

Automation of paper dossier production for Independent Review Charter

Ting Ma, Pharmacyclics, Sunnyvale, CA
Jeff Cai, Pharmacyclics, Sunnyvale, CA
Michelle Zhang, Pharmacyclics, Sunnyvale, CA
Lina Gau, Pharmacyclics, Sunnyvale, CA

ABSTRACT

Independent review charter (IRC) is recommended by FDA to minimize bias in radiographic interpretation of radiological findings. For oncology clinical trials, paper dossiers documenting the clinical information of subjects need to be submitted to IRC for assessments on an on-going basis whenever the pre-defined milestones are reached. When information from the same subject is sent to IRC more than once, track changes may need to be displayed on the paper dossier to assist the data review. Tracking the historical submissions for every subject and generating the final delivery PDF files with highlighted changes can be quite time-consuming. This paper presents an automated process to produce the track-changed paper dossier in batch, using SAS 9.3, regardless of the different submission history for each subject. The automation should greatly improve the efficiency of this programming task for IRC.

Key words: Independent Review Charter (IRC), Dynamic Data Exchange (DDE), Visual Basic Application (VBA), X command

INTRODUCTION

The independent review of imaging and clinical data in oncology clinical trials is becoming increasingly important in supporting trial outcomes [1]. It is particularly so when response endpoints are based on clinical criteria such as physical and radiological examinations rather than validated laboratory measurements [2]. Materials that are commonly provided to IRC by programmers include SAS formatted datasets and subject specific paper dossiers. The contents and layout of an oncology paper dossier are usually well defined by a charter before the initialization of a study. Typically, such a paper dossier consists of findings from physical exams and lab values and is often sectioned into different panels chronologically by visits for easy readability. Moreover, when a subject needs to be assessed by IRC more than once, new data from the subject unavailable during the earlier review cycle and modifications made on the past data will need to be highlighted on the new document.

In brief, to support the IRC process, programmers are usually required to generate the following items for each subject:

1. Standard analysis datasets in SAS format;
2. A comprehensive paper dossier in PDF format, track-changed version required when applicable;
3. A restricted paper dossier in PDF format, track-changed version required when applicable.

Two techniques are commonly used in the SAS community when the outputs from different sources are compared. One of them is to compare the SAS datasets that were used to generate the outputs. This technique allows the differences of the two datasets to be printed clearly on a plain file. However, integrating these differences into the existing layout of the outputs in a reader-friendly fashion requires expertise beyond SAS. The second technique is to import outputs from different sources into a foreign application such as Microsoft Word and compare the outputs using embedded VBA codes. This technique generates nice looking outputs with differences between the two sources highlighted in a fashion familiar to most readers. However, with the existing tools available in Microsoft Word, individual files have to be compared one by one, manually, and the working environment has to be switched between SAS and Microsoft Office.

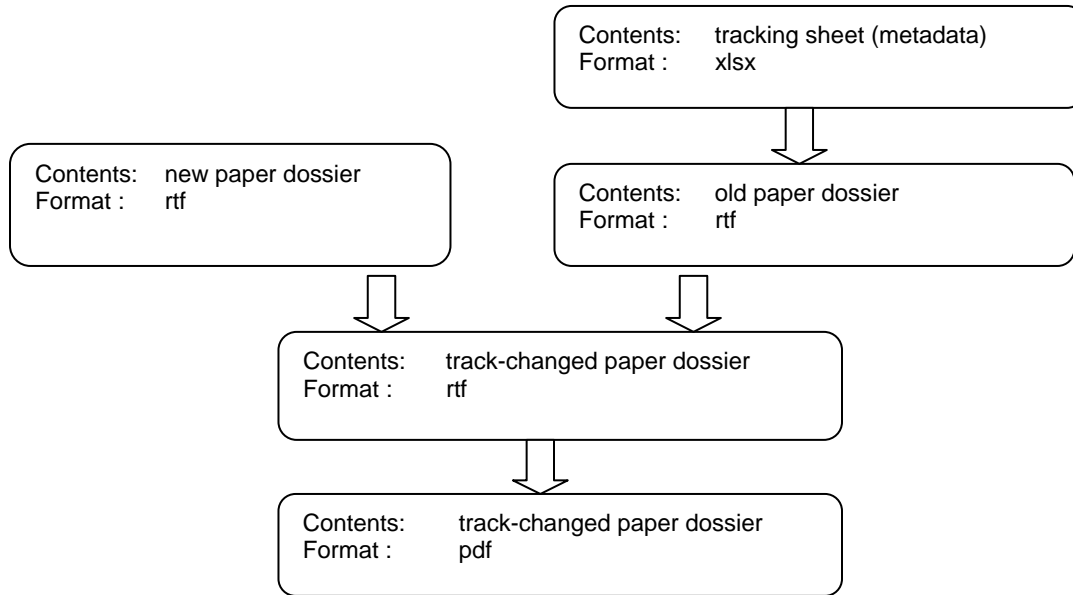
Given the facts that in a complex study, the number of subjects to be assessed by IRC in a single review cycle can be as many as a few hundred and the frequency of the need for IRC evaluation can be weekly or daily, an automated process is greatly needed to ensure a timely delivery of the above materials.

This paper addresses the challenges of producing paper dossiers with track changes for a large number of subjects whose historical submission dates and file storage locations are different. We assume that SAS programs are already developed to produce individual paper dossier without highlighted changes. Within this presented process, VBA codes are called in from SAS to generate the user friendly layout. This automated process enables the identification of each subject's latest submitted paper dossier, compares it against the newly generated paper

dossier, highlights the differences and batch-converts them to PDF format. The techniques adopted in this process include Dynamic Data Exchange (DDE), Visual Basic Application (VBA), X command.

OVERVIEW

The flow chart below demonstrates the automation process to generate the track-changed paper dossier.



ESTABLISHING A TRACKING SHEET

IRC reviews are called upon when patients in a study reach either a clinical milestone or a study milestone. Since patients reach these milestones at different time points, submission history to IRC is unique at patient level. In order to keep track of submission history, maintaining an accurate tracking system is essential. We recommend that submission history be maintained in a Microsoft Excel sheet mainly because of its easy accessibility within SAS. The minimum information contained in the tracking sheet should include subject ID, date of delivery to IRC, permanent storage location of paper dossier.

| | A | B | C |
|---|----------------------|----------------|---|
| 1 | Delivery Date | Subject | Tracking |
| 2 | 20131002 | 1000-000-001 | C:\IRC\outputs\20131002\1000-000-001_20131002.rtf |
| 3 | 20131127 | 1000-000-002 | C:\IRC\outputs\20131127\1000-000-002_20131127.rtf |
| 4 | 20131221 | 1000-000-003 | C:\IRC\outputs\20131221\1000-000-003_20131221.rtf |
| 5 | 20131221 | 1000-000-001 | C:\IRC\outputs\20131221\1000-000-001_20131221.rtf |

Display 1. Example of a tracking sheet (file name: trackingsheet.xlsx)

GETTING THE LATEST SUBMITTED PAPER DOSSIER

Subjects who were already reviewed by IRC during the earlier review cycle will need to have their current paper dossier compared with their previous version. The following codes identify such subjects and locate the latest submission using the tracking sheet.

The first step is to read the tracking sheet into SAS:

```
SYSTASK command "copy &__IRC.\IRC_tracking_sheet.xlsx
&__work.\IRC_tracking_sheet.xlsx" wait status=copyfl;

PROC IMPORT FILE = "&__work.\trackingsheet.xlsx"
      OUT = archive replace;
      SHEET = "&sheet" ;
      GETNAMES = yes ;

RUN ;
```

1

1: The SYSTASK statement copies the tracking sheet from the indicated storage location to the WORK directory of SAS. The wait option instructs SAS to suspend the current session until the task has completed. The status of the current task is stored into a macro variable copyfl. The specific sheet containing the tracking information is converted to SAS a dataset named "archive"

The second step is to locate the latest previous submission for each subject. The tracking information is sorted by submission date and only the latest submission for each subject is kept.

```
PROC SORT data=archive;
      BY subject delivery_date;
RUN;
```

```
/*Keep the latest previous delivery for each subject only*/
```

```
DATA lastd;
      LENGTH olddir $200;
      SET archive(where=(^missing(tracking)));
      BY subject delivery_date;
      IF last.subject;
      olddir=strip(tracking);
RUN;
```

2

2: The variable "olddir" contains the location of the .rtf version of the latest submitted paper dossier for each subject.

GETTING THE CURRENT PAPER DOSSIER

Assume the fresh paper dossiers for the requested patients have been produced by an established program in rtf format and saved in a directory called 'temoutloc'. The following code reads all of these rtf files in this directory.

```
/*Read all rtf filename under a directory */
X "dir temoutloc\*.rtf /b /l > temoutloc\ircreq.txt";
```

3

```
DATA ircreq;
      LENGTH filename $200;
      INFILE "temoutloc\ircreq.txt";
      INPUT filename;
      IF ^missing(filename);
      filename=strip(filename);
      subject=substr(filename,17,12);
RUN;
```

3: file names ending with .rtf in the temoutloc directory are saved on the ircreq.txt file and read into SAS dataset next.

FINDING THE LIST OF SUBJECTS THAT NEED DOCUMENT COMPARE

Subjects that are common on the current request list and on the archive will be written to the dataset 'comp'. This dataset contains both two directories that the compare will be performed and is the file that the VBA macro will use as an input.

```

PROC SORT data=ircreq;
BY subjid;
RUN;

DATA comp;
MERGE ircreq(in=a) lastd(in=b);
  BY subjid;
  LENGTH newf oldf trackf $400;
  newf="temoutloc\"||strip(filenm);
  IF a and b THEN DO;
    oldf=strip(olddir);
    trackf="temoutloc\"||scan(strip(filenm),1, '.')||'_track_chg.rtf';
    output comp;
  END;
KEEP subjid filenm newf oldf trackf;
RUN;

```

4

- 4: variable “newf” contains the path and full name of the current .rtf file.
 variable “oldf” contains the path and full name of the last submitted .rtf file.
 variable “trackf” contains the path and full name of the final .rtf file with the track-changes.

USING VBA TO COMPARE RTF DOCUMENTS

The comparison of RFT files is done using a Microsoft VBA macro that is called by SAS. The VBA script (CompareDocs.vbs) is outlined below. Refer to **Appendix A** for the source code of the macro m_DocComparison.

VBA script (CompareDocs.vbs) :

```

Dim old_file, new_file, out_file

```

5

```

new_file = WScript.Arguments(0) 'Full path and old file name
old_file = WScript.Arguments(1) 'Full path and new file name
out_file = WScript.Arguments(2) 'Full path and comparison file name
Sub DocComparison(oldFile,newFile,outFile)
  With CreateObject("Word.Application")
    .Visible=-1
    .Documents.Open oldFile
    .ActiveDocument.Compare newFile
    .ActiveDocument.SaveAs outFile
    .ActiveDocument.Close
    .Quit
  End With
End Sub

```

6

7

```

Call DocComparison(old_file, new_file, out_file);

```

8

- 5: Declares 3 variables which will be used as parameters for the sub procedure.
 6: The three parameters (old_file, new_file, out_file) entered need to have full path and file name.
 7: This sub instructs Microsoft operating system to complete the following actions: open Microsoft Word application, open the old_file, compare the old_file with the new_file, save the compared file as out_file, and close Microsoft Word application.
 8: Calls the above defined sub to compile.

Calling VBA script from SAS: The above VBA script is wrapped into a macro named “m_DocComparison” and called by SAS.

```

DATA _null_;
  SET comp;
  LENGTH command $2000;
  command='%m_DocComparison(newfile=%str('||strip(newf)||'),
    oldfile=%str('||strip(oldf)||'),outfile=%str('||strip(trackf)||'))';
  PUT command;

```

```
CALL execute (command);  
RUN;  
QUIT;
```

CONVERT RTF FILES TO PDF FILES

The rtf files are converted to PDF files in batch. The macro "m_word2pdf" is called within SAS. Refer to **Appendix B** for the source code of the macro word2pdf.

```
DATA _null_;  
  SET ircreq;  
  LENGTH command $2000 loc $1000;  
  loc="temoutloc";  
  command='%m_word2pdf(inDir=%str('||strip(loc)||'),  
  inFile=%str('||strip(filename)||'),  
  outFile=%str('||strip(subjid)||"&dt..rtf"||'))';  
  PUT command;  
  CALL execute(command);  
run;
```

CONCLUSION

Programming tasks to support IRC process generate the need to integrate SAS programs and functions performed outside of SAS. VBA scripts and DDE are useful tools in dealing with such assignments. Automation of the process flow is the key to produce consistently high quality outputs in a relatively short time-line.

REFERENCES

1. U.S. Department of Health and Human Services, Food and Drug Administration. Guidance for Industry Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics. 18-22. May2007. Available at <http://www.fda.gov/cder/guidance/index.htm>.
2. Kradjian, Steve. Development of a Charter for an Endpoint Assessment and Adjudication Committee. Therapeutic Innovation and Regulatory Science Jan2005; 39(1):53-61.

ACKNOWLEDGMENTS

The authors thank Bing Hu, Senior Statistical Analyst, for development of the VBA scripts and Statistical Programming Department at Pharmacyclics for the support throughout the efforts to develop and share this work.

CONTACT INFORMATION

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

Name: Ting Ma
Enterprise: Pharmacyclics
Address: 999 East Arques Ave, Sunnyvale, CA 94085
Work Phone: (408) 215-3577
E-mail: tma@pcyc.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.

Appendix A- macro to compare two rft files and generate a new file with tracking changes

```
%macro m_DocComparison(newfile=, oldfile=, outfile=);
option noxwait xsync;
%let a="&__tools.\CompareDocs.vbs' ""&newfile"" ""&oldfile"" ""&outfile"" ";
%put &a;
X &a;
%mend;
```

Appendix B- macro to transform a Microsoft WORD document to a PDF document

```
%macro m_word2pdf(inDir=, inFile=, outDir=, outFile=, delete=N);
options noxsync noxwait xmin;

%if %sysfunc(substr(&inDir, %length(&inDir),1)) ne \ %then %let inDir=&inDir\;
%if %index(&inFile,..) eq 0 %then %let inFile=&inFile..doc;

%if %length(&outDir) eq 0 %then %let outDir=&inDir;
%else %if %sysfunc(substr(&outDir, %length(&outDir),1)) ne \ %then %let outDir=&outDir\;
%if %length(&outFile) eq 0 %then %let outFile=&inFile;
%let filename=%sysfunc(scan(&outFile,1,'.));

%* start word;
filename sas2word dde 'winword|system';
data _null_;
length fid rc start stop time 8;
fid=fopen('sas2word','s');
if (fid le 0) then do;
rc=system('start winword');
start=datetime();
stop=start+20;
do while (fid le 0);
fid=fopen('sas2word','s');
time=datetime();
if (time ge stop) then fid=1;
end;
end;
rc=fclose(fid);
run;

%* write doc file to ps driver and create a temp ps file;
data _null_;
```

```
file sas2word;
put '[AppMinimize]';
put '[FileOpen.Name = " "&inDir&inFile" " ]';
put '[DocMinimize]';
put '[FilePrintSetUp .Printer="Adobe PDF",.DoNotSetAsSysDefault=1]';
put '[FilePrint'
    '.Background=1, '
    '.Range=0,'
    '.Pages="0",'
    '.PrintToFile=1,'
    '.PrToFileName="'&outDir&filename" '.ps",'
    '.NumCopies=1]';
put '[FileSave]';
*put '[FileSaveAs.Name = " "&inDir&inFile" " ]';
put '[FileClose]';
put '[FileExit]';
run;

options xsync xwait;
x "'E:\Program Files (x86)\Adobe Acrobat\Acrobat\AcroDist.exe' /n/q ""&outDir&filename..ps""";
options noxsync noxwait;
x "del/q ""&outDir&filename..ps"" ""&outDir&filename..log""";

%if &delete=Y %then %do;
  x "del/q ""&inDir&inFile""";
%end;
%mend m_word2pdf;
```