PhUSE 2008

Paper AD09

Automated table generation: The SAS macro %GEEWHIZ

Guido Wendland, UCB Pharma, Monheim, Germany

ABSTRACT

The macro %GEEWHIZ creates tables in PDF-format in which observational units ("subjects") and observations ("events") are counted in categories. The subject count is usually related to the total number of subjects observed. The categories constitute columns of the table. In most cases the treatment group is the categorical variable but other categorizations are possible.

The main purpose of the macro is to be able to produce a large amount of tables with just one macro in a controlled fashion. The macro combines content with layout and thus makes table production more efficient and fast.

Macro parameters are defined in an xml-interface. The paper presents the architecture and provides examples.

INTRODUCTION

Table programming accounts for a high percentage of the statistical programmer trial activities. These activities consist of a couple of repetitive tasks that could be automated. There is usually a large variety of tables involved in clinical trials that require various programming approaches. Tables could however be categorized and for single categories a global macro could be useful. This global macro would automate a lot of steps and programming time would be vastly reduced.

The macro %GEEWHIZ covers the very frequent category of tables in which subjects or events are counted across different groups (e.g. treatment groups). It was attempted to cover almost all tables for which this vague description applies. In the process of development the scope of the macro was expanded to meet the table requirements that occurred inside our studies. While the macro was initially developed to handle adverse events data, it is not restricted to this type of data.

The objective of this macro is to combine ease of use with great flexibility. The paper describes the used techniques to solve this potential conflict of goals. Furthermore the macro %GEEWHIZ combines content with layout.

The aspired flexibility of the macro can be summarized in couple of features:

- Allows for a lot of user-specified options at the same time.
- Creates tables for different domains (such as adverse events, vitals etc.)
- Can be used for development and validation programming (including the option of an automatic compare) in case double programming is used.
- Supports various content-related issues (like subgrouping for intensity, subsetting of drug-associated AEs, inclusion of total and active treatment total columns etc., by group processing, varying denominators etc.)
- Allows for overlapping treatment categories (e.g. time of onset tables).

It was attempted to reach ease of use by

- An xml-interface that allows the user to specify the macro parameters. Thus it is easy to use previous examples and copy and paste information.
- Generating helpful error or warning messages in case of inappropriate parameterization or use.
- Supporting a standard table layout by using default parameter settings. User need not to care about layout issues (like column width, alignment, etc.).
- A comprehensive user manual.

THE MACRO ARCHITECTURE

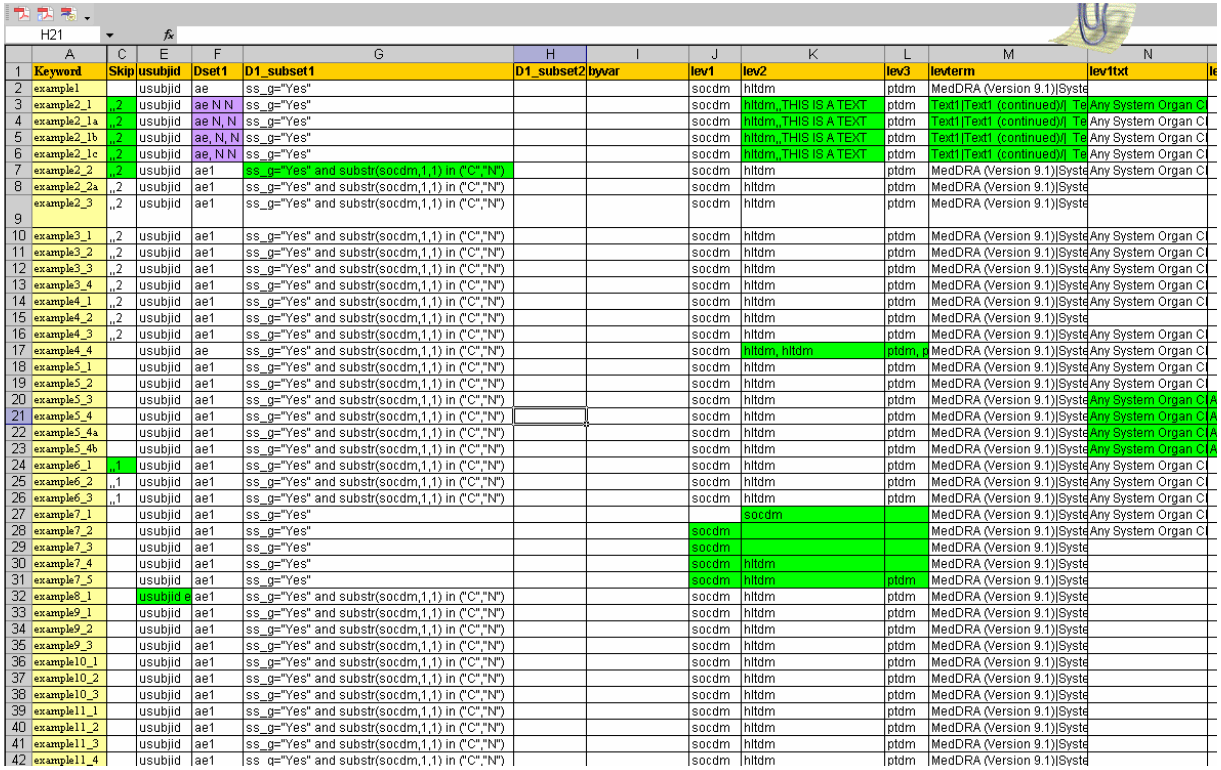
The macro parameters will usually be populated via an Excel-spreadsheet that should be stored as xml-file consisting of two sheets. The macro version of %GEEWHIZ to be used is also one parameter in the xml-file. All parameters of the xml-file are passed to the macro %GEEWHIZ via an interface macro %SETTABS. The main purpose of the macro %SETTABS is to allow for the use of different versions of %GEEWHIZ. The macro %SETTABS transforms the xml-files into SAS® datasets and calls the corresponding version of %GEEWHIZ.

THE XML INTERFACE

It was decided to use an external xml-interface to allow easier manipulations of a large number of parameters and tables, to provide a better overview of the tables to be produced and as a consequence to speed up the table production process. The xml-file can be opened for manipulation in Excel in the same way as an xls-file.

This xml-file consists of two sheets, Sheet1 and Sheet2. The name of the xml-file is user-defined. Sheet1 controls the rows of the table and the general appearance. The column headers of Sheet1 (in row 1) represent the approximately 40 macro parameter names to be used with %GEEWHIZ. For most of these parameters the default specification can be used in many tables.

Figure 1: Excerpt from sheet1 of the xml-Spreadsheet.



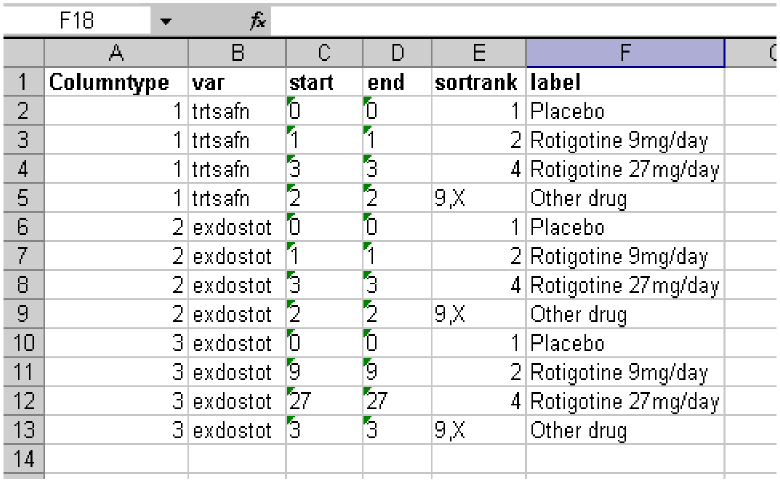
H21

	A	C	E	F	G	H	I	J	K	L	M	N	O
1	Keyword	Skip	usubjid	Dset1	D1_subset1	D1_subset2	byvar	lev1	lev2	lev3	levterm	lev1txt	le
2	example1		usubjid	ae	ss_g="Yes"			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		
3	example2_1	,,2	usubjid	ae N N	ss_g="Yes"			socdm	hltdm,,THIS IS A TEXT	ptdm	Text1|Text1 (continued)/| Te	Any System Organ C	
4	example2_1a	,,2	usubjid	ae N, N	ss_g="Yes"			socdm	hltdm,,THIS IS A TEXT	ptdm	Text1|Text1 (continued)/| Te	Any System Organ C	
5	example2_1b	,,2	usubjid	ae, N, N	ss_g="Yes"			socdm	hltdm,,THIS IS A TEXT	ptdm	Text1|Text1 (continued)/| Te	Any System Organ C	
6	example2_1c	,,2	usubjid	ae, N N	ss_g="Yes"			socdm	hltdm,,THIS IS A TEXT	ptdm	Text1|Text1 (continued)/| Te	Any System Organ C	
7	example2_2	,,2	usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	
8	example2_2a	,,2	usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		
9	example2_3	,,2	usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		
10	example3_1	,,2	usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	
11	example3_2	,,2	usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	
12	example3_3	,,2	usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	
13	example3_4	,,2	usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	
14	example4_1	,,2	usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	
15	example4_2	,,2	usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		
16	example4_3	,,2	usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	
17	example4_4		usubjid	ae	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm, hltdm	ptdm, p	MedDRA (Version 9.1)|Syste	Any System Organ C	
18	example5_1		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	
19	example5_2		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	
20	example5_3		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	A
21	example5_4		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	A
22	example5_4a		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	A
23	example5_4b		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	A
24	example6_1	,,1	usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	
25	example6_2	,,1	usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	
26	example6_3	,,1	usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste	Any System Organ C	
27	example7_1		usubjid	ae1	ss_g="Yes"				socdm		MedDRA (Version 9.1)|Syste	Any System Organ C	
28	example7_2		usubjid	ae1	ss_g="Yes"			socdm			MedDRA (Version 9.1)|Syste	Any System Organ C	
29	example7_3		usubjid	ae1	ss_g="Yes"			socdm			MedDRA (Version 9.1)|Syste		
30	example7_4		usubjid	ae1	ss_g="Yes"			socdm	hltdm		MedDRA (Version 9.1)|Syste		
31	example7_5		usubjid	ae1	ss_g="Yes"			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		
32	example8_1		usubjid e	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		
33	example9_1		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		
34	example9_2		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		
35	example9_3		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		
36	example10_1		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		
37	example10_2		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		
38	example10_3		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		
39	example11_1		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		
40	example11_2		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		
41	example11_3		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		
42	example11_4		usubjid	ae1	ss_g="Yes" and substr(socdm,1,1) in ("C","N")			socdm	hltdm	ptdm	MedDRA (Version 9.1)|Syste		

All other rows usually correspond to one table via the macro parameter “KEYWORD” which links to the corresponding entry in another xml-file that defines the titles and footnotes of the tables.

Sheet2 controls the columns of the intended table and also contains a first row of parameters. Because the number of (treatment) columns is a variable number sheet two is organized differently to Sheet1. In Sheet2 each row provides a column description of the table.

Figure 2: Example of sheet 2 of the xml-Spreadsheet.



F18

	A	B	C	D	E	F
1	Columntype	var	start	end	sortrank	label
2	1	trtsafn	0	0	1	Placebo
3	1	trtsafn	1	1	2	Rotigotine 9mg/day
4	1	trtsafn	3	3	4	Rotigotine 27mg/day
5	1	trtsafn	2	2	9,X	Other drug
6	2	exdostot	0	0	1	Placebo
7	2	exdostot	1	1	2	Rotigotine 9mg/day
8	2	exdostot	3	3	4	Rotigotine 27mg/day
9	2	exdostot	2	2	9,X	Other drug
10	3	exdostot	0	0	1	Placebo
11	3	exdostot	9	9	2	Rotigotine 9mg/day
12	3	exdostot	27	27	4	Rotigotine 27mg/day
13	3	exdostot	3	3	9,X	Other drug
14						

Sheet1 and Sheet2 are linked via the macro parameter “COLUMNTYPE”. So you can use different column categorizations across the set of tables defined in the xml-file.
There are a few requirements to bear in mind in order to read xml-files:

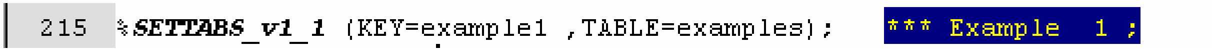
Xml-files can automatically be opened with Excel and you can set this up as a default (right click on the file, Open with Microsoft excel and check “Always use this program to open these files”; alternatively specify specifying Tools-> Folder Options -> File Type etc).

- To ensure proper conversion into a SAS file no empty field is allowed in the xml-file. If you want to have no information in an xml-field add a blank character (“ “) to it, e.g. by choosing Edit -> Replace and "replace all".
- The macro %XLXP2SAS from SAS Institute (-> DelGobbo 2006) needs to be available. A typical invocation would look like this: % XLXP2SAS (excelfile=&PATH…./titles.xml);

Under windows you have to close the xml-spreadsheet in order to retrieve the information from it. Users that prefer to enter the parameter values directly can alternatively use the macro %SETTABS (see below).

THE MACRO %SETTABS

The user has to invoke the macro %SETTABS to create the tables. This macro %SETTABS constitutes a link between the xml-file and the macro %GEEWHIZ, which is called automatically.



```
%SETTABS_v1_1 (KEY=example1 ,TABLE=examples);    *** Example  1 ;
```

This invocation of %SETTABS shows its minimal parameterization: The parameter TABLE refers to the xml-file that is used (in the same folder as the program) and the parameter KEY refers to the row in the xml-file and determines the table to be produced. It has to match the corresponding KEY that is specified inside the xml-file. Usually other macro parameters can be used as default. You can use %SETTABS also to overwrite macro parameter setting from the xml-file by specifying OVERWRITE=%QUOTE(macro parameter1=value1, macro parameter2=value2,…, macro parameterN=valueN), e.g. if you quickly want to test the effect of a parameter setting without having to change the values in the xml-file.

%SETTABS creates two temporary SAS datasets from Sheet1 and Sheet2 of the xml-file via the Macro %XLXP2SAS. Then it calls the corresponding macro version of %GEEWHIZ that is specified in the xml-file. One call of %SETTABS corresponds to one table.

THE MACRO %GEEWHIZ

The macro %GEEWHIZ is usually only called automatically via %SETTABS and the user does not have to care about it. To implement this, specify MACROTYPE=GEEWHIZ inside the xml-file. Inside the xml-File you can refer to different versions of the macro. The macro %GEEWHIZ combines the information of Sheet1 and the corresponding column information from Sheet2 (via the link variable COLUMNTYPE). It then creates the table using the macro parameter specifications from the xml-File. If the macro %GEEWHIZ is used as intended it creates a PDF-file for each table. The internals of the %GEEWHIZ macro are built around proc report and ods PDF.

PREPARATION OF THE DATA

Sometimes the datasets that are used as macro parameter inside the xml-interface need to be prepared.

Usually two data files are needed to create a table with %GEEWHIZ: The first SAS dataset (e.g. ae) contains the subjects and corresponding events to be counted. In the following this will be referred to as event dataset. The second SAS dataset (e.g. disposit or demo) contains the number of subjects under risk categorized by group. In the following, this will be referred to as disposit dataset. Both datasets are merged during the process of table creation.

Prior to calling %SETTABS these two (temporary) files sometimes have to be prepared by

- Deriving variables in the event dataset to be used in where-conditions or as sorting variables
- Slicing the event dataset into by-groups (e.g. Week at onset).
- Create the disposit dataset from multiple sources in case of overlapping groups (e.g. the additional use of an exposure dataset for dose at onset analyses if a subject may be at risk in multiple groups).
- Adding by-group variables to the event dataset (in case the by-group variable is a merge variable).
- Duplicating the records of the disposit dataset in case of overlapping by-groups.
- Using additional SAS-dataset to determine subjects at risk (e.g. in particular periods).
- Derive variables that are used to specify a particular aspect of the layout (e.g. a variable that indicates where to include an empty line)
- Derive numeric sort variables if you want to bring character values in a pre-defined order.

In many cases you can –as a matter of choice - alternatively address data manipulation issues by preparing the files or by choosing the macro parameters appropriately. You can for example subset your event dataset prior to calling %SETTABS or you can do it by specifying a specific macro parameter.

A general advice is hard to give, as the more efficient way of handling things depends on the context. A good practical example of which pre-processing steps can be necessary is shown in Example 6 below.

A BRIEF DESCRIPTION OF MACRO PARAMETERS

This paper can only provide a selection of the functionalities. The functions could be divided into:

- Layout (such as column widths, empty lines, display of frequencies and percentages).
- Content (data selection, subsetting etc.)
- Text labels of columns and categories (e.g. text that appears in the header of the table).

Table 1 and Table 2 provide a brief overview of the general functions of the macro parameters used inside %GEEWHIZ. In order to address many different aspects of the data and requirements of the customers additional options, features and functions can be specified for most of the macro parameters. These will usually be appended by commas and will internally be split up into multiple macro parameters. You can for example add an additional variable that is used for sorting to the first level variable (see parameter LEV1).

Table 1: Brief description of most of the parameters in sheet1 and their basic functions

	Parameter	Brief description
1	Keyword	Variable that identifies the entry and matches with the entry from the file containing the titles
2	Skip	Controls page breaks and allows you to enter blank lines
3	Macrotype	This is the macro version that is used to process the data
4	Usubjid	Usually the subject identifier. Could be a combination of variables if a subject falls into multiple treatment groups (e.g. dose at onset of adverse event)
5	Dset1	The SAS-dataset which contains the relevant information to be tabulated (“event” dataset).
6	D1_subset1	Contains a where condition for SAS-dataset DSET1 to subset the data
7	D1_subset2	Special subsetting of DSET1, e.g. to determine the numbers at risk or to exclude certain by-groups
8	Byvar	Determines the variable (usually in DSET2) that is used for by-group processing
9	Lev1-Lev3	Up to three variables used to display the different levels (e.g. soc, hlt, and pt). Each new level is indented by two additional blank characters
10	Levterm	This is the column header of the first column of the table. Use “|” as split character to wrap lines.
11	Lev1txt	The text of this parameter is displayed on top of the table (non-indented) as summary for the 1st level variable. If text1 is omitted, no summary for the 1st level variable is used.
12	Col1min, Col1max	This is the minimum (maximum) width in cm reserved for the first column of the table, which contains the first, second and sometimes a 3rd level variable. If no value is specified COL1MIN is automatically set to 20 cm
13	Subgroup	This is the subgroup variable (optional) whose values are displayed in the second column of the table. The values of the subgroup variable are defined in → SUBVALS.
14	Subgrouplabel	This is the text for the column header of the subgroup variable (e.g. relationship, intensity, maximum intensity)
15	Subvals	The comma-separated list of values usually determines the values for the subgroup column.
16	Subvalorder	Determines the order of priority for SUBVALS in the presence of a value for SUBGROUP.
17	Dset2	The SAS-dataset which contains the treatment group information and information concerning the populations (e.g. disposition dataset).
18	D2_subset	Contains a where condition for SAS-dataset DSET2 to subset the disposition dataset.
19	Header_display_n	This determines whether of not the N’s should be displayed in the header block of the table
20	Selectrows	This parameter allows you to exclude rows from the table display without deleting data from the dataset that is to be displayed.
21	Overall	If this parameter is not blank, it determines the text for the Overall column(s) on the right side of the table. If the parameter is blank, no overall column is created.
22	Orderbytotal	Determines the rank of the overall column in specifying the sort order of the 2nd and 3rd level variable in the first column.
23	Columntype	Determines the appearance of the treatment columns and refers the same variable in sheet2 of the xml-file.
24	Displaytype	Determines how frequencies and percentages and events are displayed
25	Percentformat	Determines the number of digits used in the percentage display.
26	Minimumwidth	Allow you to switch between a standard column width and a minimum column width to display the treatment groups.
27	Spanningtext	This optional text is placed on top of the treatment columns as spanning text
28	Showbody	Determines whether the frequencies and percentages are summarized for the 1st and 2nd level.
29	Cretable	Indicates whether or not the PDF-file should be produced (for validaton programmers a PDF is not required)
30	Newtarget	Allows you to start e.g. an automated compare.

Table 2: Description of the parameters in Sheet2

Parameter	Brief description
Columntype	Integer Number for the table that matches with the corresponding value of sheet1
Var	This is the name of variable that is used for grouping the data into the columns. This is usually a numeric treatment group variable. Character variable are also allowed if each distinct value is displayed as column. It is also possible to specify a dose variable if each column should represent a single or combined dose class.
Start	Determines the start value for the category. The categories must be mutually exclusive.
End	Determines the end value for the category. To include only a single value into a groups END should be the same value as START. The categories must be mutually exclusive. The overall column(s) are created inside sheet1.
Sortrank	Determines the rank of the column in specifying the sort order of the 2nd level variable (first column). The provided numbers must be unique but do not need to be successive. If a group should be excluded from the active treatment group total add: ",x" or any other character.
Label	Provides the label for the column that appears in the header of the table.

AUTOMATED COMPARE

One advantage of the xml-user interface is that a large number of tables can be handled efficiently. That is way the macro is often used to create a large number of tables with just one program.

The macro %GEEWHIZ can be used for development and validation of tables. For validation programmers (in case of double programming) it offers the possibility to perform an automated comparison to the developed file provided that some requirements are met (e.g. the variable names of the compared variables need to be equal). The requirements are met if both the development and the validation of a table were done with the macro.

With the automated compare (→ parameter NEWTARGET in sheet 1) the comparison procedure is invoked for a table and the *.lst outputs created by the compare procedure are concatenated. If the number of tables is large this yields a big file that can be tedious to check for findings.

An additional SAS macro was designed to extract the relevant information form the *.lst file and summarize the results in an xml-file. It is optionally called as the last step of the validation program, i.e. after the *.lst was already created.

EXAMPLES

The adverse event data presented in the following examples is completely fictional and only serves the purpose to show particular test cases and to illustrate how the macro % GEEWHIZ handles these specific test cases. The macro is particularly suited to process adverse events data. The last example shows its application to real-life laboratory data.

This chapter can only cover a couple of different fields of application. If shows which changes would be necessary if you are starting from a base case. Each example will be started with an overview of the changes in the parameter setting and followed by the table that would be generated with the parameter setting. In the tables a couple of comments are inserted to further explain the effect of a certain parameter specification.

EXAMPLE 1

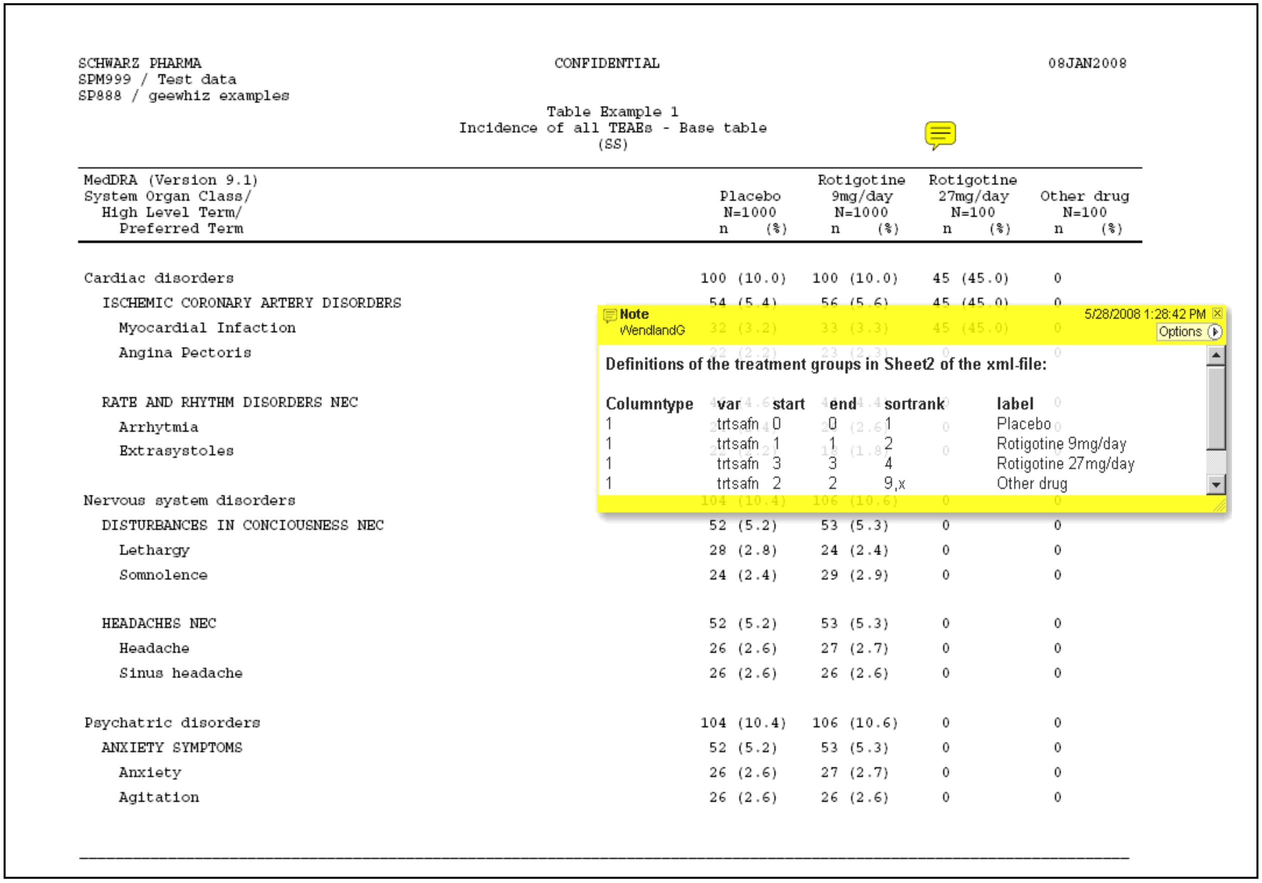
The following macro parameters would be specified for a "standard" table. In the following parameter overview the columns of sheets1 are expressed as rows:

Table 3: Parameter settings in sheet 1 for a standard table

Column	Example 1	Incidence of all Treatment-Emergent Adverse Events (TEAEs)
1	Keyword	example 1
3	macrotype	geewhiz_v1_5
4	usubjid	usubjid
5	Dset1	ae
6	D1_subset1	ss_g="Yes"
9	lev1	socdm
9	lev2	hltdm
9	lev3	ptdm
10	levterm	MedDRA (V9.1)| System Organ Class/| High Level Term/| Preferred Term
12	col1min	4
12	col1max	15
17	Dset2	disposit
18	D2_subset	ss_g="Yes"
23	Columntype	1
24	Displaytype	1
25	Percentformat	percnew
26	Minimumwidth	yes
28	showbody	yes
29	cretable	yes

This parameter setting yields the following result:

Figure 3: A "Standard" table



```
SCHWARZ PHARMA                              CONFIDENTIAL                                  08JAN2008
SPM999 / Test data
SP888 / geewhiz examples
                                          Table Example 1
                                 Incidence of all TEAEs - Base table
                                                (SS)
__________________________________________________________________________________________________
MedDRA (Version 9.1)                                          Rotigotine   Rotigotine
System Organ Class/                                 Placebo    9mg/day     27mg/day    Other drug
  High Level Term/                                   N=1000     N=1000       N=100       N=100
    Preferred Term                                  n     (%)  n     (%)   n     (%)   n     (%)
__________________________________________________________________________________________________

Cardiac disorders                                  100 (10.0)  100 (10.0)  45 (45.0)   0
  ISCHEMIC CORONARY ARTERY DISORDERS                54 (5.4)    56 (5.6)   45 (45.0)   0
    Myocardial Infaction                            32 (3.2)    33 (3.3)   45 (45.0)   0
    Angina Pectoris                                 22 (2.2)    23 (2.3)    0          0

  RATE AND RHYTHM DISORDERS NEC                     46 (4.6)    44 (4.4)    0          0
    Arrhytmia                                       24 (2.4)    26 (2.6)    0          0
    Extrasystoles                                   22 (2.2)    18 (1.8)    0          0

Nervous system disorders                           104 (10.4)  106 (10.6)   0          0
  DISTURBANCES IN CONCIOUSNESS NEC                  52 (5.2)    53 (5.3)    0          0
    Lethargy                                        28 (2.8)    24 (2.4)    0          0
    Somnolence                                      24 (2.4)    29 (2.9)    0          0

  HEADACHES NEC                                     52 (5.2)    53 (5.3)    0          0
    Headache                                        26 (2.6)    27 (2.7)    0          0
    Sinus headache                                  26 (2.6)    26 (2.6)    0          0

Psychatric disorders                               104 (10.4)  106 (10.6)   0          0
  ANXIETY SYMPTOMS                                  52 (5.2)    53 (5.3)    0          0
    Anxiety                                         26 (2.6)    27 (2.7)    0          0
    Agitation                                       26 (2.6)    26 (2.6)    0          0

__________________________________________________________________________________________________
```

Definitions of the treatment groups in Sheet2 of the xml-file:

Columntype	var	start	end	sortrank	label
1	trtsafn	0	0	1	Placebo
1	trtsafn	1	1	2	Rotigotine 9mg/day
1	trtsafn	3	3	4	Rotigotine 27mg/day
1	trtsafn	2	2	9,x	Other drug

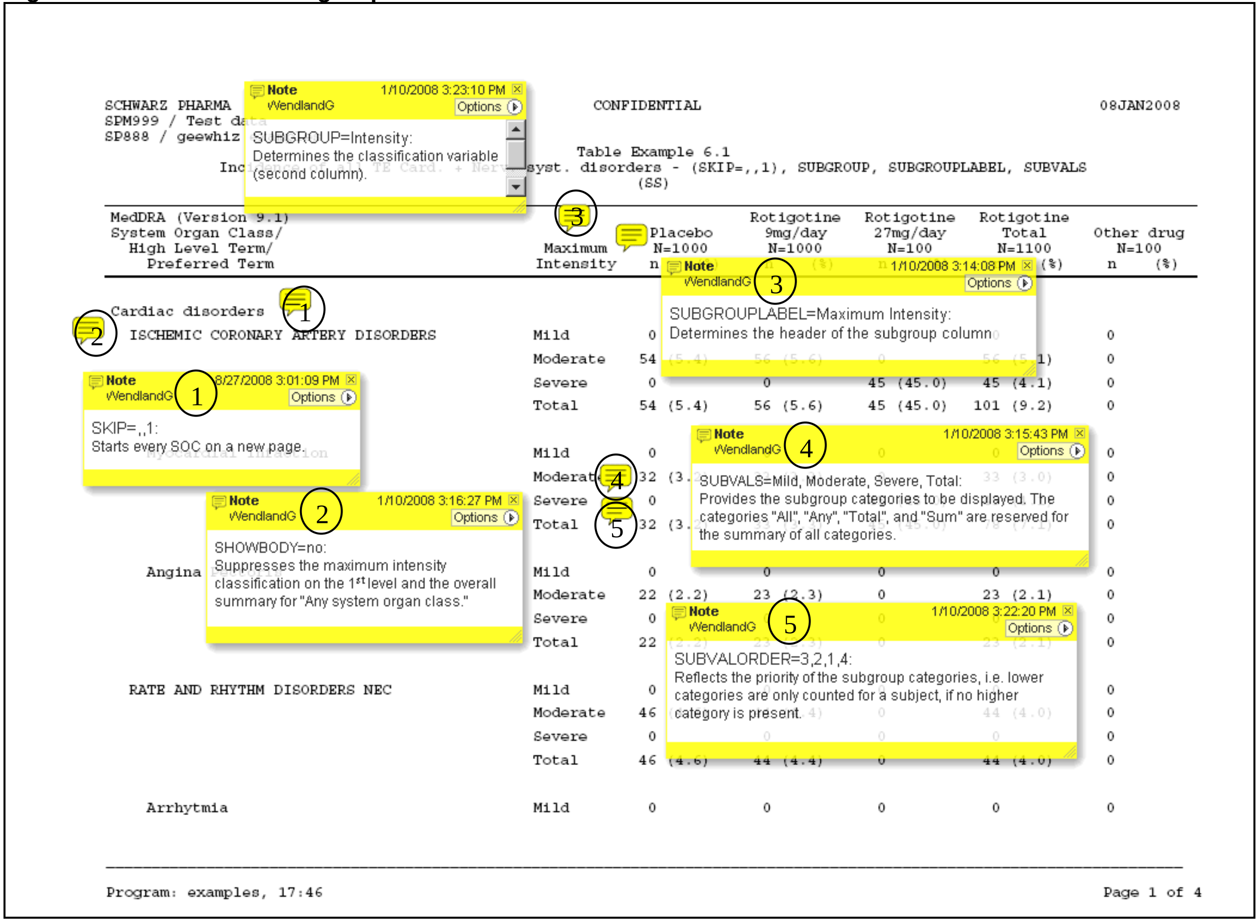
EXAMPLE 2

Now the adverse events are to be categorized by maximum intensity via an additional subgroup variable. Table 4 highlights the changes as compared to the base case

Table 4: Inclusion of a subgroup variable to categorize intensity

	Table 6_1	Incidence of all TE Card. + Nerv. syst. disorders - (SKIP=,,1), SUBGROUP, SUBGROUPLABEL, SUBVALS
1	Keyword	Example6_1
2	Skip	,,1
5	Dset1	ae1
6	D1_subset1	ss_g="yes" and substr(socdm,1,1) in ("C","N")
11	Lev1txt	Any System Organ Class
13	Subgroup	Intensity
14	Subgrouplabel	Maximum Intensity
15	Subvals	Mild, Moderate, Severe, Total
16	Subvalorder	3,2,1,4
21	Overall	,Rotigotine Total,1
28	Showbody	no

Figure 4: A table with a subgroup column



```
SCHWARZ PHARMA                                  CONFIDENTIAL                                    08JAN2008
SPM999 / Test data
SP888 / geewhiz
                                                 Table Example 6.1
          Incidence of all TE Card. + Nerv. syst. disorders - (SKIP=,,1), SUBGROUP, SUBGROUPLABEL, SUBVALS
                                                        (SS)
 ---------------------------------------------------------------------------------------------------------
 MedDRA (Version 9.1)                                               Rotigotine  Rotigotine  Rotigotine
 System Organ Class/                                     Placebo     9mg/day     27mg/day     Total      Other drug
   High Level Term/                            Maximum    N=1000      N=1000      N=100       N=1100       N=100
     Preferred Term                           Intensity   n   (%)     n   (%)     n   (%)     n   (%)     n   (%)
 ---------------------------------------------------------------------------------------------------------

 Cardiac disorders
   ISCHEMIC CORONARY ARTERY DISORDERS          Mild        0                                               0
                                               Moderate   54 (5.4)   56 (5.6)     0          56 (5.1)     0
                                               Severe      0          0          45 (45.0)   45 (4.1)     0
                                               Total      54 (5.4)   56 (5.6)    45 (45.0)  101 (9.2)     0

     Myocardial infarction                     Mild        0                                               0
                                               Moderate   32 (3.2)                           33 (3.0)     0
                                               Severe      0                                               0
                                               Total      32 (3.2)   33 (3.3)    45 (45.0)   78 (7.1)     0

     Angina Pectoris                           Mild        0          0           0           0           0
                                               Moderate   22 (2.2)   23 (2.3)     0          23 (2.1)     0
                                               Severe      0                                               0
                                               Total      22 (2.2)   23 (2.3)     0          23 (2.1)     0

   RATE AND RHYTHM DISORDERS NEC               Mild        0                                               0
                                               Moderate   46 (4.6)   44 (4.4)     0          44 (4.0)     0
                                               Severe      0          0           0           0           0
                                               Total      46 (4.6)   44 (4.4)     0          44 (4.0)     0

     Arrhytmia                                 Mild        0          0           0           0           0

 _________________________________________________________________________________________________________
 Program: examples, 17:46                                                                     Page 1 of 4
```

Note 1/10/2008 3:23:10 PM WendlandG
SUBGROUP=Intensity:
Determines the classification variable (second column).

Note 1/10/2008 3:14:08 PM WendlandG (3)
SUBGROUPLABEL=Maximum Intensity:
Determines the header of the subgroup column

Note 8/27/2008 3:01:09 PM WendlandG (1)
SKIP=,,1:
Starts every SOC on a new page.

Note 1/10/2008 3:16:27 PM WendlandG (2)
SHOWBODY=no:
Suppresses the maximum intensity classification on the 1st level and the overall summary for "Any system organ class."

Note 1/10/2008 3:15:43 PM WendlandG (4)
SUBVALS=Mild, Moderate, Severe, Total:
Provides the subgroup categories to be displayed. The categories "All", "Any", "Total", and "Sum" are reserved for the summary of all categories.

Note 1/10/2008 3:22:20 PM WendlandG (5)
SUBVALORDER=3,2,1,4:
Reflects the priority of the subgroup categories, i.e. lower categories are only counted for a subject, if no higher category is present.

The parameter SHOWBODY=no suppresses the display of frequencies and percentages for the first level entries. It also suppresses the display of the summary of the first level entries (e.g. System Organ Class) specified by LEV1TXT=Any system organ class (see also the next example where this summary line is displayed). This example shows that sometimes parameters interact with each other.

The parameters SUBGROUP, SUBGROUPLABEL, SUBVALS, and SUBVALORDER are often used in combination. Specifying a value for SUBVALORDER implies that the categories follow a priority: If a subject experiences an adverse event twice the lower category would not be counted.

The parameter OVERALL is starting with a comma, which means that no overall total is calculated but an active treatment total. The adjacent "1" indicates that the last group is to be excluded from the active treatment total.

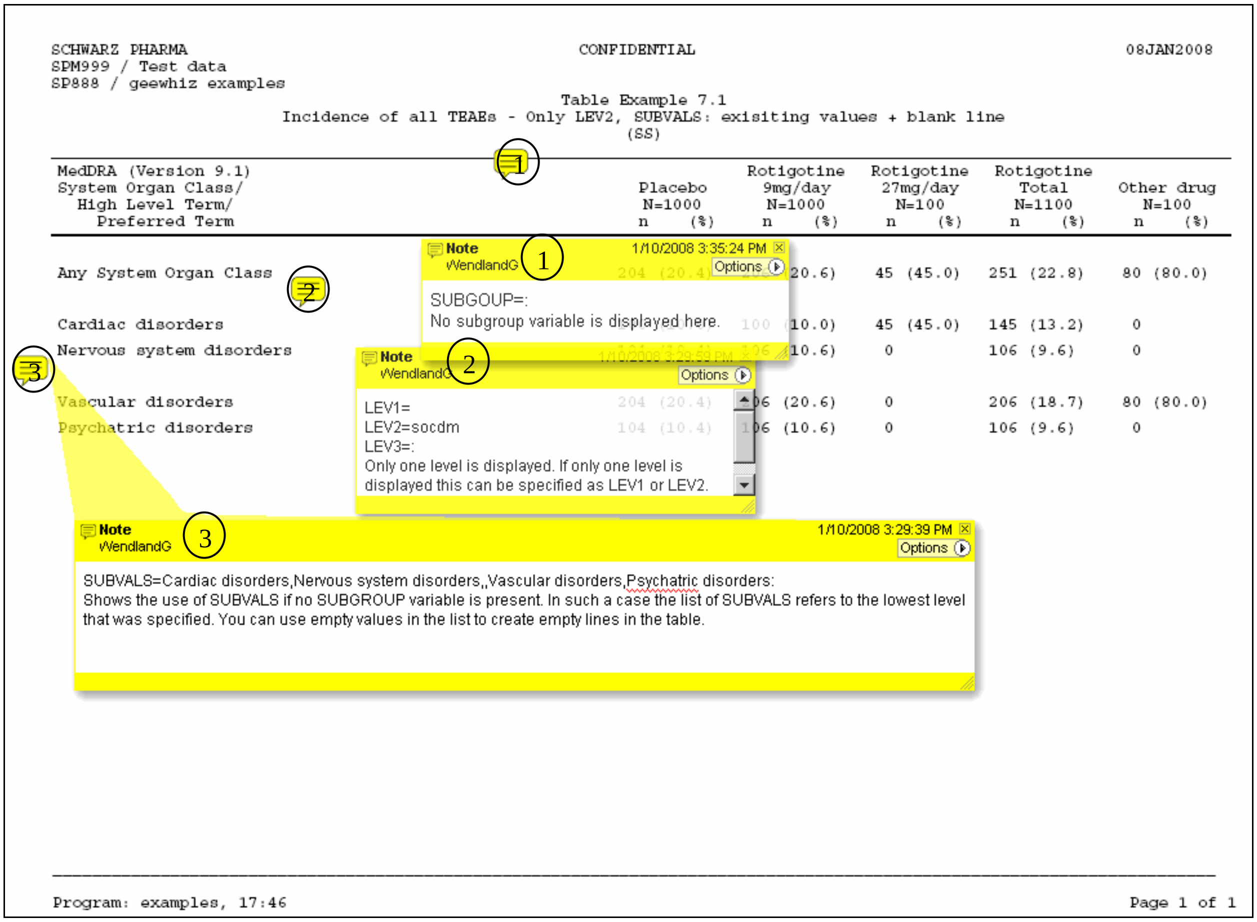
EXAMPLE 3

This example shows another case where parameters interact with each other. Here the parameter SUBVALS could also be used without SUBGROUP:

Table 5: Use of parameter SUBVALS without parameter SUBGROUP

	Table 7_1	Incidence of all TEAEs - Only LEV2, SUBVALS: existing values + blank line
1	Keyword	Example7_1
5	Dset1	ae1
9	Lev1	
9	Lev2	Socdm
9	Lev3	
11	Lev1txt	Any System Organ Class
13	Subgroup	
15	Subvals	Cardiac disorders,Nervous system disorders,,Vascular disorders,Psychatric disorders
21	Overall	,Rotigotine Total,1

Figure 5: Using the SUBVALS parameter to manipulate the first column



```
SCHWARZ PHARMA                                 CONFIDENTIAL                                 08JAN2008
SPM999 / Test data
SP888 / geewhiz examples
                                              Table Example 7.1
                       Incidence of all TEAEs - Only LEV2, SUBVALS: exisiting values + blank line
                                                    (SS)
_________________________________________________________________________________________________________
 MedDRA (Version 9.1)                                            Rotigotine  Rotigotine  Rotigotine
 System Organ Class/                                 Placebo      9mg/day    27mg/day      Total     Other drug
   High Level Term/                                  N=1000       N=1000      N=100       N=1100       N=100
     Preferred Term                                  n    (%)     n    (%)    n    (%)    n    (%)     n    (%)
_________________________________________________________________________________________________________

 Any System Organ Class                           204  (20.4)       20.6)    45  (45.0)   251  (22.8)   80  (80.0)

 Cardiac disorders                                             100  10.0)    45  (45.0)   145  (13.2)    0
 Nervous system disorders                                      106  10.6)    0            106  (9.6)     0

 Vascular disorders                               204  (20.4)  206  (20.6)   0            206  (18.7)   80  (80.0)
 Psychatric disorders                             104  (10.4)  106  (10.6)   0            106  (9.6)     0

_________________________________________________________________________________________________________
Program: examples, 17:46                                                                      Page 1 of 1
```

Note 1 — WendlandG — 1/10/2008 3:35:24 PM — Options

SUBGOUP=:
No subgroup variable is displayed here.

Note 2 — WendlandG — 1/10/2008 3:29:59 PM — Options

LEV1=
LEV2=socdm
LEV3=:
Only one level is displayed. If only one level is displayed this can be specified as LEV1 or LEV2.

Note 3 — WendlandG — 1/10/2008 3:29:39 PM — Options

SUBVALS=Cardiac disorders,Nervous system disorders,,Vascular disorders,Psychatric disorders:
Shows the use of SUBVALS if no SUBGROUP variable is present. In such a case the list of SUBVALS refers to the lowest level that was specified. You can use empty values in the list to create empty lines in the table.

EXAMPLE 4

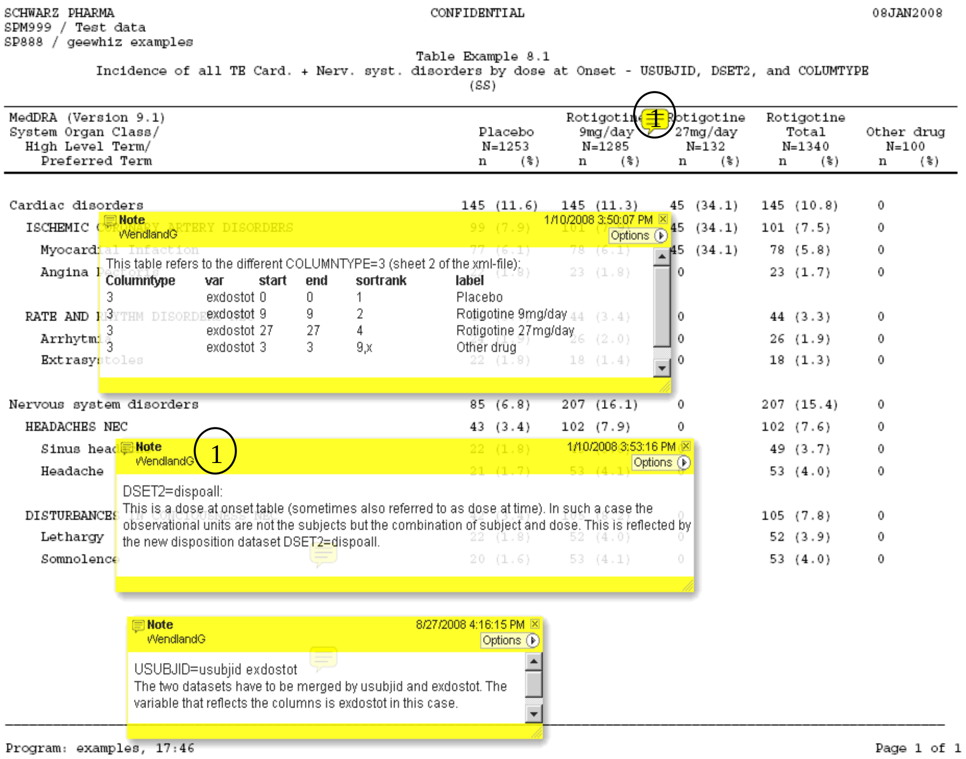
Sometimes treatment groups are not mutually exclusive, e.g. if you want to display adverse events by dose at onset (variable exdostot), a subject might experience the same adverse at different doses if the subject is titrated to a maintenance dose.

Table 6: Specifications if treatment columns are not mutually exclusive

	Table 8_1	Incidence of all TE Card. + Nerv. syst. disorders by dose at Onset - USUBJID, DSET2, and COLUMTYPE
1	Keyword	Example8_1
4	Usubjid	Usubjid exdostot
5	Dset1	ae1
6	D1_subset1	ss_g=”yes” and substr(socdm,1,1) in (“C”,”N”)
17	Dset2	dispoall
21	Overall	,Rotigotine Total,1
23	Columntype	3

As the dose at onset is used here sheet 1 refers to a different value of COLUMNTYPE in sheet 2 (see specification in the figure below).

Figure 6: Using a combination of two variables to merge the “event” dataset with the “disposition” dataset.



```
SCHWARZ PHARMA                                  CONFIDENTIAL                                   08JAN2008
SPM999 / Test data
SP888 / geewhiz examples
                                              Table Example 8.1
        Incidence of all TE Card. + Nerv. syst. disorders by dose at Onset - USUBJID, DSET2, and COLUMTYPE
                                                    (SS)
---------------------------------------------------------------------------------------------------------
MedDRA (Version 9.1)                                           Rotigotine  Rotigotine  Rotigotine
System Organ Class/                                 Placebo     9mg/day    27mg/day     Total      Other drug
  High Level Term/                                  N=1253      N=1285      N=132       N=1340      N=100
    Preferred Term                                  n    (%)    n    (%)    n    (%)    n    (%)    n    (%)
---------------------------------------------------------------------------------------------------------

Cardiac disorders                                 145 (11.6)  145 (11.3)   45 (34.1)  145 (10.8)    0
  ISCHEMIC CORONARY ARTERY DISORDERS               99 (7.9)   101 (7.9)    45 (34.1)  101 (7.5)     0
    Myocardial Infarction                          77 (6.1)    78 (6.1)    45 (34.1)   78 (5.8)     0
    Angina Pectoris                                22 (1.8)    23 (1.8)     0          23 (1.7)     0

  RATE AND RHYTHM DISORDERS NEC                    [illegible]  44 (3.4)     0          44 (3.3)     0
    Arrhytmia                                      24 (1.9)    26 (2.0)     0          26 (1.9)     0
    Extrasystoles                                  22 (1.8)    18 (1.4)     0          18 (1.3)     0

Nervous system disorders                           85 (6.8)   207 (16.1)    0         207 (15.4)    0
  HEADACHES NEC                                    43 (3.4)   102 (7.9)     0         102 (7.6)     0
    Sinus headache                                 22 (1.8)   [illegible]               49 (3.7)     0
    Headache                                       21 (1.7)    53 (4.1)                53 (4.0)     0

  DISTURBANCES IN CONCIOUSNESS NEC                 42 (3.4)   105 (8.2)     0         105 (7.8)     0
    Lethargy                                       22 (1.8)    52 (4.0)     0          52 (3.9)     0
    Somnolence                                     20 (1.6)    53 (4.1)     0          53 (4.0)     0

---------------------------------------------------------------------------------------------------------
Program: examples, 17:46                                                                   Page 1 of 1
```

Note
WendlandG
1/10/2008 3:50:07 PM

This table refers to the different COLUMNTYPE=3 (sheet 2 of the xml-file):

Columntype	var	start	end	sortrank	label
3	exdostot 0	0		1	Placebo
3	exdostot 9	9		2	Rotigotine 9mg/day
3	exdostot 27	27		4	Rotigotine 27mg/day
3	exdostot 3	3		9,x	Other drug

Note
WendlandG
1/10/2008 3:53:16 PM

DSET2=dispoall:
This is a dose at onset table (sometimes also referred to as dose at time). In such a case the observational units are not the subjects but the combination of subject and dose. This is reflected by the new disposition dataset DSET2=dispoall.

Note
WendlandG
8/27/2008 4:16:15 PM

USUBJID=usubjid exdostot
The two datasets have to be merged by usubjid and exdostot. The variable that reflects the columns is exdostot in this case.

EXAMPLE 5

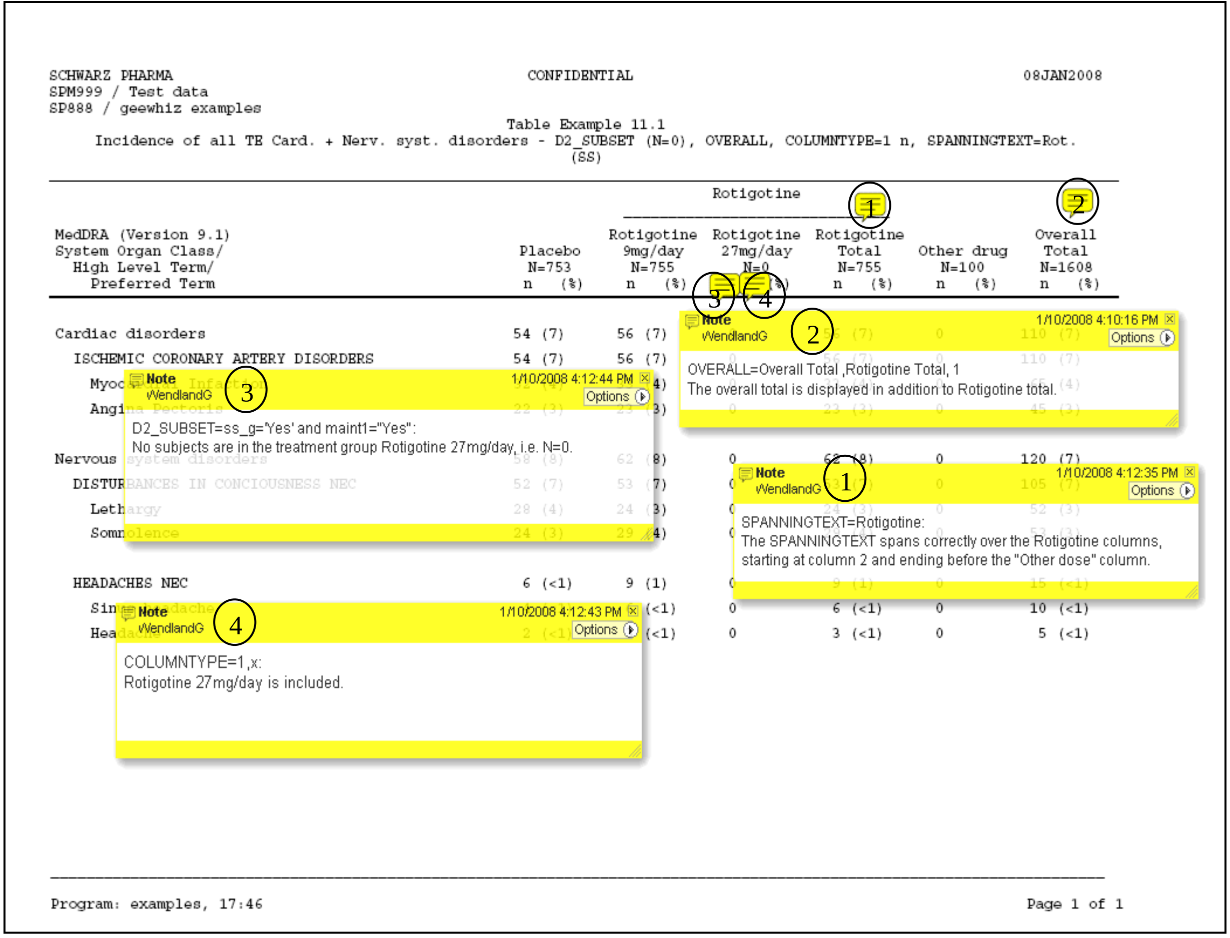
If you want to group column together (e.g. if you have multiple doses of the same active treatment) it makes sense to use a spanning header to indicate that certain columns belong together. Furthermore, it is shown how to include treatment columns into the table that usually would disappear because no subject is in that treatment group. In the following example this occurs because it is restricted to subjects that reach a certain period in the study.

Table 7: Implementation of a spanning header and inclusion of columns with N=0

	Table 11_1	Incidence of all TE Card. + Nerv. syst. disorders - D2_SUBSET (N=0), OVERALL, COLUMNTYPE=1 n, SPANNINGTEXT=Rot.
1	Keyword	Example11_1
5	Dset1	ae1
6	D1_subset1	ss_g=”yes” and substr(socdm,1,1) in (“C”,”N”)
18	D2_subset	ss_g='Yes' and maint1="Yes"
21	Overall	Overall Total ,Rotigotine Total, 1
23	Columntype	1,x
25	Percentformat	Perc
27	Spanningtext	Rotigotine

Here the parameter OVERALL interacts automatically with the parameter SPANNINGTEXT. The text indicates that the overall total and the active treatment total (i.e. Rotigotine total) are to be displayed. Furthermore a “1” is added to indicate that one additional column apart from Placebo is excluded from the active treatment total. This is recognized by the parameter SPANNINGTEXT which begins at the 2nd treatment column (default) and ends with the column “Rotigotine Total”. Percentages are displayed here without any decimals to save space (PERCENTFORMAT=perc). The setting COLUMNTYPE=1,x ensures that all (intended) treatment groups are displayed even if no subjects fall into a group.

Figure 7: Using a spanning header and including treatment columns with no subjects (N=0).



```
SCHWARZ PHARMA                                   CONFIDENTIAL                                    08JAN2008
SPM999 / Test data
SP888 / geewhiz examples
                                                  Table Example 11.1
     Incidence of all TE Card. + Nerv. syst. disorders - D2_SUBSET (N=0), OVERALL, COLUMNTYPE=1 n, SPANNINGTEXT=Rot.
                                                        (SS)
____________________________________________________________________________________________________________
                                                                    Rotigotine
                                                         ____________________________________
MedDRA (Version 9.1)                                     Rotigotine  Rotigotine  Rotigotine               Overall
System Organ Class/                             Placebo    9mg/day    27mg/day     Total     Other drug    Total
  High Level Term/                               N=753      N=755       N=0        N=755       N=100      N=1608
    Preferred Term                              n   (%)    n   (%)    n   (%)    n   (%)    n   (%)    n   (%)
____________________________________________________________________________________________________________

Cardiac disorders                               54  (7)    56  (7)    0          56  (7)    0          110  (7)
  ISCHEMIC CORONARY ARTERY DISORDERS            54  (7)    56  (7)    0          56  (7)    0          110  (7)
    Myocardial Infarction                       32  (4)    33  (4)    0          33  (4)    0           65  (4)
    Angina Pectoris                             22  (3)    23  (3)    0          23  (3)    0           45  (3)

Nervous system disorders                        58  (8)    62  (8)    0          62  (8)    0          120  (7)
  DISTURBANCES IN CONCIOUSNESS NEC              52  (7)    53  (7)    0          53  (7)    0          105  (7)
    Lethargy                                    28  (4)    24  (3)    0          24  (3)    0           52  (3)
    Somnolence                                  24  (3)    29  (4)    0          29  (4)    0           53  (3)

  HEADACHES NEC                                  6 (<1)     9  (1)    0           9  (1)    0           15 (<1)
    Sinus headache                               4 (<1)     6 (<1)    0           6 (<1)    0           10 (<1)
    Headache                                     2 (<1)     3 (<1)    0           3 (<1)    0            5 (<1)

____________________________________________________________________________________________________________
Program: examples, 17:46                                                                         Page 1 of 1
```

Note WendlandG 1/10/2008 4:12:35 PM

SPANNINGTEXT=Rotigotine:
The SPANNINGTEXT spans correctly over the Rotigotine columns, starting at column 2 and ending before the "Other dose" column.

Note WendlandG 1/10/2008 4:10:16 PM

OVERALL=Overall Total ,Rotigotine Total, 1
The overall total is displayed in addition to Rotigotine total.

Note WendlandG 1/10/2008 4:12:44 PM

D2_SUBSET=ss_g='Yes' and maint1="Yes":
No subjects are in the treatment group Rotigotine 27mg/day, i.e. N=0.

Note WendlandG 1/10/2008 4:12:43 PM

COLUMNTYPE=1,x:
Rotigotine 27mg/day is included.

EXAMPLE 6

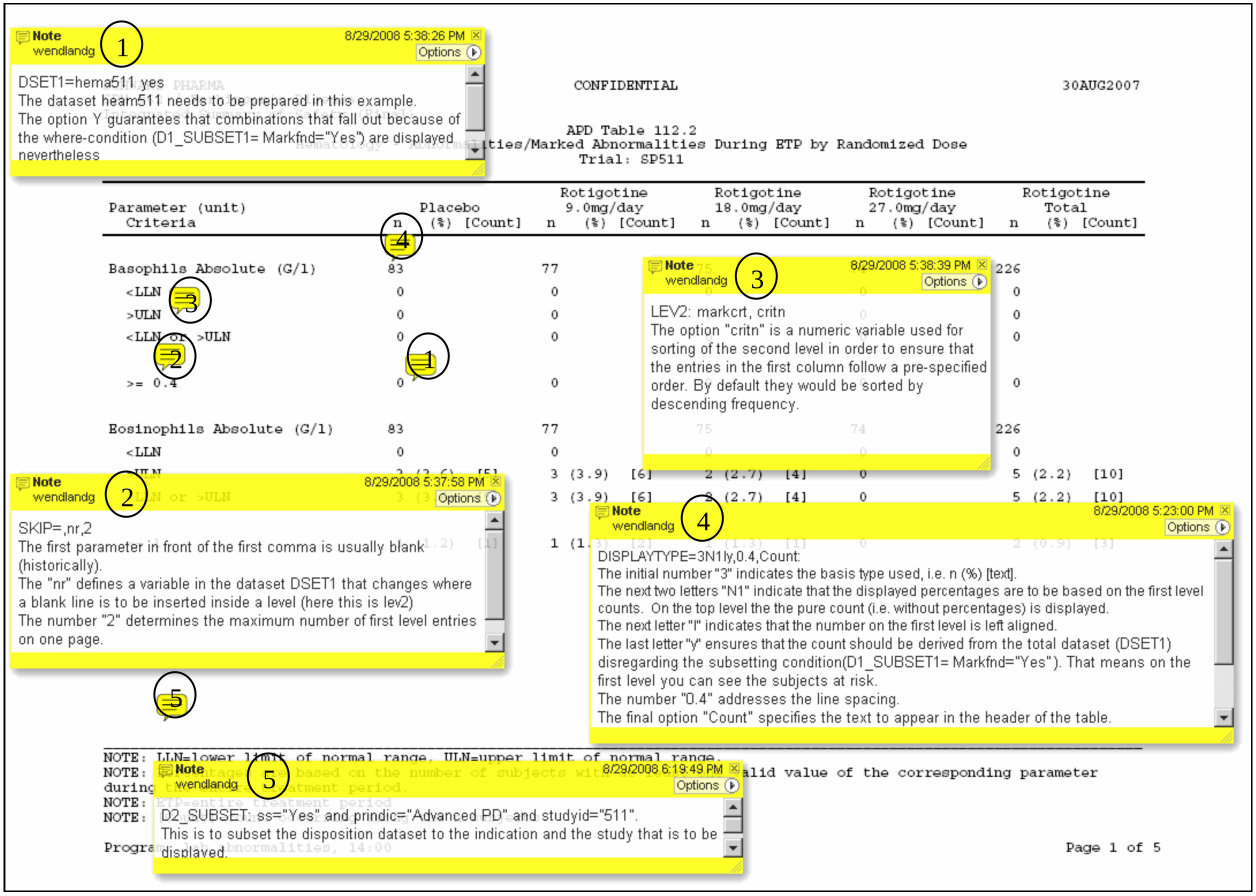
The last example handles a more complex real-life situation and shows that %GEEWHIZ is not only restricted to adverse events data. This is a more sophisticated use of the macro.

The used data needs to be pre-processed here (usually in the SAS program that contains the %SETTABS invocations) in order to generate the variable “critn” to sort LEV2, to create a variable “nr” that defines empty lines between the abnormalities and the marked abnormalities, and to generate a dataset “hema511” that only contains the hematology parameters of the study 511. The latter is required because the full input data (disregarding the subsetting condition D1_SUBSET1) is used to determine the subject counts (see comment 4 below).

Table 8: Implementation of special laboratory table

	Table 112_2	Hematology – Abnormalities/Marked Abnormalities
1	Keyword	iss_ap_lab_hem_rd_511
2	Skip	,nr,2
5	Dset1	Hema511 yes
6	D1_subset1	Markfnd=”Yes”
9	Lev1	lbtest
9	Lev2	Markcrt, critn
10	Levterm	Parameter (unit)| Criteria
17	Dset2	Subject
18	D2_subset	ss='Yes' and prindic = 'Advanced PD' and studyid='511'
21	Overall	,Rotigotine Total
24	Displaytype	3N1ly,0.4,Count

Figure 8: Displaying abnormalities/ marked abnormalities with special features



```
                                   CONFIDENTIAL                                              30AUG2007

                                  APD Table 112.2
                        ...ties/Marked Abnormalities During ETP by Randomized Dose
                                   Trial: SP511

                                            Rotigotine           Rotigotine          Rotigotine          Rotigotine
Parameter (unit)              Placebo        9.0mg/day           18.0mg/day          27.0mg/day            Total
  Criteria                n   (%)  [Count]   n   (%)  [Count]   n   (%)  [Count]   n   (%)  [Count]   n   (%)  [Count]

Basophils Absolute (G/l)  83                77                                                       226
  <LLN                     0                 0                                                         0
  >ULN                     0                 0                                                         0
  <LLN or >ULN             0                 0                                                         0

  >= 0.4                   0                 0                                                         0

Eosinophils Absolute (G/l) 83               77                                                       226
  <LLN                     0                 0                                                         0
                                             3  (3.9)   [6]      2  (2.7)   [4]     0                  5  (2.2)   [10]
                                             3  (3.9)   [6]      2  (2.7)   [4]     0                  5  (2.2)   [10]

                                             1  (1.

NOTE: LLN=lower limit of normal range, ULN=upper limit of normal range.
NOTE:                                                            ...alid value of the corresponding parameter
during
NOTE:
NOTE:

Progra                                                                                            Page 1 of 5
```

Note wendlandg 8/29/2008 5:38:26 PM (1)

DSET1=hema511 yes
The dataset heam511 needs to be prepared in this example.
The option Y guarantees that combinations that fall out because of the where-condition (D1_SUBSET1= Markfnd="Yes") are displayed nevertheless

Note wendlandg 8/29/2008 5:38:39 PM (3)

LEV2: markcrt, critn
The option "critn" is a numeric variable used for sorting of the second level in order to ensure that the entries in the first column follow a pre-specified order. By default they would be sorted by descending frequency.

Note wendlandg 8/29/2008 5:37:58 PM (2)

SKIP=,nr,2
The first parameter in front of the first comma is usually blank (historically).
The "nr" defines a variable in the dataset DSET1 that changes where a blank line is to be inserted inside a level (here this is lev2)
The number "2" determines the maximum number of first level entries on one page.

Note wendlandg 8/29/2008 5:23:00 PM (4)

DISPLAYTYPE=3N1ly,0.4,Count:
The initial number "3" indicates the basis type used, i.e. n (%) [text].
The next two letters "N1" indicate that the displayed percentages are to be based on the first level counts. On the top level the the pure count (i.e. without percentages) is displayed.
The next letter "l" indicates that the number on the first level is left aligned.
The last letter "y" ensures that the count should be derived from the total dataset (DSET1) disregarding the subsetting condition(D1_SUBSET1= Markfnd="Yes"). That means on the first level you can see the subjects at risk.
The number "0.4" addresses the line spacing.
The final option "Count" specifies the text to appear in the header of the table.

Note wendlandg 8/29/2008 6:19:49 PM (5)

D2_SUBSET: ss="Yes" and prindic="Advanced PD" and studyid="511".
This is to subset the disposition dataset to the indication and the study that is to be displayed.

INTERACTIONS WITH OTHER MACROS

The invocations of additional macros are embedded into the macro %GEEWHIZ. This includes two error check macros, that check the macro parameters and the used datasets and variables for validity and a few other macros, e.g. a macro that takes care that indented text wraps appropriately. All of these macros are part of a macro library whose members are automatically compiled.

CONCLUSION

The macro %GEEWHIZ tries to combine user friendliness with a great deal of functionality and flexibility. Users report that it requires a bit of a learning curve to get used to it. However, once you have some experience it pays off rather quickly because you can generate a large number of tables in a more or less automated fashion.

Its basic use is pretty straightforward and with some experience you can implement more sophisticated solutions.

The experience has shown that the %GEEWHIZ macro can handle approximately 90% of the adverse events tables, concomitant medication and concomitant diseases tables. Furthermore, it was also used to create tables from a lot of other domains such as laboratory tables or vital sign tables.

The macro has evolved over time with a lot of features and solutions being implemented into an already existing macro. The new functions were built-in as options by appending values to existing macro parameters over the years and over the projects. New macro parameters were not introduced into the xml-spreadsheet. On the one hand this procedure allows the use of the same set of macro parameters for simple tables and more complex tables. On the other hand the correct parameter specification often requires a look into the user manual because it can be complex.

As a drawback due to the seemingly complexity of the macro, some users are hesitant to start using the macro and it needs some promotion to persuade these users.

REFERENCES

DelGobbo, V. 2006. "Creating AND Importing Multi-Sheet Excel Workbooks the Easy Way with SAS® ". Proceedings of the Thirty-First Annual SAS Users Group International Conference, 31. CD-ROM. Paper 115. Available http://www2.sas.com/proceedings/sugi31/115-31.pdf.

CONTACT INFORMATION

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

Guido WENDLAND
UCB Pharma SA
Alfred-Nobel Str 10
40789 Monheim
Germany
Work Phone: + 49-2173-48 1944
Fax: + 49-2173-48 1947
Email: guido.wendland@ucb-group.com
Web: www.ucb-group.com