapsed time from first dose of treatment, and is equivalent to TAFD if the first Observation record for a subject occurs after the first dose. If Observation records occur before the first dose in any subject (e.g. screening values are requested), then all Observation and Dosing records within the study data set are positively shifted to ensure that all STAFD values are positive. E.g. if the earliest screening Observation record in the study occurred 14 days before the first dose, then STAFD is TAFD+14 days, assuming a unit of days per time. This will ensure that the value of STAFD is always non-negative.

PhUSE 2014

7.5 TALD - Actual Elapsed Time from Previous Dose

Represents actual elapsed time from most recent previous dose of treatment, and may be useful for purposes of plotting, doing visual predictive checks etc. Additionally, these times are useful for NCA analysis of PK data.

For Single dosing studies, TAFD = TALD.

7.6 NTAFD - Nominal Time after First Dose

Represents nominal time after first dose, and may be useful for purposes of plotting, doing visual predictive checks etc. Observation records prior to the first dose will be reported as negative.

8. OTHER SUPPORTING VARIABLES IN NONMEM INPUT DATASET

8.1 C-Comments

This variable is used to flag a record with specific comment code with respect to any specific variable/s value. Whenever no comment code is applicable to a NONMEM record then C is set to missing ".". If multiple comment codes are applicable to the same NONMEM record, then Comment codes are sorted in ascending order and separated them by dot.

For example, if codes 1, 2, and 9 are applicable to a given NONMEM record then C = "1.2.9"

9. HANDLING MISSING DATA

Missing dosing records, observation records and covariate information may lead to issues while analyzing and so there is need to have conventions and rules to handle missing data.

The BQL and MDV fields are available to flag the records as missing and to provide for the possibility that the Observation record is missing due to assay qualities. No specific fields are available to address missing Dosing records.

- MDV: Discussed above in section 6.3
- BQL: BQL is a special case of a missing Observation record, and represents observation values that are below the limit of quantitation. That is, the measurement was made, but the result was below the assays lower limit of quantitation. Modern advances in estimation techniques allow this record to play a role in parameter estimation.

Imputation of Observation or Dosing records is an area of active research, and specifying a single set of imputation rules that would satisfy every imputation scenario for all possible study-specific situations is not possible. Imputation rules are defined by Drug Metabolism and Pharmacokinetics (DMPK) team.

9.1 Imputation of Observation Records:

Imputation of observation records may be needed to make sure there is no information missing related to time etc. If any of the observation related information is missing, then appropriate imputation rules will be provided in study specific document.

9.2 Imputation of Dosing Records:

Imputation of dosing records may be needed to make sure there is no information missing related to dosing time and amount. If any of the dosing information is missing, then appropriate imputation rules will be provided in study specific document.

Based on need, different imputation methods are used.

- LOCF - Last Observation Carry Forward
- WOCF - Worst Observation Carry Forward
- FOCB - Forward Observation carry backward
- COA - Closest Observation Algorithm

10. NONMEM DATASET EXCLUSIONS

This information is provided by Drug Metabolism and Pharmacokinetics team as required for Population PK/PD analysis. Generally all dose missing records and records with derived variables are excluded at preprocessing level. As per analysis requirements mostly the Screen failure records, date missing records and duplicate records on Key variables listed by DMPK team are excluded.

11. CONCLUSION

The role of the Pop PK/PD dataset programmer is not only just following a set of rules and data set specifications, but also understanding the dataset development work flow, gaining thorough knowledge on the structure and requirements of a NONMEM dataset. This paper will act as a reference document for all naive programmers developing NONMEM input dataset.

REFERENCES

1. Dennis Fisher, Steven Shafer. Basic Concepts- Pharmacokinetic and Pharmacodynamic Analysis with NONMEM
2. NONMEM User Guide by Alison J. Boeckmann, Lewis B. Sheiner, Stuart L. Beal
3. http://shaficonsultancy.com/papers/Producing_NONMEM_Datasets_Poster.pdf

ACKNOWLEDGMENTS

I would like to thank all Cytel colleagues who provided their valuable input for this paper.

I would like to acknowledge Shafi Chowdhury, et al., poster on “Producing NONMEM dataset using a standard SAS® program” presented at PAGE meeting which helped me as a reference material.

CONTACT INFORMATION

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

Raghu Kishor Koukuntla
Cytel Statistical Software & Services Pvt. Ltd, India
raghukishor.k@cytel.com
<http://www.cytel.com>