PhUSE 2011

Paper TS02

Using Dynamic Lists in SAS Stored Processes for Genetic Toxicity Historical Control Data

Volker Harm, Bayer Pharma AG, Berlin, Germany

ABSTRACT

In standard tests for genetic toxicity historical control data are used to assess the validity of the experiments as well as additional information for the statistical analysis of the tests. Thus compilation, management and reporting of this data gain ever more importance.

Existing SAS Stored Processes can be significantly enhanced by the use of dynamic lists to cope with an increasing number of test types, types of controls, and laboratories.

The paper will introduce the data that is handled, show how the application is built up, describe the use of dynamic lists in SAS Stored Processes and explain how the users benefit from this new flexibility.

INTRODUCTION

STANDARDIZED GENETIC TOXICITY TESTING

Genetic Toxicity Testing is part of the safety assessment in Phase I of drug development. It is mostly done on animals and required by the authorities after some very serious incidents (e.g. Contergan, 1961). To ensure a defined level of quality and comparability these experiments are highly standardized. They are based on a fixed protocol recommended by OECD guidelines and strictly audited by the authorities.

The typical structure for a genetic toxicity test is as follows:

Group	Treatment	Dose level
1	Medium/solvent control	
2	Test substance	Low dose
3	Test substance	Mid dose
4	Test substance	High dose
5	Positive control	

Each group usually consists of a constant number of animals. Responses of three doses of a test substance are compared to the responses of medium/solvent control to asses an effect of the substance. Responses of the positive control are compared to the responses of the medium/solvent control to asses the sensitivity of the assay. This setting may be repeated under different experimental conditions as with or without metabolic activation and pulse or continuous treatment.

CONTROL GROUPS

A plethora of variables (e.g. genetic, environmental, infectious agents) can potentially affect the outcome of studies performed with animals. It is therefore critical to use control animals to minimize the impact of these extraneous variables or to recognize the possible presence of unwanted variables. Control groups are not treated with the substance to be tested, they are used to control the test system. Responses seen in control groups should not vary from study to study, but could vary in different experimental settings (e.g. different labs).

Control types considered here are medium (completely untreated), solvent (treated with the solvent of the substance), and positive (treated with a dose of a substance with known effect).

THE TASK

Historical Control Data are lists of control group data of previous studies. Statistics of the control group data (means, standard deviations, ranges, confidence limits) are compared to the results of an actual study. These statistics are reported for each study that is conducted. In most cases historical control data are evaluated per year or the last N experiments are taken into account.

Our department developed an application based on the SAS Stored Process Web Application providing a graphical user interface to collect the data and report the statistics. As it is a web application it is accessible by an internet browser through the Intranet.

This application shall be extended for each test and each lab of Bayer´s genotoxicity testing groups.

EXISTING VERSION OF A STORED PROCESS WEB APPLICATION USING STATIC LISTS

In the following I will shortly describe the existing application to handle historical control data. This is done more detailed in a previous paper (Harm, 2009).
The following table shows an example of the list of historical control data for the micronucleus test in vitro.

Data base of historical control data for the In Vitro Micronucleus Test sorted in descending order by study year and study number

Study Year	Study Number	Metabolic Activation	Control Type	Experimental Part	Measurement Number	Cells Scored	Micronucleus containing cells [%]
2009	TOXT0079950	With	Medium	Pulse treatment	1	1000	0.8
2009	TOXT0079950	With	Medium	Pulse treatment	2	1000	1.1
2009	TOXT0079950	With	DMSO (1%)	Pulse treatment	1	1000	1.2
2009	TOXT0079950	With	DMSO (1%)	Pulse treatment	2	1000	1
2009	TOXT0079950	With	Cylclophosphamide (2 µg/mL)	Pulse treatment	1	1000	17
2009	TOXT0079950	With	Cylclophosphamide (2 µg/mL)	Pulse treatment	2	1000	19.2
2009	TOXT0079950	Without	Medium	Continuous treatment	1	1000	0.6
2009	TOXT0079950	Without	Medium	Continuous treatment	2	1000	1.6

For each measurement of control group data the study year, the study number, the experimental setting for metabolic activation, the control type, the experimental part, and the number of the measurements, if there are more than one per experiment, are listed to uniquely identify the measurement. In columns Cells Scored and Micronucleus containing cells [%] the measurement values are given.
The structure of a data set to hold such a list is as follows:

Variables in Table: HCDMNTINVITROBERLIN

Variable Number	Name	Type	Format	Label	Length
1	StudyNo	Character		Study Number	12
2	MetabolicActivation	Numeric	METABOLICACTIVATIONF	Metabolic Activation	8
3	ExperimentalPart	Numeric	EXPERIMENTALPARTF	Experimental Part	8
4	MeasurementNo	Numeric		Measurement Number	8
5	ControlType	Numeric	CONTROLSMNTINVITROF	Control Type	8
6	CellsScored	Numeric		Cells Scored	8
7	MeasurementValue	Numeric		Micronucleus containing cells [%]	8
8	StudyYear	Numeric		Study Year	8

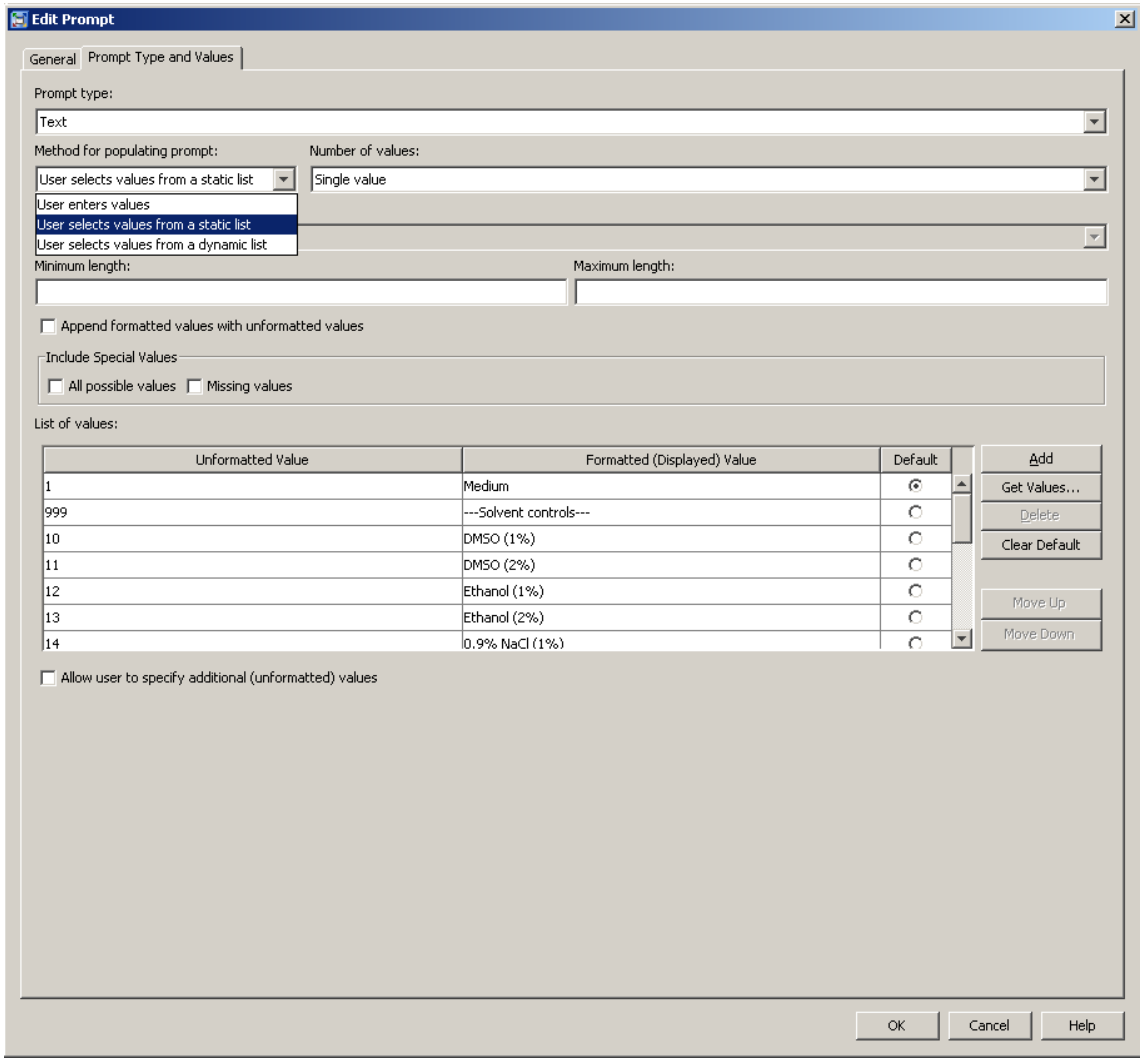
As usual for the different experimental settings formats were used to code the different realizations:

```
proc format;
    value MetabolicActivationf
        1 = "With"
        2 = "Without";
    value ExperimentalPartf
        1 = "Pulse treatment"
        2 = "Continuous treatment";
run;
```

The control substances that would possibly occur were also collected in a format. For reporting purposes also the type of control was implemented in the codes (0-9 medium, 10-19 solvent, 20-29 positive controls). And for experiments in the future we already considered extra positive controls, where the concentration to be used was not yet determined.

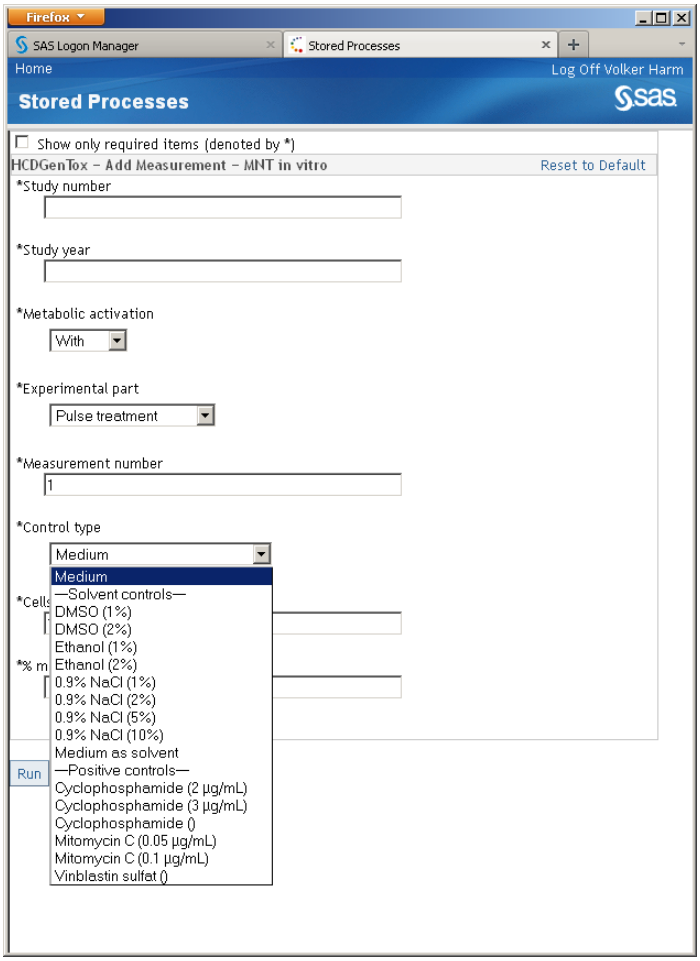
```
proc format;
    value ControlsMNTInVitrof
        1  = "Medium"
        10 = "DMSO (1%)"
        11 = "DMSO (2%)"
        12 = "Ethanol (1%)"
        13 = "Ethanol (2%)"
        14 = "0.9% NaCl (1%)"
        15 = "0.9% NaCl (2%)"
        16 = "0.9% NaCl (5%)"
        17 = "0.9% NaCl (10%)"
        18 = "Medium as solvent"
        20 = "Cylclophosphamide (2 µg/mL)"
        21 = "Cylclophosphamide (3 µg/mL)"
        22 = "Cylclophosphamide (Conc. still open)"
        23 = "Mitomycin C (0.05 µg/mL)"
        24 = "Mitomycin C (0.1 µg/mL)"
        25 = "Vinblastin sulfat (Conc. still open)";
run;
```

This setup fits nicely into the concept of static lists for the SAS Stored Processes prompts as can be seen in the following screenshot of the Edit Prompt dialog:



For better readability also subheaders for the control the control type were inserted by using special codes.

These definitions lead to the interface of a SAS Stored Process “Add new measurement – MNT in vitro” as shown in the following screen shot.

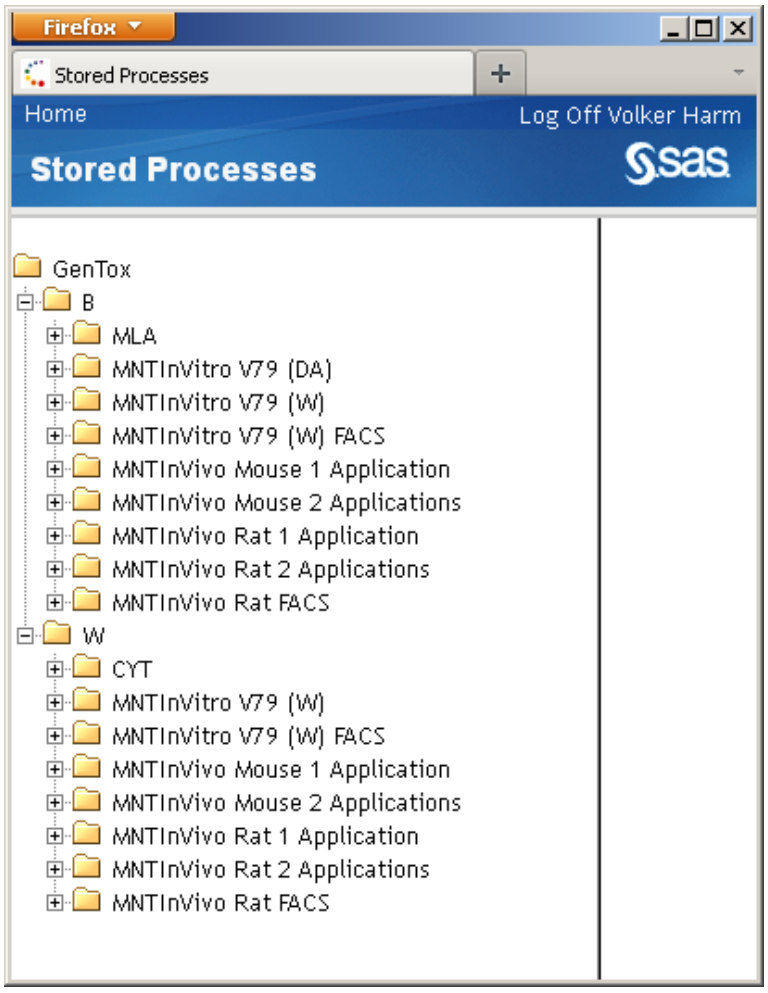


On execution of the SAS Stored Process the macro variables necessary for executing are returned to the program of the SAS Stored Process so that the following code can be executed to update the database:

```
proc sql;
    insert into org.&DBName. (
        StudyNo,
        StudyYear,
        MetabolicActivation,
        ExperimentalPart,
        ControlType,
        MeasurementNo,
        CellsScored,
        MeasurementValue
        )
    values (
        "&StudyNo.",
        &StudyYear.,
        &MetabolicActivation.,
        &ExperimentalPart.,
        &ControlType.,
        &MeasurementNo.,
        &CellsScored.,
        &MeasurementValue.);
quit;
```

ENHANCED VERSION OF A STORED PROCESS WEB APPLICATION USING DYNAMIC LISTS

Using the application over time the design showed several flaws. At first we defined some control types, for which the concentration was still not specified. So the list was longer than necessary.
Furthermore we had to cope with new complexity by adding additional control substances, including new tests, and providing different laboratories with the application. Indeed the next goal is the following setup of test systems:



Above this the first application is running now in production. Each change therefore requires a new roll out of the application. So the requirement arose: Is there a way to make the design more dynamic by presenting only the established control substances in the selection lists, presenting the control substances based on a preselected control type, and allowing the user himself to define new control substances.
To achieve a simple solution for this, the structure of the data set containing the historical control data was modified to resolve the formats of the control substances: The types of the control groups got an own variable with the following format:

```
proc format;
    value ControlTypef
        1 = 'Medium Control'
        2 = 'Solvent Control'
        3 = 'Positive Control';
run;
```

The variable which controls the used substances is now of type Character. Thus we get the following structure for our database:

Variables in Table: HCDMNTINVITROBERLIN

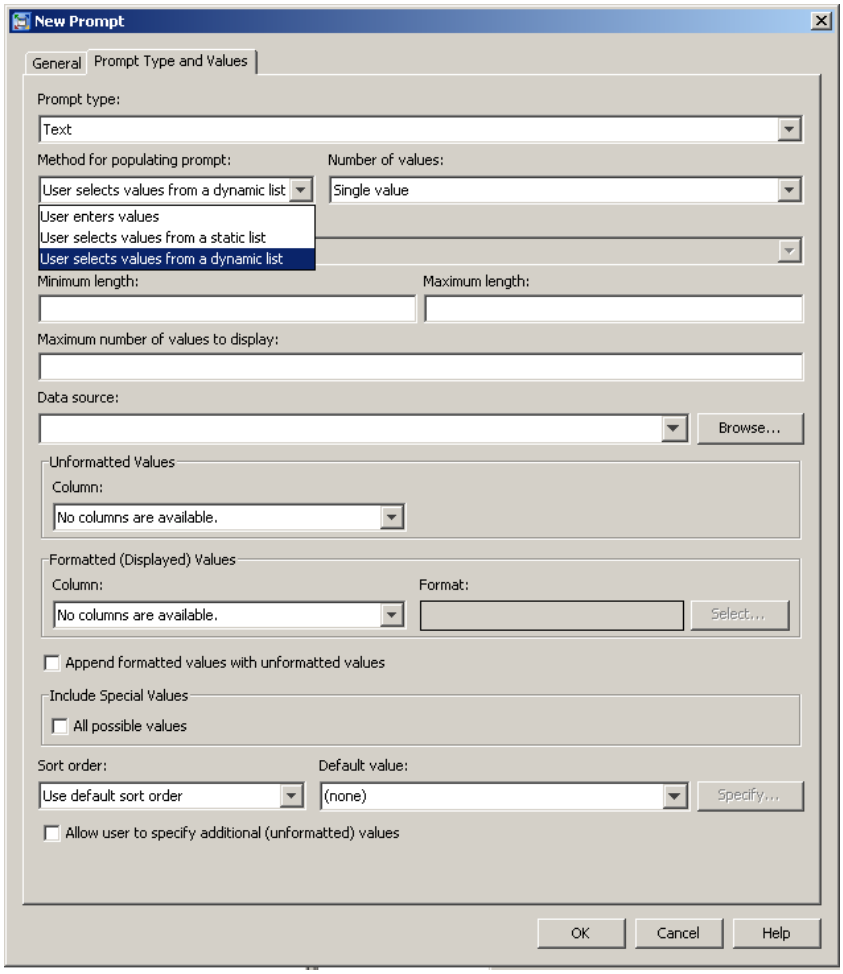
Variable Number	Name	Type	Format	Label	Length
1	StudyNo	Character		Study Number	20
2	StudyYear	Numeric		Study Year	8
3	MetabolicActivation	Numeric	METABOLICACTIVATIONF	Metabolic Activation	8
4	ExperimentalPart	Numeric	EXPERIMENTALPARTF	Experimental Part	8
5	ControlType	Numeric	CONTROLTYPEF	Control Type	8
6	XControlSubstance	Character		Control Substance	100
7	MeasurementNo	Numeric		Measurement Number	8
8	CellsScored	Numeric		Cells Scored	8
9	MeasurementValue	Numeric		Micronucleus containing cells [%]	8

DYNAMIC LISTS

With this modified data structure we can exploit two new features, which the SAS 9.2 prompt framework introduced: Dynamic Prompts, which allow the lookup of possible prompt values from a data set and Cascading Prompts, which allow possible prompt values to be dependent on the selected value of another prompt .

These features are enabled by setting the Method for populating prompts to “User selects values from a dynamic list” in the prompt editor.

As can be seen in the following screen shot the prompt needs for this type of selection a data source and the name of the column, which contains the values that the selection will be build of. It is possible to use unformatted and formatted values. There is also a check box, which can be checked to allow the user to specify additional values. This is possible only when we work with unformatted values. One more reason for the modified data structure.

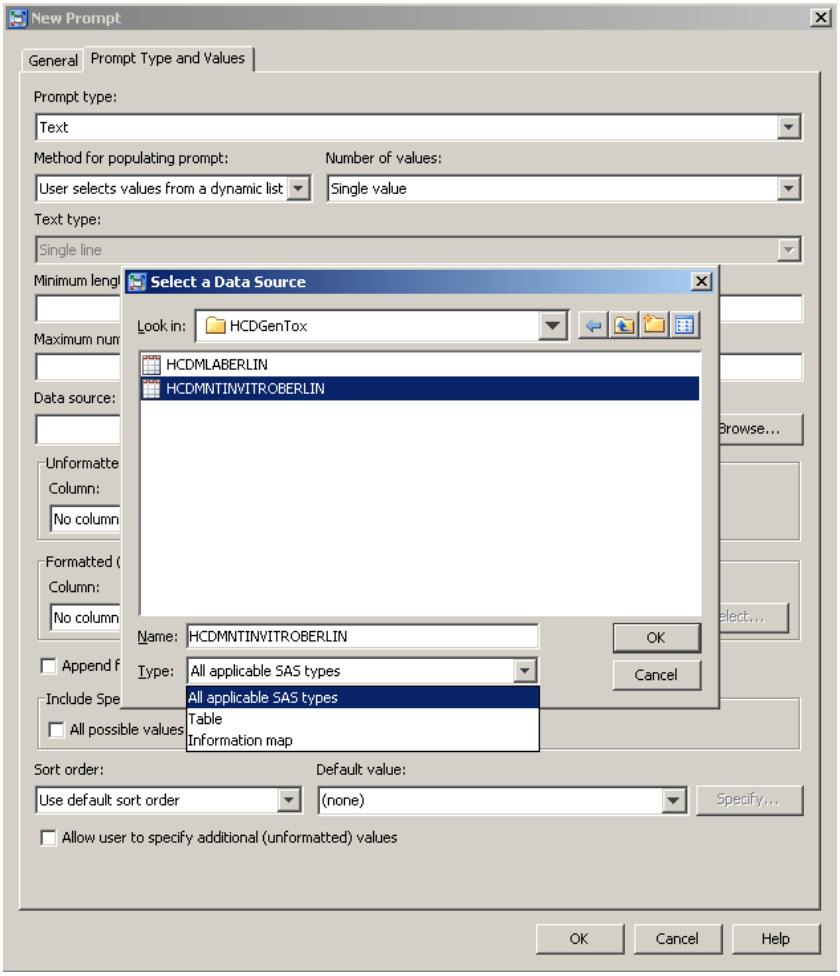


DATA SOURCES

As the prompts are part of the metadata of the SAS Stored Process their definitions are stored in the SAS Open Metadata Repository. Thus they cannot use the data assignments done during SAS program execution. Therefore to use data within the prompt framework they must be known within the metadata; they must be “registered” in the SAS Open Metadata Repository. There are several SAS objects that can be registered as data source. The simplest and most common is a SAS data set.

Registering data sets is a two-step procedure done with the Data Library Manager in SAS Management Console. Following the known logic of SAS as a first step a library must be defined in the metadata, which corresponds to an existing SAS library, and as a second step within this defined library an existing SAS data set must be registered. During registration of the data set metadata information as variable names and type are stored in the SAS Open Metadata Repository. A “Getting started” how to do that can be found in the document “SAS Information Map Studio 4.2 Getting Started with SAS Information Maps”. More detailed instructions can be found in the document “SAS 9.2 Intelligence Platform Data Administration Guide” and minimal but sufficient hints in the Data Library Manager help.

Once a data source is registered in the SAS Open Metadata, it can be selected in the prompt editor as shown in the following screen shot. In our case it is the data set containing the historical control data.

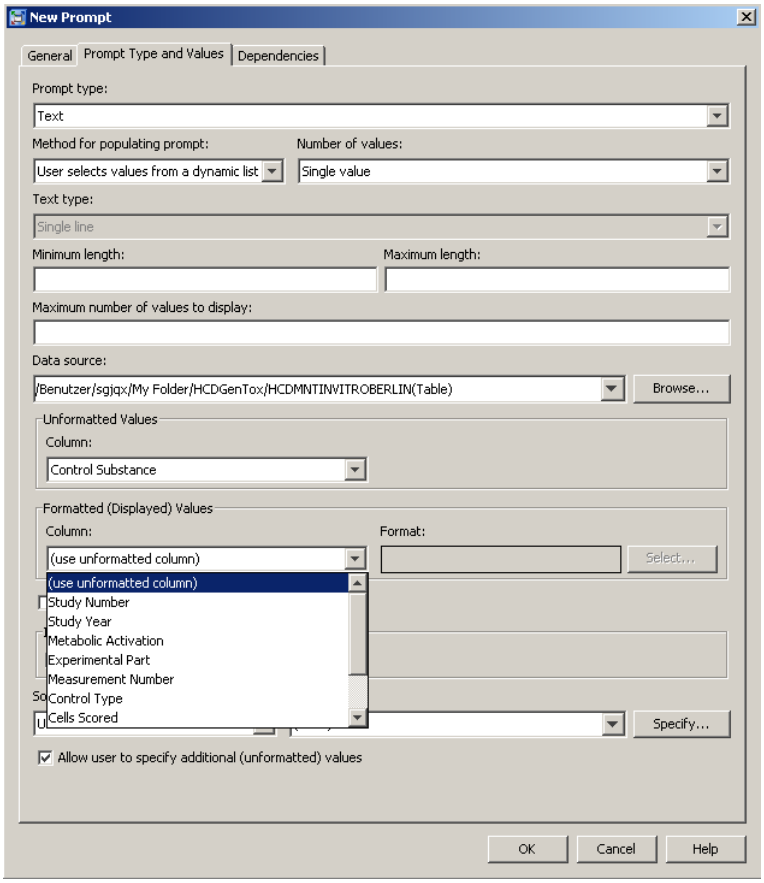


THE NEW CONTROL SUBSTANCE SELECTION LIST

Once the data source is selected, the information about the data source stored in the SAS Open Metadata Repository is available to the prompt. The sections Unformatted Values and Formatted Values present selection lists containing the labels of the appropriate variables. For the example given here we use only the unformatted variable XControlSubstance with the label Control Substance.

To use also formatted variables the formats have to be collected in a format catalog. This catalog has to be stored in specific place. Further information can be obtained in the SAS Data Administration Guide.

The following screen shot shows the situation for the historical control database. (Note that “Allow user to specify additional (unformatted) values” is checked.)



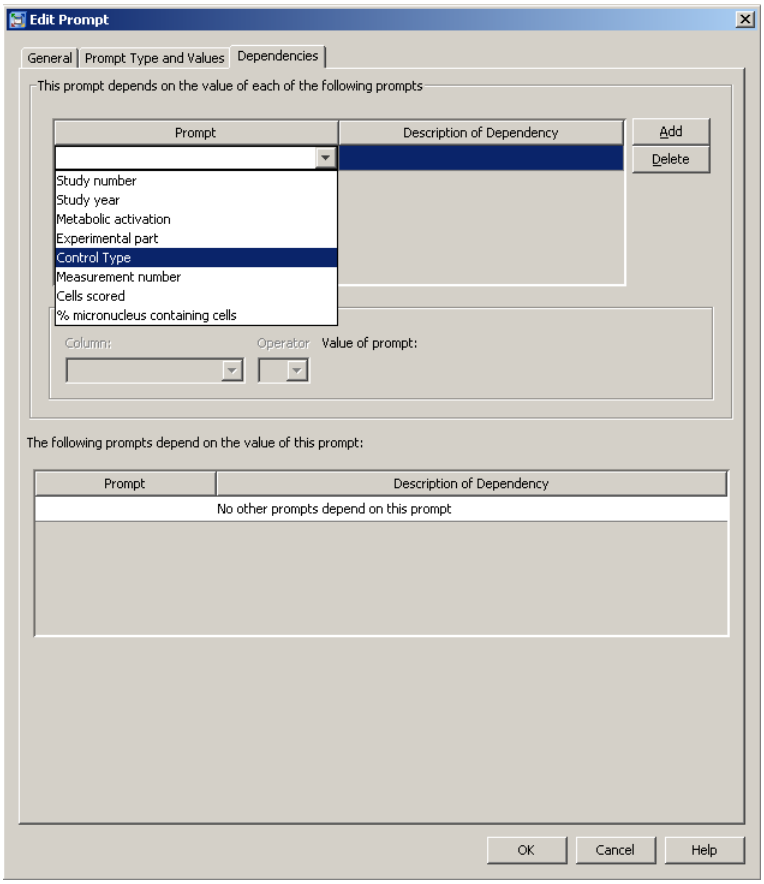
With these definitions the prompt is advised to uniquely select values from the specified column and allow the user to enter new control substances.

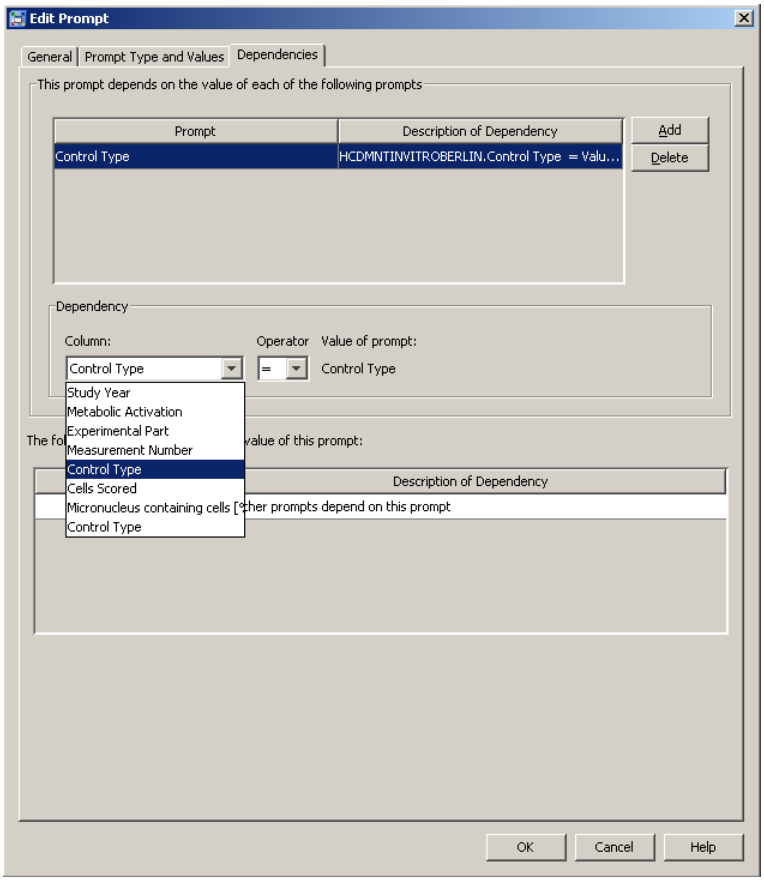
DEPENDENCY ON THE CONTROL TYPE

What is also seen in the screen shot above is a new tab Dependencies that shows up after specifying the data source. This feature is used to make the selection list of this prompt dependent on the selected control type.

In a first step the Control Type prompt is selected from a list of available prompts that shows up after clicking “Add” in the section “This prompt depends on the value of the following prompts”. Control Type is selected as seen in the first of the two following screen shot.

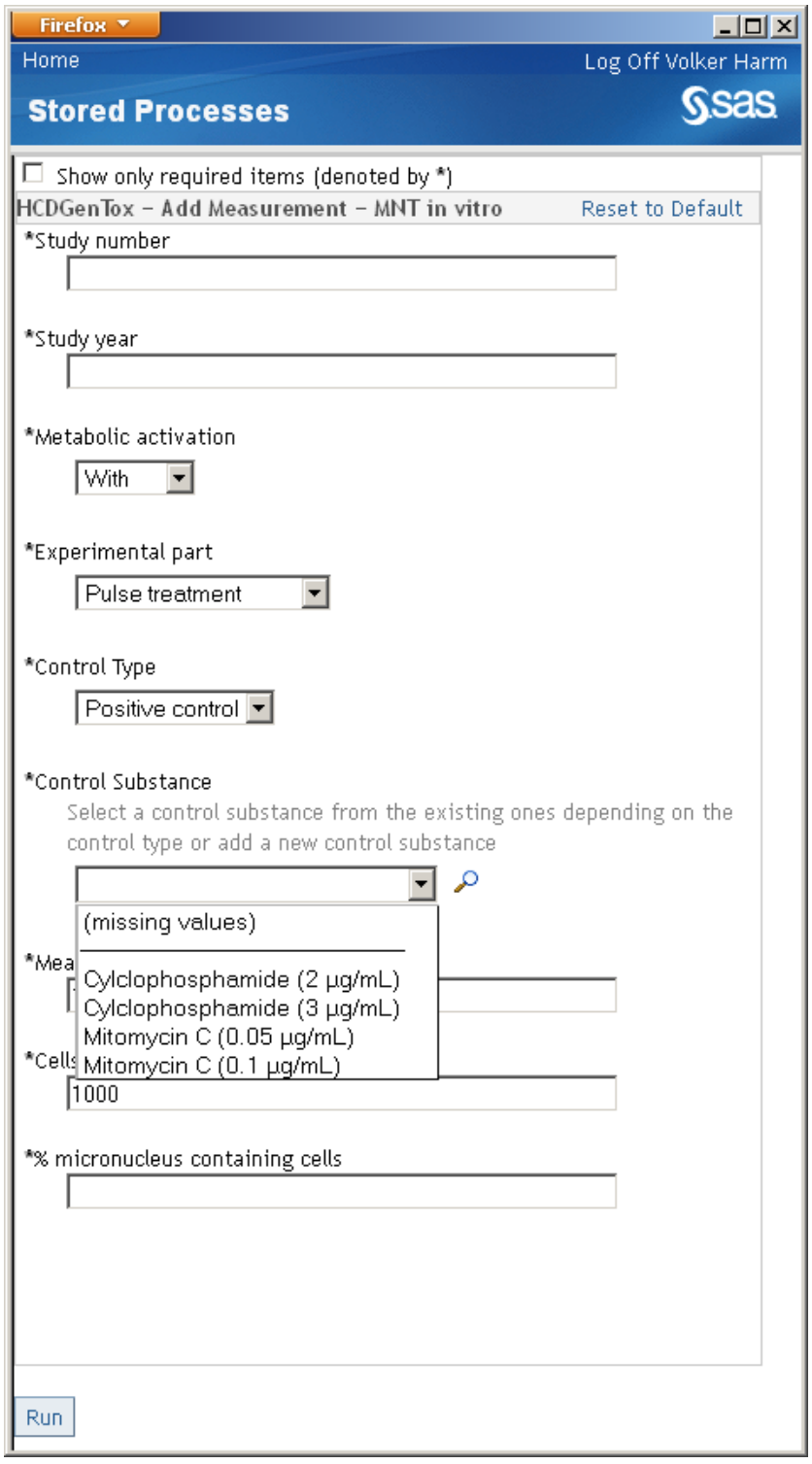
In a second step the type of dependency is set. This is shown in the second screen shot.


Edit Prompt
General
Prompt Type and Values
Dependencies
This prompt depends on the value of each of the following prompts
Prompt
Description of Dependency
Add
Delete
Study number
Study year
Metabolic activation
Experimental part
Control Type
Measurement number
Cells scored
% micronucleus containing cells
Column:
Operator
Value of prompt:
The following prompts depend on the value of this prompt:
Prompt
Description of Dependency
No other prompts depend on this prompt
OK
Cancel
Help


Edit Prompt
General
Prompt Type and Values
Dependencies
This prompt depends on the value of each of the following prompts
Prompt
Description of Dependency
Add
Delete
Control Type
HCDMNTINVITROBERLIN.Control Type = Valu...
Dependency
Column:
Operator
Value of prompt:
Control Type
=
Control Type
Study Year
Metabolic Activation
Experimental Part
Measurement Number
Control Type
Cells Scored
Micronucleus containing cells
Control Type
Description of Dependency
OK
Cancel
Help

THE MODIFIED SAS STORED PROCESS

As a last step we adapt the description for our new Control Substance Prompt to give some usage information under the General tab and get the following new interface for our SAS Stored Process.



CONCLUSION

Looking at the list of tests for which I have to implement the Historical Control Databases I feel much more comfortable with this new flexibility!

REFERENCES

Davidson, Kevin and Duong, Minh . 2010. Using Dynamic and Cascading Prompts in SAS Enterprise Guide®". Proceedings of the SAS® Global Forum 2010 Conference

Flynn Joe. 2009. SAS® Stored Processes: Going Beyond the Current Capabilities of the Stored Process Wizard. Proceedings of the SAS® Global Forum 2009 Conference

Harm, Volker. 2009. Providing Historical Control Data for Toxicology With the SAS Stored Process Web Application. PhUSE Conference 2009

Johnson PD, Baseless DG. 2002. Practical aspects to experimental design in animal research. ILAR J 43:202-206

SAS Institute Inc. 2009. SAS® Information Map Studio 4.2: Getting Started with SAS® Information Maps. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2009. SAS® 9.2 Intelligence Platform: Data Administration Guide. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2009. SAS® 9.2 Stored Processes: Developer's Guide. Cary, NC: SAS Institute Inc.

CONTACT INFORMATION

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

Volker Harm
Bayer Pharma AG
Müllerstraße 178
13353 Berlin
Work Phone: +49-30-468-11208
Fax: +49-30-468-91208
Email: volker.harm@bayer.com
Web: http://www.bayerpharma.com