

How to determine carriage controls in SAS® transfers?

Eric Kammer, Novartis Pharmaceuticals, East Hanover, NJ

Tiantian Sun, Novartis Pharmaceuticals, East Hanover, NJ

ABSTRACT

It is not uncommon for the pharmaceutical industry to outsource data management activities. However, when the data comes back to the sponsor company there are sometimes hidden surprises that cause problems in the SAS data sets that are transferred back from the CRO. In particular sometimes carriage controls are hidden in the SAS data sets from the transfers that occur or from the database applications used at the CRO. A technique to identify the carriage controls in data will be presented in this paper.

INTRODUCTION

In data cleaning activities, the focus for review in CRFs such as concomitant medications or adverse experiences, which have entered text, is only data that can be seen by the “naked eye.” It is not easy for data management to determine the carriage control problems that are then communicated by the statistical programming group, as these are invisible and can not be seen by the “naked eye.” This presentation will show an approach to detect hidden carriage controls using SAS, give clear direction of data cleaning to Data Management, and the end result will be less issues that arrive back for tabulations and the clinical study reports. Otherwise, the carriage controls cause printing problems when the tables and listings are produced. The application uses SAS 9.2 in Jreview® so the users can run the report as needed for any study within a Unix® operating system.

BASIC IDEA OF CARRIAGE CONTROL

Carriage controls can easily be overlooked when data managers review data as it is hidden. However, data review activity turns into a disaster when statistical programmers run programs which do not execute perfectly for the data with carriage controls.

SAS can not recognize them automatically but Word documents sure do.

Example of SAS Output where Line Feed Can Not Be Seen

```
Obs                                MEDNAME
1      ESMEEL(ACONITUM TINCTURE ARNICA TINC BELLADONNA TINC)
```

If this same line is copy and pasted into a Word Document the following would be seen

```
Obs                                MEDNAME
1      ESMEEL(ACONI TUM TI NCTURE
ARNI CA TINC
BELLADONNA TINC)
```

Also a carriage return can be specified by its hexadecimal value. For ASCII system, the value is '0D'.

Example of the SAS Output showing the contents in Hexadecimal:

0D = Carriage Control

```
45534D45454C2841434F4E4954554D2054494E4354555245200D0A41524E4943412054494E4320200D0A42454C4C4
```

How to determine carriage controls in SAS transfers, continued

SAS PROGRAMMING OVERVIEW

We apply the macro “_chk_special_characters” to read SAS data, export SAS data to txt file, find the location where carriage control occurs, and count patients having the issue. The programming logic of this macro could be divided into several steps as follows:

Step 1: Read SAS data set

At the first glance, there is no difference in the values of Medication/Non-drug therapy. The SAS data set contains the carriage control as normal data.

SAS data set: temp

	Study + SID	Medication / Non-drug therapy
6103	Study_Number_0281_00012	HARTMANN GLUCOSE
6104	Study_Number_0281_00012	SMOFLIPID(PURIFIED SOYBEAN OIL,FISH OIL,OLIVE OIL,TRIGLYCERIDE)
6105	Study_Number_0281_00012	PROBEN(MELILOTUS EXTRACT,PROXYPHYLLINE)
6106	Study_Number_0281_00012	ESMEEL(ACONITUM TINCTURE ARNICA TINC BELLADONNA TINC)
6107	Study_Number_0281_00012	FREAMIN 10%
6108	Study_Number_0281_00012	FREAMIN 10%
6109	Study_Number_0281_00012	H2(CIMETIDINE)
6110	Study_Number_0281_00012	H2(CIMETIDINE)

Step 2: Export to text file

After this SAS data set is exported to txt file, the truncation occurs, which capture the user's attention: Carriage control is hidden here! Note that the study number is not shown in the two lines below.

Temp_txt.txt

```
Study_Number_0281_00012 SMOFLIPID(PURIFIED SOYBEAN OIL,FISH OIL,OLIVE OIL,TRIGLYCERIDE)
Study_Number_0281_00012 PROBEN(MELILOTUS EXTRACT,PROXYPHYLLINE)
Study_Number_0281_00012 ESMEEL(ACONITUM TINCTURE
ARNICA TINC
BELLADONNA TINC)
```

Step 3: Import back to SAS

From line 6107, the value of patient record (STYSID1A) does not follow the way it should be as the carriage control has been located at the end of “ACONITUM TINCTURE” (cmdnam1a) and it jumps into next line immediately after importing back into SAS from the ASCII file.

SAS data set: temp2

	stysid1a	cmdnam1a
6106	Study_Number_0281_00012	ESMEEL(ACONITUM TINCTURE
6107	ARNICA TINC	
6108	BELLADONNA TINC]	

Step 4: Find the location of carriage return

Using the macro variable "pb" to explore the location of the records in which the carriage control is stored and then trace back to find which patient has this issue.

SAS data set: temp3

	pb	stysid1a	cmdnam1a
6106		Study_Number_0281_00012	ESMEEL(ACONITUM TINCTURE
6107	6107	ARNICA TINC	
6108	6107 6108	BELLADONNA TINC]	

How to determine carriage controls in SAS transfers, continued

Step 5: Sum up total patients having the record with carriage control

Macro variable "pat" is here to count total patients who have carriage control issue:

SAS data set: temp4

	pat	stysid1a	cmdnam1a
1	1	Study_Number_0281_00012	ESMEEL(ACONITUM TINCTURE
2	1	ARNICA TINC	
3	1	BELLADONNA TINC)	

Step 6: print out the listings containing records with carriage return issues.

Example of Source Code

```
%macro _chk_special_characters(libin= ,datain=, keepvar=, csvout=);

/*---&keepvar is the variable potentially containing carriage return---*/
/*-----&libin is library name-----*/
/*-----&datain is data set name-----*/

data temp;
    set &libin..&datain;
    keep stysid1a &keepvar;
run;

/*--- Export the data set as a TXT file in order to identify easily the new
line, carriage return... characters ---*/
proc export data= temp
    outfile= "/proj/rational/home/temp_txt.txt"
    dbms=tab replace;

run;
proc print data = temp;run;
/*--- Re-import the TXT file ---*/
data temp2;
    infile "&gendata./misc/temp_txt.txt" dsd dlm='09'x missover dsd firstobs=2
trunccover;
    input stysid1a $23. cmdnam1a $70.;
run;

/*--- Flag the lines in the data set not beginning by a patient number ---*/
data temp3;
    length pb $300.;
    format pb $300.;
    retain pb '';
    set temp2;
    if substr(stysid1a,1,12) ne "Study_Number" then pb=strip(pb)||"
"||compress(put(_n_,8.));
    call symput('pb',strip(pb));
run;

%if &pb. eq %then %do;
data _null_;
    file print;
    put "No such case of &keepvar in &datain panel";
run;
```

How to determine carriage controls in SAS transfers, continued

```
%end;

%else %do;

%let pb2=&pb;
%let nbpb1=%words(&pb.);

/*--- Put the lines number in a macro variable and get the line-1 in order to
get the patient number ---*/
%do i=&nbpb1. %to 2 %by -1;
    %if %eval(%scan(&pb2.,&i.)-1) ne %scan(&pb2.,%eval(&i.-1)) %then %do;
        %let pb2=&pb2. %eval(%scan(&pb2.,&i.)-1);
    %end;
%end;

%let pb3=%eval(%scan(&pb2.,1)-1) &pb2;
%let nbpb2=%words(&pb3.);

%put &pb.;
%put &nbpb1;
%put &pb2.;
%put &nbpb2;
%put &pb3.;
/*--- Keep only the problems with the number of patient ---*/
data temp4;
    retain pat 0;
    set temp2 end=last;
    %do i=1 %to &nbpb2.;
        if _n_ eq %scan(&pb3.,&i.) then do;
            if length(compress(stysidla)) ge 12 then do;
                if substr(compress(stysidla),1,12) eq "Study_Number"
then pat+1;
                    end;
                    output;
                end;
            %end;
            if last then call symputx('nbpat',pat);
run;

/*--- Blank lines for a better readability ---*/
data blank;
    length patid $27.;
    do pat=1 to &nbpat.;
        patid = "";
        output;
    end;
run;

data temp5(drop= pat);
    set blank temp4;
    by pat;
run;

proc export data= temp5
    outfile= "/proj/rational/home/temp.txt"
    dbms=tab replace;
run;
```

How to determine carriage controls in SAS transfers, continued

```
%mend _chk_special_characters;
```

CONCLUSION

The above codes not only help us to find carriage control locations, but also creates a TXT file including patient id, line number of the record having carriage control, which are convenient for data managers to quickly review from many records of clinical data.

In summary, using PROC EXPORT to send output to an ASCII (.TXT) file a customer can easily see carriage controls located in a file. If the sequence of patient, field name (e.g. medication) is not correct and there are spaces then this will be shown as an inconsistency. Next, using INFILE statement to send this TXT file back to SAS is a start to create a clear report which simply gives direction to records including carriage control. Finally, customers can then easily use the report to communicate with the CRO and get the carriage controls removed.

SAS can be used to show the inconsistencies of a patient id and then a program can be developed to identify these inconsistencies.

RECOMMENDED READING

- Base SAS® Procedures Guide
- SAS® For Dummies®
- SAS FAQ # 1617 for Words macro

CONTACT INFORMATION

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

Eric Kammer
180 Park Avenue
Florham Park, NJ, 07928
862-778-6172
E-mail: eric.kammer@novartis.com

Tiantian Sun
180 Park Avenue
Florham Park, NJ, 07928
862-778-0725
grace.sun@novartis.com

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

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