

# Producing NONMEM dataset using a standard SAS® program

Shafi Chowdhury (1), Alexander Staab (2), Karl-Heinz Liesenfeld (2), Carmen Burger (2), Modesta Wiersema (2)

(1) Shafi Consultancy Limited; (2) Boehringer Ingelheim Pharma GmbH & Co. KG

## Objectives

The overall objective is to implement a system which will ensure results from NONMEM analysis to be ready in time for decisions made about further drug development. One main problem is a bottleneck in the data building part of the process, especially with the complex datasets in phase 2 and 3. The aim is to produce a standard validated program which greatly reduces the time taken to prepare a dataset for NONMEM from the raw data in the database.

## Method

In order to develop a standard program we had to ensure that the input, process and output parts of the program were standard across studies. To this end, key data components were identified for the NONMEM dataset, and a standard structure was developed for each of the components. These included the dosing history (non steady state and/or steady state; intravenous and/or extravascular administration), PK (plasma and urine) and/or PD observations, time independent covariates (e.g. sex), time dependent covariates (e.g. weight), lab covariates and covariates such as concomitant therapies which the patient may or may not be taking at any given point in time. Standard rules were developed with regards to handling and identifying missing data, and how to merge various types of covariates to the dosing and observation records. All these specifications were detailed in a standard NONMEM specifications document. These ensured consistency across studies in the final NONMEM dataset both in terms of structure and definition of various covariates. It also helped to reduce program modification between different studies.

## Result

A set of programs were developed consisting of 15 macros and over 6500 lines of code. They were all called from one main program, which checked data and applied standard rules to handle data issues, created standard input datasets from the various raw data, and transformed the standard input datasets into the structure required by NONMEM. Six standard input datasets are created from the various source datasets within the study database. These are then used in the rest of the program which will remain mainly unchanged. The dosing and observation data is set together to create the main body of the NONMEM dataset. The covariates with any missing values replaced by standard rules are then merged to the main dataset. Time independent covariates are merged by patient number and time-dependent covariates are merged by patient and visit number. Laboratory covariates are then merged based on patient number, visit number and date, as it is possible for a patient to have repeated lab measurements within a visit. Finally the co-medications are merged by patient number, date and time, and a flag is assigned for each CT group to identify if the patient was taking something from that particular medication group at each dosing and observation record.

## Conclusion

The time taken to produce the NONMEM dataset was reduced to a quarter of what it took in the past, and it is reduced even further once the user is familiar with the process. This has ensured the NONMEM analysis can be performed very soon after the database is locked, and in time for decisions made about future drug development.

## Flowchart of the standard SAS® program

