

Program Documentation Process Automated with SAS® Macros to ease the pain...

Chris Hord, Quintiles, Inc., Kansas City, Missouri
Jianlin (Jay) Zhou, Quintiles, Inc., Kansas City, Missouri

ABSTRACT

It is no secret that regulatory authorities in the pharmaceutical industry require more documentation than most programmers would like to provide. To ease the pain, we've developed two SAS macros that help automate the documentation process. The first macro creates a standard header documentation shell at the top of each program, to collect the project code, author's name, date created, output type, QC reviewer's name, date of QC, etc. The second macro extracts the information from the header of each program in a specified directory and creates a SAS data set from which a Word Document is created. This table provides vital tracking information for each program written for a project, which fulfills a company reporting requirement, and provides an invaluable productivity tool for the programmer and the entire project team.

INTRODUCTION

Although programmers come from all walks of life and all corners of the globe, one trait seems almost universal: we don't like to document. Since this trait is diametrically opposed to the expectations of the pharmaceutical industry and regulatory agencies, we decided to do something to help our programmers meet documentation requirements.

When a company SOP was issued that required (for each protocol) a table providing descriptive data for each program used for reporting and creation of analysis files (see Appendix 1 for abridged version), we spotted a golden opportunity to automate the table population process with SAS macros. Our vision was to develop macro code to extract header documentation data from each program in a directory and use that data to populate the required table. It is not uncommon for pharmaceutical studies to be supported by 100 or more reporting programs requiring documentation in the aforementioned table. By automating the table data entry process, we could reap the compound benefits of eliminating the time programmers would spend keying data into the table, plus preempt the need for QC of all that manually entered data.

We recognized early on that the success of our endeavor rested initially upon development of a standard header documentation template that programmers would be willing to adopt to preface each program. Once established, we wrote two SAS macros designed to be run in tandem. The first macro (%QHeader) creates the standard header documentation shell for each program. The second macro (%QDocum) extracts the header documentation data from each program in a specified path, and creates the program documentation output file as specified by the user.

PREREQUISITE: DOCUMENTATION TEMPLATE

Since the data provided in each program's header documentation block would be extracted by macro code and used to populate the output file, it was essential to establish a standard header documentation template to be used uniformly in all programs for all projects. This was the most practical way to accommodate extraction of header data elements from many programs. Hence our first step prior to writing any macro code was to examine company SOP's to distill the standard set of documentation elements that were required for each program. We then produced the standard header documentation prototype, where each documentation element is captured on a separate record to

simplify the entry of data by programmers. We incorporated feedback from programmers and management to accommodate a broad spectrum of usage, and thereby transformed the header prototype into the standard header documentation template (see Appendix 2).

MACRO %QHEADER

The %QHeader SAS macro was written for the NT and Unix to automate the creation of the standard header documentation shell. The %QHeader macro dedicates one or two lines to each of the following header documentation elements: drug project name, program name (including the full path), author's name, date started, SAS version and operating system, validating programmer, date validated, description, design reference, calling routine, called routine, input data, output files, comments, and revision history.

%QHeader uses the SAS automatic macro variable SYSDATE to populate the "Date Started" field. Several of the documentation elements can be populated by specifying their values when the macro is called, as described below. Values not specified in the macro call and all value additions/revisions are simply keyed by the programmer within the file editor. As a final step of the %QHeader macro execution, the macro launches an edit session for the shell just created, where the programmer can immediately begin entering the header documentation data and initiate programming.

The call to the %QHeader macro takes the following form:

```
%qheader(help,
           dir=,
           progname=,
           project=,
           author=,
           desc1=,
           desc2=,
           desc3=,
           desc4=,
           desc5=)
```

The %QHeader macro has a mixed parameter list, with one positional parameter and nine keyword parameters, as described below:

help	prints to the SAS log window a description of how to use the macro. Activated by providing '?', 'h', or 'help' in the macro call. Example: %qheader(help).
dir	identifies the directory path where the SAS program (for which the program header is being created) will reside. Example: %qheader(dir= d:\dev\test).
progname	name of SAS program file being created. Example: %qheader(progname= mytable.sas). Alternatively, the user could omit the dir parameter and use the progname parameter to define the complete path, including the program name. Example: %qheader(progname= d:\dev\test\mytable.sas).
project	project for which the program is being written. Example: %qheader(project= Company/Project/Protocol).

author	author(s) of the program. Example: %qheader(author= Chris Hord and Jay Zhou).
desc1	brief description of purpose/objectives of program, first line. If one line is insufficient, additional lines are available via keyword parameters desc2-desc5.
desc2	brief description of purpose/objectives of program, second line.
desc3	brief description of purpose/objectives of program, third line.
desc4	brief description of purpose/objectives of program, fourth line.
desc5	brief description of purpose/objectives of program, fifth line.

It is important to note that the %QHeader macro is designed to be called only at the inception of writing a specific program, as a mechanism to create the program file and build the header documentation shell. From that point on, the documentation header is updated using the program editor. Hence, when a call to the %QHeader macro is issued, the path and filename specified in the call (using **dir** and **progrname** keywords) should describe a file that does not yet exist in the specified path. Otherwise, if the file already exists, a safeguard within the macro is triggered (to protect against overwriting a pre-existing file); macro execution will be aborted, and the following notes will appear in the SAS log:

Note:The specified file already exists in the directory specified.
Note:The macro call has been terminated.

MACRO %QDOCUM

The %QDocum macro is designed to run after all desired programs in the target directory have been prefaced with the standard documentation header (produced earlier by running the %QHeader macro). It is important to note that any programs in the target directory that do not have the standard documentation header will be ignored by the %QDocum macro, and will be omitted from any output files %QDocum produces. When %QDocum executes, it extracts the program documentation data from each program in the target directory and then populates the specific output file(s) as specified by the user in the macro call.

The call to the %QDocum macro takes the following form (defaults shown):

```
%qdocum(dir=,
        fname=,
        fpath=,
        history= n)
```

The %QDocum macro has a parameter list of four keyword parameters, as described below:

dir	identifies the directory path where the program (from which the header documentation will be extracted) resides. Example: %qdocum(dir= d:\dev\test).
fname	defines the name of the output file(s) that will be created and populated with data extracted from the program header (normally, the output file will be populated from the program headers of many programs). Example: %qdocum(fname= pgmsdoct).
fpath	identifies the path of the output file(s) that will be created. Example: %qdocum(fpath= l:\dev\doc).

history	indicates whether the Revision History output file will be created. Possible values are n (no - default); y (yes); null. Example: %qdocum(history=n).
----------------	---

The %QDocum macro is currently designed to furnish two summary outputs in several optional file formats (ie., SAS dataset, Word, Excel), but we have just begun to explore many other possibilities, in terms of additional outputs and file formats. The two current outputs are the SAS Program Documentation file and the Revision History file; see Appendices 3 and 4. These Word documents are produced using the Quintiles Output Generation System (OGS).

These summaries serve as valuable inventory tools to track the progress of programs through the life-cycle of a project. Their value is amplified for projects where many programs are required or where a team of programmers is involved. A current summary can be produced quickly and easily by simply running %QDocum.

These summaries offer tremendous advantages over the time-consuming approach where the summary data are manually entered and QC'd for each requested update. Attempts to provide frequent updates on-demand via the manual process are laborious and error-prone at best. The macros overcome these drags on productivity by eliminating the manual data entry and QC steps through automation.

EXAMPLE

See Appendix 5 for an example of calls to macros %QHeader and %QDocum.

CONCLUSIONS

In the procedures presented, SAS macros provide an automated method to compile documentation from selected programs in a target directory into a summary output file. Execution of the %QHeader macro creates a standard shell that makes it easy for the programmer to provide header documentation within each program. Execution of the %QDocum macro extracts the header documentation data from each program in the path specified and transforms this set of program-specific data into convenient, comprehensive summary output files. These summary outputs can be readily adapted to meet the user's needs, and current summaries can be produced quickly by simply running %QDocum and printing the summary outputs.

This methodology will reduce program documentation time and should allow more time for programmers to actually write code...thus achieving our ultimate goal of reducing programmers' pain!

REFERENCES

SAS Institute Inc., *SAS Macro Language: Reference, Version 6, First Edition*, Cary, NC: SAS Institute, Inc., 1997.

SAS Institute Inc., *SAS Language: Reference, Version 6, First Edition*, Cary, NC: SAS Institute, Inc., 1990.

ACKNOWLEDGMENTS

The authors would like to acknowledge Dawn DuBois for reviewing and editing this paper.

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of the SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

CONTACT INFORMATION

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

Chris Hord

Quintiles, Inc.
10245 Hickman Mills Drive
Kansas City, MO 64134
Work Phone: 816-767-6414
Fax: 816-767-7347
Email: chris.hord@quintiles.com

Jay Zhou

Quintiles, Inc.
10245 Hickman Mills Drive
Kansas City, MO 64134
Work Phone: 816-767-6768
Fax: 816-767-7347
Email: jay.zhou@quintiles.com
zhou_jay@yahoo.com

Tracking Document for Derived Datasets and Statistical Outputs

Customer: _____

Project Code: _____

TLG No.	TLG Title/ Description	TLG Filename	Programmer	Output Produced Date	QC Reviewer	QC Review Date

DD = Derived Dataset, TLG = Table, Listing or Graph, QCR = Quality Control (QC) Review

Initials/Signatures: _____

(Note: This Form is a representation of a spreadsheet. Columns may be added, relabeled or deleted depending on the project scope of work.

Appendix 2 Standard program header template

```

/*****
CONFIDENTIAL AND PROPRIETARY
Copyright © 2001 Quintiles Transnational Corp.
The contents of this program are confidential and proprietary to Quintiles
Transnational Corp. Use, disclosure or reproduction without the prior
written consent of Quintiles Transnational Corp. is strictly prohibited.
*****/

Project      ::
Program name ::
Authors name ::
Date started ::
SAS Version/OS::
Validated by  ::
Date validated::
Description   ::

Design reference:

Calling routine:

Called routine:

Input data:

Output files:

Revision history:
Date      Author      Description of the change

| *****/
```

Appendix 3

Program Location: c:\conference\prog

Protocol Title

Page 1 of 1

SAS Program Documentation File for Project XXX

Project Path: c:\conference\prog

Program Name	Programming		Validation		Date Last Updated	Description
	Programmer Name	Date	Programmer Name	Date		
lpkconc.sas	John Doe	27DEC2000	J.S.	21MAY2001	21MAY2001	Plasma concentration (ng/mL) and time
lpkpara.sas	John Doe	28NOV2000	J.S.	21MAY2001	21MAY2001	list of plasmacokinetic parameters
tpdemax.sas	John Doe	25NOV2000	J.S.	21MAY2001	21MAY2001	Test scores following oral administration of drugname
tpdpara.sas	John Doe	25NOV2000	J.S.	21MAY2001	21MAY2001	Test scores following oral administration of drugname
tpkconc.sas	John Doe	27DEC2000	J.S.	21MAY2001	21MAY2001	Plasma concentration (ng/mL) of each analyte
tpkpara.sas	John Doe	11OCT2000			21MAY2001	PK parameters following oral administration of drugname

Appendix 4

Program Location: c:\conference\prog

Protocol Title

Page 1 of 1

Revision History of SAS Programs for Project XXX

Project Path: c:\conference\prog

Program Name	Programming		Revision History
	Programmer Name	Date	
tpdemax.sas	John Doe	25NOV2000	20FEB01 M. E. Changed to ANOVA 27FEB01 M. E. Changed to ANOVA: Removed Main effects
tpdpara.sas	John Doe	25NOV2000	30NOV2001 J.S. Added one decimal to the range values per J.C.
tpkpara.sas	John Doe	11OCT2000	30NOV2001 J.S. Added one decimal to the range values per J.C.

Appendix 5 Example calls to %QHeader and %QDocum

```
*** Compile the macro programs ***;
%inc 'c:\sastools\maclib\qheader.sas';
%inc 'c:\sastools\maclib\qdocum.sas';

*** Create a program header for Program tdiscon.sas ***;
%qheader(progname=c:\conference\prog\tdiscont.sas,
         project=Client Name/Project Name/Protocol Number,
         author=John Doe,
         desc1=Table of Study Completion/Discontinuation Information,
         desc2=The table will be used as in-text table
        )
```

The following program header was generated with the above macro call:

```
/*****
CONFIDENTIAL AND PROPRIETARY
Copyright © 2001 Quintiles Transnational Corp.
The contents of this program are confidential and proprietary to Quintiles
Transnational Corp. Use, disclosure or reproduction without the prior
written consent of Quintiles Transnational Corp. is strictly prohibited.
*****/

Project      :: Client/Project/Protocol
Program name  :: c:\conference\prog\tdiscont.sas
Authors name  :: John Doe
Date started  :: 21MAY2001
SAS Version/OS:: SAS V6.12/WIN_NT
Validated by  ::
Date validated::
Description   :: Table 1 Study Completion/Discontinuation Information
               :: The table will be used as in-text table

Design reference:

Calling routine:

Called routine:

Input data:

Output files:

Comments:

Revision history:
Date      Author      Description of the change

/*****/
```

```
***** Create Word document files *****;
%qdocum(dir=c:\conference\prog,
        fname=demo,
        fpath=c:\conference\demo,
        history=y
       )
```