Using ‘Big’ NONMEM Dataset Generated from Standard SDTM to Review Data and Accelerate Model Building Process

Kantish Chowdhury, Md. Rajwanur Rahman, Shafi Consultancy Bangladesh
Shafi Chowdhury, Shafi Consultancy Limited

Objectives:

The use of standard SDTM structured data has provided the opportunity to generate one ‘Big’ NONMEM dataset by merging key SDTM datasets together. It is possible to merge all related data together by using the key identifying variables within those datasets, which are always the same. The advantage of doing this is that it is then possible to easily perform modelling and simulation on the big data very early on and multiple times during the trial without requiring too much resource. Although this will be based on dummy treatments, it can help to clean key data and finalise the NONMEM dataset specification prior to the database lock.

Methods:

The ‘Big’ NONMEM dataset is generated using the standard NONMEM procedure, by setting all the observations and dosing records together and then by merging groups of covariates to them, one group at a time based on the type of covariate. Usually the covariates to be merged are pre-specified, but in this case, all covariates from key datasets can be added as it requires no further effort. The first few versions of the NONMEM dataset generated using this method will not be perfect; there will be data issues that require resolutions. However, it will provide the Pharmacometrician a good insight into the trial data. They will then be in a better position to minimise and finalise the specification prior to the database lock.

Results:

The statistical analysis performed automatically on all covariates to measure the association between them shows if some of the covariates were related to each other. This allows some covariates to be dropped from the final specification where this matches the pre-study expectations. The graphical versions of Patient profiles also helps to review and understand the data. .

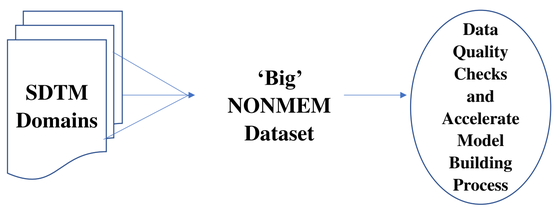
Conclusions:

Looking at all associations and summary data of individual covariates makes it easy to see if there are any data issues, and which covariates are more logical for inclusion in the NONMEM dataset

specification. As the ‘Big’ NONMEM dataset was already generated multiple times, the reduced final NONMEM dataset is produced with minimum additional time. As associations between covariates are also known, this can further aid the modelling process, helping to accelerate the process.

This helps to bring some of the data exploration aspects of modelling prior to the database lock, thus ensuring the Pharmacometrician is more familiar with the data by the time they start the modelling process. This too can speed up the delivery of the analysis and means the results will be available in time to influence future decisions.

1. Basic Process Flow with ‘Big’ NONMEM Dataset



2. SDTM to ‘Big’ NONMEM Dataset

The following are the steps for creating the ‘Big’ NONMEM dataset from SDTM:

- Demographic and randomisation data from DM and DS domains are collected and processed.
- Administration data from EC, EX, SU and SE domains are collected and combined together.
- Observational records for PK, PD and urine are collected from PC, PP and LB domains, and merged with administration data to create a combined dosing and observational dataset.
- Demographic time dependent covariates information is collected from VS and SV domain, and merged with dosing and observational dataset.
- Demographic time independent and randomisation data are then merged with earlier the dataset to create a basic ‘Big’ NONMEM dataset.
- “Findings” information from laboratory, vital signs and other categories is collected from LB, VS and EG domain, processed, and then merged with basic NONMEM dataset.
- Concomitant medication information is collected from CM domain and processed, and merged with earlier dataset.
- Study-specific variables are collected and added with the earlier dataset to create the ‘Big’ NONMEM dataset.

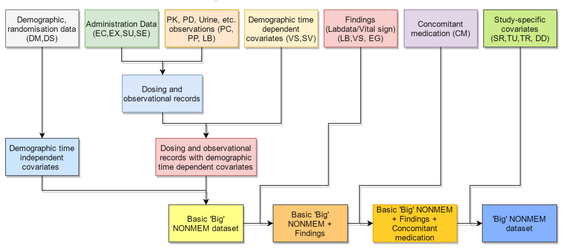
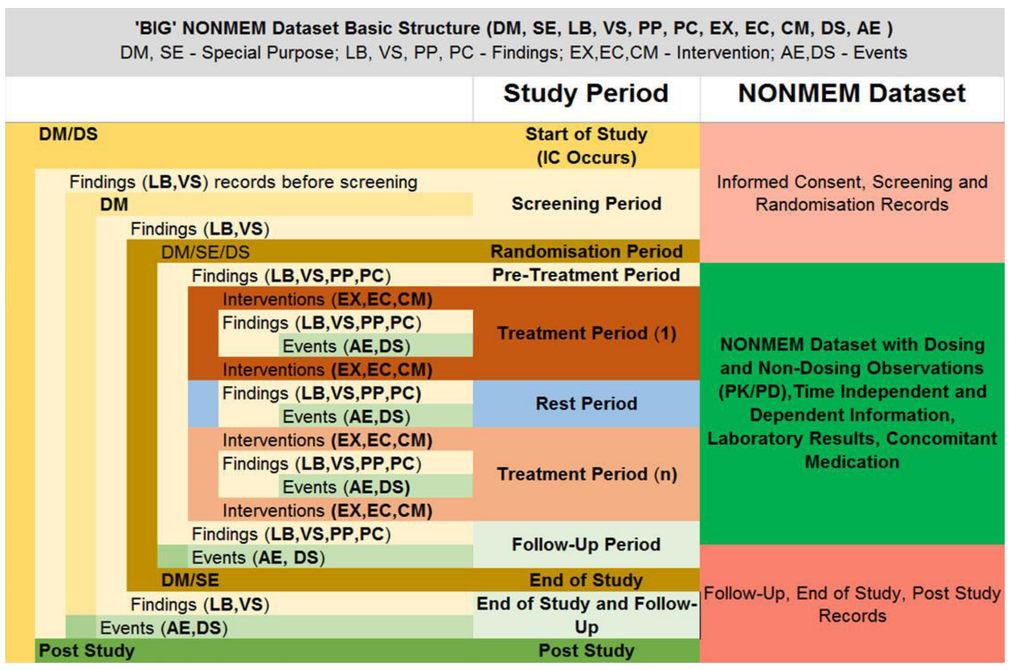


Figure: SDTM to ‘Big’ NONMEM



3. Why call it ‘Big’?

A ‘Big’ NONMEM dataset contains all of the observational data in a study from the beginning to the end. All observational records are merged together by corresponding key variables and sorted by date/time variables along with other key variables.

All SDTM domains and necessary information for NONMEM can easily be found in one place. Special-purpose domains information (e.g. demography, subject visits, subject elements), intervention taken (e.g. exposure collected, exposures, concomitant medication), baseline and related findings (e.g. vital sign and laboratory results, pharmacokinetic parameters and concentrations) before and/or during and/or after intervention taken, adverse and dispositional events of the subjects in a clinical study are all present. From the ‘Big’ dataset, it can easily be found which records are going to come in the NONMEM dataset with PK/PD observations, time independent and dependent information, related laboratory findings and concomitant therapy related flag covariates.

4. Why use ‘Big’ NONMEM Dataset?

SDTM has a specific data structure and it is already known which variables are coming from which domain. As ‘Big’ NONMEM dataset is based on SDTM, it can be created based on a default specification prior to having the study-specific specification, which will accelerate the overall programming process.

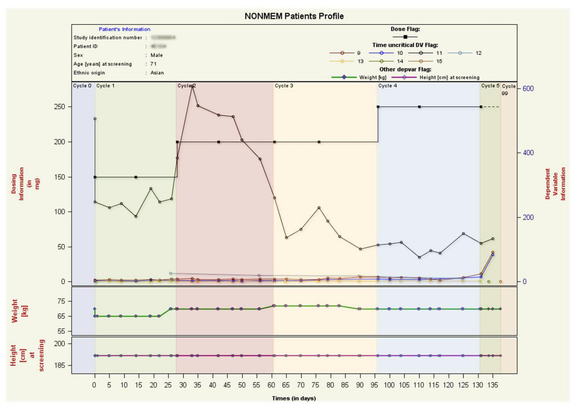


Figure: Patient Profile

Based on this ‘Big’ dataset, a programmer will be able to quickly identify any missing information that are required for the final NONMEM dataset or can point out any data issues which may exist within the data prior to the database lock. The data manager can then quickly rectify these issues and the overall quality of data will be improved.

Project statistician and/or programmer can easily generate charts from this ‘Big’ dataset to identify any irregularities or any pattern which may help making significant decisions regarding the data. Important variables can also be analysed over time continuously to identify potential issues. Project managers or statisticians can focus on the improvement of specification building process based on these findings.

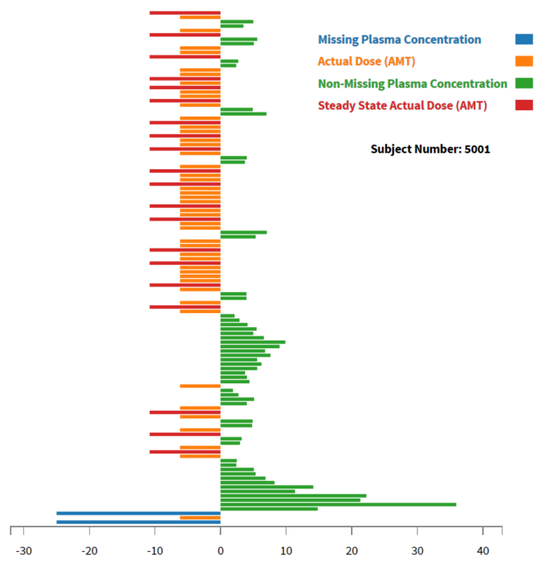


Figure: Dosing / Non-dosing records

This standard NONMEM dataset structure that is produced repeatedly over time allows the Pharmacometrician to review the data using different graphical tools. They can review patient profiles and other graphical forms to confirm that the data makes sense and to identify any issues that may still be present.

Using SDTM datasets to generate the NONMEM dataset means it can be produced very quickly both at the start of the trial and reproduced at the end after database lock. The specifications will be quick for each new trial, and having a standard tool for the generation of NONMEM dataset means not only does it save time, but it will be consistent between trials, be able to produce project NONMEM datasets and have the highest quality.

Please feel free to contact us for further information.

www.shaficonsultancy.com

Shafi Consultancy Ltd.