

Automating Validation of QC Plan and Creating SAS® Programs with Standardized Headers

Venkatesh Nemalipuri, Ateet Shah, Vertex Pharmaceuticals, Boston, MA

ABSTRACT

Clinical trial data computing environment can be part of audit by regulatory agencies. Validation of QC plan dates against the date stamps of Production SAS® programs is often a manual or a non-existent process and manual checking can take a considerable amount of time and resources, not to mention the risks of errors in process. The first macro can compare date stamps of Production SAS® programs against the QC Plan and if there are any mismatches, it can email the list of issues to their authors automatically. Also, it can alert programmers and study lead if program is present in the QC plan and not in the folders or vice versa and can write the Date QC passed in the QC Plan based on date stamps of the QC SAS® programs. Manually creating hundreds of standardized SAS® programs in short time period can be a tedious task. The second macro will generate SAS® programs with standard header for SDTM, ADaM, TFL's and Adhoc based on QC plan information and will populate the Purpose field for TFL's using titles in the QC plan. Both of these macros help to automatize these quality control and validation processes in a clinical trial data computing environment.

INTRODUCTION

Quality Control Plan (QC Plan) also referred as Quality Control Tracker (QC Tracker) is commonly used in the industry as an official document for tracking quality control process of biometrics deliverables. It typically entails details of production and validation like dates and names of programs and programmers, level of validation, statistician's review for each output generated along with QC related notes/comments. Due to multiple rounds of production and validation during the development and finalization, the completion of all the entries in QC plan is not only tedious, but also error prone.

Another important practice in the industry is to have standard program headers. These headers provide information about program name, project, study, analysis, and purpose of the program to name a few. During the development cycle any changes to the Titles for TFLs need to be reflected in the **Purpose** field of the program headers.

This paper describes two macros to help programmers automate these manual tasks. The first macro %M_QCPLAN_DATE checks and validates the dates Ready for QC and programs entered into QC plan, sends emails to appropriate programmer(s) when any mismatches are detected and populates dates QC Passed. The second macro generates blank SAS® programs with standard headers for all outputs and also pre-populates all the required fields based on the information in the QC plan. The macro also provides functionalities of updating the Purpose field in the program header, but keeping all other contents intact when a title in the QC plan gets changed and the corresponding SAS® program already exists as well creating new programs shells as and when information in the QC Plan is updated.

LOGISTICS OF MACRO - %M_QCPLAN_DATE

Typical QC plan (See Appendix 1) contains name of the programs and programmers and dates for both production and QC. Date columns in particular are required to be updated constantly during the development phase and once at the end of the study.

This macro:

1. Checks for any mismatches between Date Ready for QC and Programs between the QC Plan and what is available in the corresponding program folders. This macro should be run after all development cycles are completed.

2. If any mismatches are found, the macro will automatically email an attachment **Qcplan_Datecheck.rtf** (See Appendix 2) detailing appropriate discrepancies to the programmer(s) mentioned in the “**production programmer**” column of the QC Plan. The email is sent from the email id mentioned in the parameter **From** of the macro. Study lead can also be added to the emails using the parameter **SSP**.
3. Lastly the macro automatically populates “QC Passed Date” in the QC Plan based on the date stamp of proc compare QC outputs (.lst files). For this purpose the macro should be run after confirmation that all program logs are clean and all outputs have passed QC to further reduce burden on the programmers.

The output generated by the macro **Qcplan_datecheck.rtf** contains the project name, study name and reporting effort as header making it is easier to locate the program filenames with issues as programmers are working on multiple reporting efforts under different studies. For example, the header is **123 FINAL QC_FILE REPORT** for Final folder of study 123 from project ABC. **The Qcplan_datecheck.rtf** contains following columns:

1. Sheet Name
2. File Name
3. Programmer Name
4. Date Ready for QC
5. File Saved Date
6. Comment

These details help programmer to pinpoint the programs with issues and update them as appropriately. All the checks and corresponding comments are described below in Table 1.

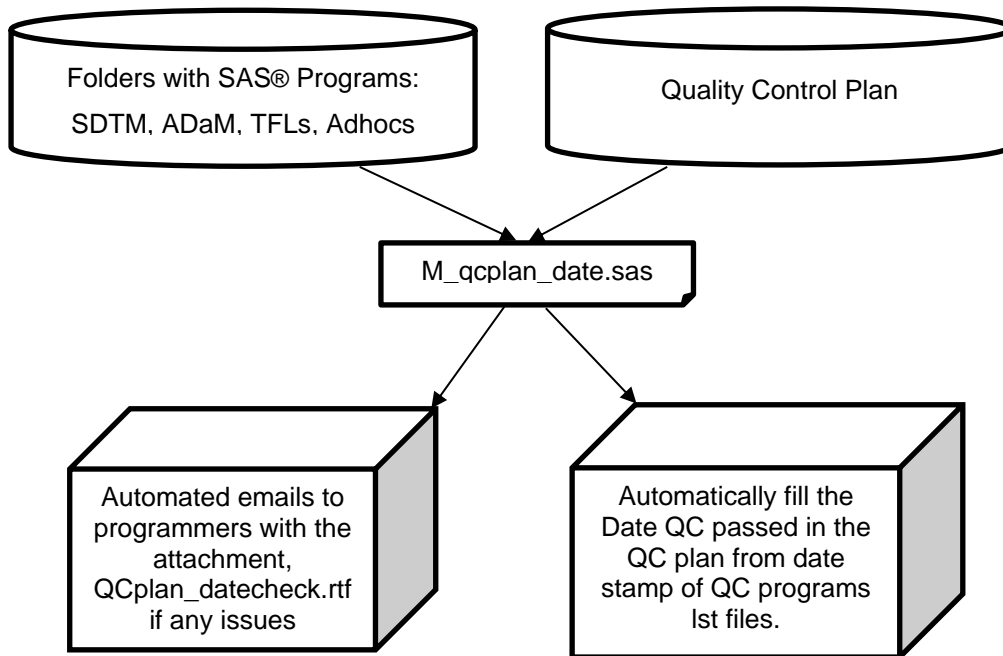
Table 1 Comments from Qcplan_datecheck.rtf

Checks	Comments
Program last update date not matching between QC Plan and program folders	DATE READY FOR QC AND FILE SAVED DATES ARE NOT EQUAL
Program not created in the folder yet	PROGRAM IS PRESENT IN QC PLAN, BUT NOT IN FOLDER
Program not added to QC Plan	PROGRAM IS PRESENT IN FOLDER, BUT NOT IN QC PLAN
Date not populated in QC Plan	PROGRAM IS PRESENT IN FOLDER, BUT DATE READY FOR QC IS MISSING IN QC PLAN.

Dependencies are described here:

1. QC plan in excel format containing columns for production/QC program names, production/QC programmer names and Date Ready for QC/Date QC passed.
2. List of programmer names and their emails should be available.

The process is depicted in the following basic figure:



Macro parameters are described in Table 2.

Table 2 Explanation of Macro Parameters

Parameter	Description	Required Variable	Default Value	Example
Qcplan	The name of the QC plan.	Yes		Study123 QC Plan
Email	Send email notification to programmers.	Yes		Y or N
From	User running the macro call	Yes		Venkatesh_Nemalipuri@vrtx.com
StudyLead	Study lead to be included in all emails	Yes		Ateet_Shah@vrtx.com
Infiles	Folder names which are compared against the QC plan.	Yes		%str(upcase(memname) in ('SDTM', 'ADAM', 'TABLES', 'LISTINGS', 'FIGURES'ADHOC TABLES'))
Qcpassed	To automatically populate QC date passed in the QC. Y is selected when programs are validated.	Yes	N	Y or N

Debug	To debug the macro. Default is N. If set to Y, work data sets created during macro execution should be kept for checking any issues encountered by the macro.	No	N	Y or N
-------	---	----	---	--------

An example of Marco call is below:

```
%m_qcplan_date
      (qcplan=E:\clinstat_dev\VXABC\123\Final\documents\qc\VXABC-123 Final QC
Plan,
      email= Y,
      from= Venkatesh_Nemalipuri@vrtx.com,
      ssp= Ateet_Shah@vrtx.com,
      infiles= %str(upcase(memname) in ('SDTM', 'ADAM', 'TABLES',
'LISTINGS', FIGURES', 'ADHOC TABLES')),
      qcpassed = Y,
      debug= N);
```

With these three different functionalities, this macro removes the manual process and significantly reduces time required for SSP to do these checks after every run of programs. At the time of database lock when SSP is juggling between multiple tasks, these automatic checks reduce the stress associated with quality control process. It also provides the assurance of compliance if audited by internal or external auditors.

LOGISTICS OF MACRO - %M_PGM_HEADER

This macro can be used at the start of programming activities to create template programs with standard headers or after the programming activities have started to incorporate any changes specially to TFL programs. It generates programs with standard header for SDTM, ADaM, TFLs and Adhocs based on information in the QC plan and populates the Purpose field for TFL's using titles from the QC plan. The macro also adds a few lines of codes like including directory setup program and clearing out log, output files and work library to give each program a clean start.

1. At the start of programming activities, the macro creates program shells with appropriate headers (see example below for program t-dm.sas). Macro can be run multiple times as new information is added to the QC Plan to create SDTM, ADaM, TFL programs at different times as and when this information is available.

An example program header created by the above macro call is provided below:

```
/******  
* COMPOUND:                VXABC  
*  
* STUDY/REPORT:            123 - Final  
*  
* PROGRAM:                 t-dm.sas
```

```

*
* PURPOSE:
*****/

* Initialization *;
*****;

dm log "clear";
dm output "clear";

proc datasets library = work kill nolist nodetails memtype=data;
quit;

```

2. This macro can be used once programming activities are underway and any updated information needs to be incorporated, especially for TFLs due to changes to titles. It reads the programs and only updates the header portion of the appropriate program.
3. In addition, this macro adds proc compare procedure to QC programs based on the sheet of the QC plan automatically. For example, SDTM QC programs have below proc compare code:

- a. For main domain

```

proc compare base= sdtm.<dataset> compare = qcdata.<dataset>
Listall;

run;

```

- b. For Supplemental domain

```

proc compare base=sdtm.<suppqual dataset> compare = qcdata.<suppqual
dataset> Listall;

run;

```

- c. TFLs programs have below proc compare code in the QC program files:

```

proc compare base = base compare = qc listall;

run;

```

In the course of TFLs development, title changes are quite common occurrences, and the purpose field of the program header should reflect the correct title. To update the purpose field for hundreds of programs manually can take considerable amount of time. This macro can update the purpose field of program headers systemically for all programs that need to be updated.

Dependencies are described here:

1. QC plan must have Production/QC program names filled.

Macro parameters explained in Table 3.

Table 3 Explanation of Macro Parameters

Parameter	Description	Required Variable	Default Value	Example
qcplan	The name of the QC plan.	Yes		
env	Location of the files to be saved.	Yes	dev qc	
debug	To debug the macro. Default is N. If set to Y, work data sets created during macro execution will be kept.	No	N	Y or N

An example of Marco call is explained here:

```
%m_pgm_header  
    (qcplan=E:\clinstat_dev\VXABC\123\Final\documents\qc\VXABC-123 Final QC  
    Plan,  
    env= dev qc,  
    debug= N);
```

CONCLUSION

As presented above the 2 macros reduce manual and often error prone work for critical programming activities. The first macro validates entries of Date Ready in QC plan and cross check program names in the study folders vs QC plan, helps improve compliance by automatically emailing appropriate programmers a word document with summary of issues in their programs and lastly even fill out the Date Passed QC column upon finalization of all programs. This in turn helps to keep QC plan documentation in compliance if needed by internal or external auditors. The second macro creates headers in template programs at the start of the programming activities and changes the Purpose field if any changes are needed once programming activities are underway. All of these tasks are completed in a fraction of time that would take to check and update manually, therefore, providing programmers more time to focus on other important tasks.

ACKNOWLEDGMENTS

SAS is a Registered Trademark of the SAS Institute, Inc. of Cary, North Carolina.

CONTACT INFORMATION

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

Venkatesh Nimalipuri
Vertex Pharmaceuticals
Venkatesh_Nimalipuri@vrtx.com
www.vrtx.com

Ateet Shah
 Vertex Pharmaceuticals
 Ateet_Shah@vrtx.com
 www.vrtx.com

APPENDIX 1 (QC PLAN)

SDTM Worksheet Example:

SDTM Dataset	Production Program	Production Programmer	Date Ready for QC	QC Program	QC Programmer	Date QC Passed
DM,SUPP DM	dm- suppdm.sas	Ateet Shah	02-Feb-2023	v-dm- suppdm.sas	Venkatesh Nemalipuri	20-Feb-2023
AE,SUPP AE	ae- suppae.sas	Ateet Shah		v-ae- suppae.sas	Venkatesh Nemalipuri	
CM,SUPP CM	cm- suppcm.sas	Ateet Shah	06-Feb-2023	v-cm- suppcm.sas	Venkatesh Nemalipuri	20-Feb-2023
DS,SUPP DS	ds- suppds.sas	Ateet Shah	07-Feb-2023	v-ds- suppds.sas	Venkatesh Nemalipuri	20-Feb-2023

Table Worksheet Example:

Table Number	Table Title	Production Program	Production Programmer	Date Ready for QC	QC Program	QC Programmer	Date QC Passed
14.1.1	Subject Disposition OL All Subjects Set	t-ds-disp.sas	Ateet Shah	22-Feb-2023	v-t-ds-disp.sas	Venkatesh Nemalipuri	28-Feb-2023
14.1.2	Summary of Important Protocol Deviations OL Full Analysis Set	t-deviation.sas	Ateet Shah	22-Feb-2023	v-t-deviation.sas	Venkatesh Nemalipuri	28-Feb-2023
14.1.3	Subject Demographics at Parent Study Baseline OL Full Analysis Set	t-dm-demo.sas	Ateet Shah	24-Feb-2023	v-t-dm-demo.sas	Venkatesh Nemalipuri	28-Feb-2023
14.1.4	Parent Study Baseline Characteristics OL Full Analysis Set	t-dm-blchar.sas	Ateet Shah	24-Feb-2023	v-t-dm-blchar.sas	Venkatesh Nemalipuri	28-Feb-2023

APPENDIX 2 (QCPLAN_DATECHECK)

VXABC-123 FINAL QC_FILE REPORT

SHEET	FILENAME	PROGRAMMER	DATE READY FOR QC	FILE SAVED DATE	COMMENT
TABLES	t-ds-disp.sas	Ateet Shah	22-Feb-2023		PROGRAM IS PRESENT IN QC PLAN, BUT NOT IN FOLDER
SDTM	dm-suppdm.sas	Ateet Shah	02-Feb-2023	03-Feb-2023	DATE READY FOR QC AND FILE SAVED DATES ARE NOT EQUAL
SDTM	ae-suppae.sas	Ateet Shah		06-Feb-2023	PROGRAM IS PRESENT IN FOLDER, BUT DATE READY FOR QC IS MISSING IN QC PLAN. IF PROGRAMMING IS NOT COMPLETED PLEASE IGNORE COMMENT
ADHOC_TABLES	AH-T-EX-EXPO.SAS	Ateet Shah		13-Feb-2023	PROGRAM IS PRESENT IN FOLDER, BUT NOT IN QC PLAN