

# PhUSE 2007

## Paper RA07

### **A validated tool to create the Submission and Analysis Data Sets**

Claude Guyot, sanofi-aventis pharma, Chilly-Mazarin, France

#### **ABSTRACT**

The standardization of the SDS & ADS creation process involves a validated tool by following a computerized methodology. This validated tool will ensure quality and gain/improve efficiency, and will permit programmers to produce easier and quickly their standard datasets. They will mainly focus on the specificities of their study and they will have more time to include and perform the programming of it.

This validated tool must be sufficiently flexible to take into account different study designs and different structures of database.

#### **INTRODUCTION**

In a first step, this presentation will show that the study programmers being the users of the tool are involved in the conception of it. In a second step, the main principles of the tool will be presented and then we will focus on the macro which performs the mapping of the data and on the macro which provides the structure of the ADS/SDS.

In a last step, as we are dealing with a validated tool, we will see the accessibility to the tool with the linked administrative tasks and the traceability means.

#### **A TOOL DONE FOR THE USERS, BY THE USERS**

##### **INVOLVEMENT OF THE USERS IN THE COMPUTERIZED METHODOLOGY**

- The user leader.

He is one of the principal actors in the computerized methodology. He is the representative of the users' community from the launch of the project to the closure of it. He approves the main documents of the project and he is the author of the documents concerning the users' tests. He leads the users group to write the users requirements.

- Requirements specification.

This document lists all the wishes of the future users of the tools. With a users group representing the community of the study programmers (all the therapeutic areas and all the countries are represented) and with the participation of some biostatisticians, the user leader defines all the scope that the tool will must cover.

##### **USER ACCEPTANCE TESTS**

- The user leader.

He defines the user tests design and is the approver of the user acceptance tests cases (UATC). After the tests, he writes the user tests report.

- The user tests.

They are done by study programmers the most representative of the community of the programmers in term of therapeutic area and country.

##### **TRAINING**

All the future users of the tool must be trained. This training given by people closely involved in the project (project manager, developers and user leader) is mainly a practical training through individual exercises directly on PC. The participation to the training gives the access to the tool.

# PhUSE 2007

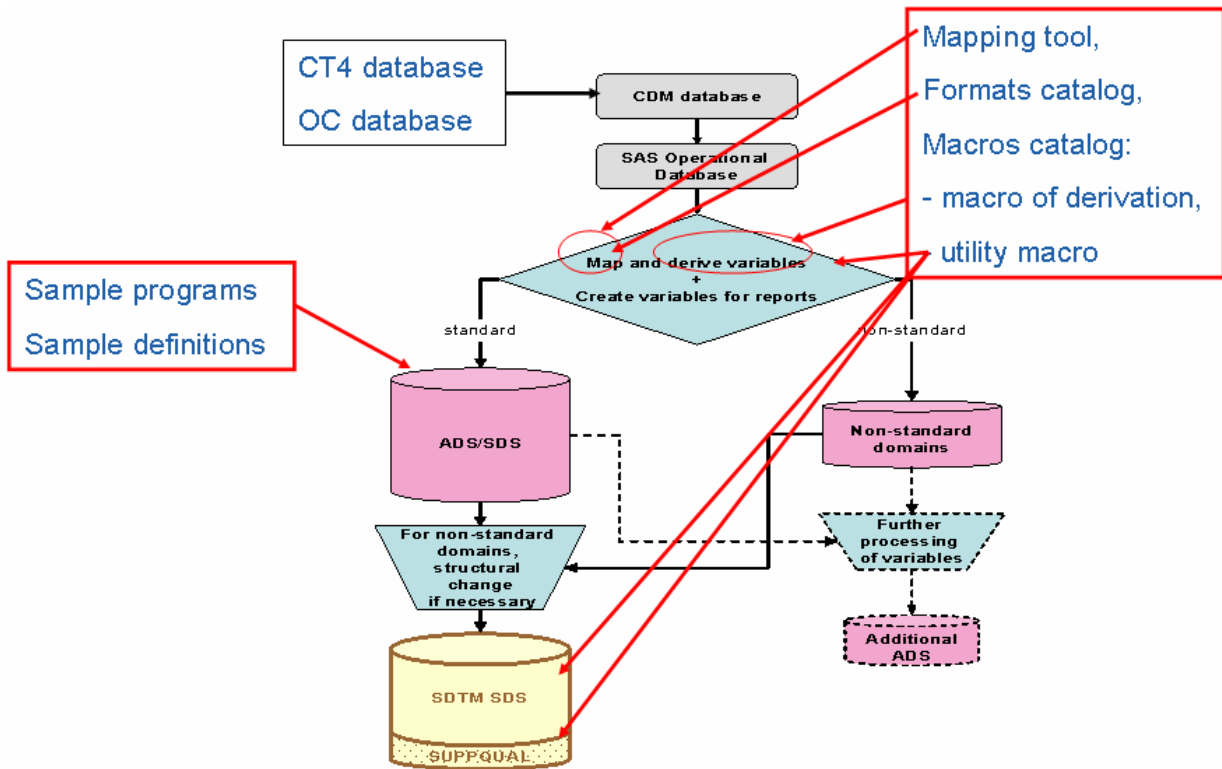
## CHANGE CONTROL

The Change Control is the procedure of the computerized methodology which allows upgrading the tool in a qualified way. The users can be the cause of a Change Control by new important requirements or by raising a big issue in the process put in production (problem in a derivation macro or in the conception of the tool for example).

The users can be involved also in the user tests in the scope of a Change Control.

## TOOL PRINCIPLES

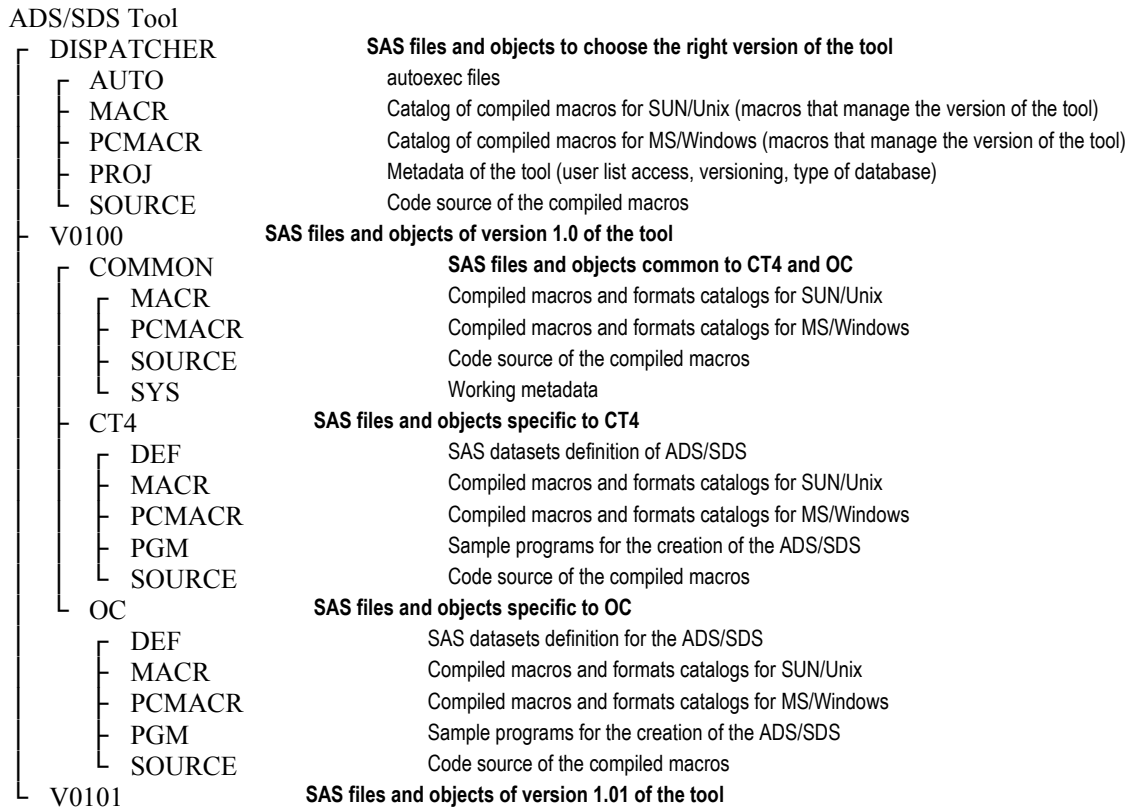
### DATA FLOW



Note: The color pink stands for ADS and are used for reporting.  
Both the ADS and the SDTM SDS/SUPPQUAL data are submitted.  
The dotted line means when needed.

# PhUSE 2007

## ARCHITECTURE OF THE TOOL



## SAS MACROS

The macros are classified according to their role:

- The derivation macros which are stored at the level of the version of the tool in COMMON or CT4 or OC folders depending if they are linked or not to the study database (Oracle Clinical or Clintrial 4).

- The utility macros which are stored either in DISPATCHER folder for the macros used in the management of the tool or in the COMMON folder of the tool version.

The macros are compiled in a catalog stored in the folders MACR and PCMACR and the code is not accessible for the users.

Each macro has a documentation which is stored in a shared e-room and can be consulted by the users.

## DATASETS OF REFERENCE

To derive the variables concerning for example the PCSAs (Potentially Clinically Significant Abnormalities), the grading of adverse events according to the NCI (National Cancer Institute) CTC (Common Terminology Criteria), the concomitance of the treatments we used reference datasets to manage easier the vast number of cases.

# PhUSE 2007

## TWO MAIN MACROS

### MACRO TO PERFORM THE MAPPING

The mapping is the simple derivation from a source variable to a target variable.

- Several kinds of mappings:

- Mappings 1 to 1
- Mappings 1 to 1 with a format application
- Mapping with a change in the type (Num to Char or Char to Num)
- The 3 previous kinds of mapping but with a target variable name equal to the source variable name

- Definition file of the dataset

The definition file is a dataset describing the structure variable per variable of the ADS/SDS dataset we want to create. Inside this definition file, for the variables for which we want to perform a mapping, the source information must be filled in (source dataset, source variable name, source variable type) and if needed the name of the format to apply to the source values to obtain the target values.

The mapping macro allows selecting the mappings to perform.

### MACRO TO CREATE THE STRUCTURE OF THE ADS/SDS

The objective of this macro is to create or to adapt the structure of the output SDS/ADS according to their definition file.

If the user does not fill in an input dataset in the macro call, only an empty ADS/SDS will be performed with the structure defined by the definition file.

If the user fills in an input dataset in the macro call, the variables properties defined in the definition file will be applied to the same variables (same name and same type) of the input dataset.

If the type is not the same, the variable from the input dataset is kept if we have values; otherwise the variable is created as defined in the definition file and is finally empty.

In the output dataset, all the variables are sorted as defined in the definition file and the other variables from the input dataset are put after.

## A VALIDATED TOOL

### MANAGEMENT OF THE TOOL

The access to the tool is controlled and managed by system datasets

- Users' access

When a user is trained, the access (or the authorization to use the tool) is given. He is recorded in the system dataset of the users.

Then, for each study program using the tool, the process verifies that the user is listed in the system dataset of the users and, if yes, the process goes on; otherwise the process stops with a message in the LOG to explain that he must be trained to have access to the tool.

- Study access

All the studies using the tool are recorded in another system dataset. This dataset defines the kinds of database of the study (Oracle Clinical or Clintrial 4) and which version of the tool is used.

Then, for each study program using the tool, the process verifies that the study is listed in the system dataset and which catalog of macros to initialize.

### TRACEABILITY

## PhUSE 2007

All the actions on the tool and the use of the tool are recorded.

- Change Controls and routine maintenance

All the corrections or modifications on the tool are recorded by the computerized methodology documentation.

- Audit on the system datasets

All the action on the system datasets (add a new user, deactivate a user, add a study, modify the version of the tool for a study ...) are recorded with a SAS audit attached on each system dataset.

- Audit on the use of the tool

In production, for each study program using the tool, all the calls of macro are recorded in a SAS dataset with the value of the macro-parameters. The information of the study and the version of the tool are also recorded with the date and time of the launching, the programmer ID, if this is the creation of the dataset or an update. The users do not have writing access on these audit datasets outside the process.

### CONCLUSION

In spite of the regulatory and quality constraints and the imposed standards, the development of such tool must be done with the main goal: satisfy the final users.

To reach this goal, the involvement of the study programmers in the project must be real to know what they expect from the tool and to explain them also that the development of a validated tool must follow a computerized methodology.

### REFERENCES

Internal Computerized System Methodology

### ACKNOWLEDGMENTS

N/A

### RECOMMENDED READING

N/A

### CONTACT INFORMATION

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

Claude Guyot  
Sanofi-aventis  
1, avenue Pierre Brossolette  
Chilly-Mazarin 91385 Cedex FRANCE  
Work Phone: (33) 1 69 79 72 72  
Email: [claude.guyot@sanofi-aventis.com](mailto:claude.guyot@sanofi-aventis.com)

Yvane Boudraa  
Sanofi-aventis  
20, avenue Raymond Aron  
Antony 92165 FRANCE  
Work Phone: (33) 1 55 71 63 95  
Email: [yvane.boudraa@sanofi-aventis.com](mailto:yvane.boudraa@sanofi-aventis.com)

Brand and product names are trademarks of their respective companies.