SAS® GLOBAL FORUM 2018

USERS PROGRAM

Producing a Format Library and Test Data for Case Report Forms using a Data Definition Table

April 8 – 11 | Denver, CO

#SASGF

Producing a Format Library and Test Data for Case Report Forms using a Data Definition Table

Suad El Burai Félix, MPH[1], Ellen J. DeMatt, MA[1], Rebecca A. Horney, BA[1]

[1]Cooperative Studies Program Coordination Center, Office of Research & Development,

U.S. Department of Veterans Affairs, Perry Point, MD

ABSTRACT

In Clinical Trials, a Data Definition Table (DDT) is a document that lists a variety of information about each case report form (CRF), such as form number, variable name, variable label or description, data type, length, and codes. When imported into SAS® it can be used to accomplish a variety of tasks. We aim to provide a workflow for creating a format library and test data for CRFs by using SAS and the DDT in Microsoft Excel. Both creating a format library and test data are necessary pieces to gather at the beginning of a study to eliminate the need of re-creating formats in each of your SAS programs (reduce code replication) and to start programming when real study data is not available yet (quicker delivery reporting time), respectively.

INTRODUCTION

Statistical programmers have the vital role of developing SAS coding in early stages of a study to ensure efficient and timely implementation of analyses. Consequently, the prompt development of SAS programs will allow being "ahead of the game" by becoming familiarized with the database (i.e., case report forms) content and producing reports, tables and listings in a timely manner. Two SAS programming practices that a statistical programmer can use to safeguard efficient and timely implementation of analyses are creating a centralized format library and test data.

We aim to describe a SAS macro that reads in an Excel format DDT file that comes from case report forms, which creates SAS data sets. In addition, we will walk you through the process of using SAS programs that utilize these SAS datasets from the macro to actually generate a format library and test data.

SETTING UP THE EXCEL DDT BEFORE RUNNING THE MACRO

In order to use the SAS macro that reads in the Microsoft Excel formatted DDT, the DDT must be one workbook containing a separate worksheet/tab for each case report form. The DDT in appendix A illustrates the columns needed in each DDT worksheet.

RUNNING THE MACRO TO READ IN THE DDT

The macro Read_DDT (appendix B) takes the data from the DDT and creates two SAS datasets per case report form (i.e., Fxx and Fxx_formats – where "xx" represents the form number). These datasets will be used to create a format library and test data. The following code calls the Read_DDT macro.

```
%let ddt = C:\SAS\DDT\;
libname ddt "C:\SAS\DDT";

%Read_DDT(frm_num=xx);
%Read_DDT(frm_num=xx); *Continue for each CRF form as needed;
```

CREATING A FORMAT LIBRARY

Run the program below, which uses the SAS datasets Fxx_formats created when running the macro to read in the DDT. This will generate a new SAS program called ProcFormat.sas located in the directory folder specified in the %let outdir statement.

```
%let outdir = C:\SAS\Formats\;

DATA allfmt_data;
   length label $300 form $8;
   set ddt.fxx_formats
       ddt.fxx_formats; *Continue for each CRF form as needed;
RUN;

PROC SORT data = allfmt_data; by form fmtname; RUN;
```

Producing a Format Library and Test Data for Case Report Forms using a Data Definition Table

Suad El Burai Félix, MPH[1], Ellen J. DeMatt, MA[1], Rebecca A. Horney, BA[1]

[1]Cooperative Studies Program Coordination Center, Office of Research & Development,

U.S. Department of Veterans Affairs, Perry Point, MD

CREATING A FORMAT LIBRARY (CONT.)

```
DATA allformats;
   set allfmt_data end=eof;
   by form fmtname;
   file "&outdir.ProcFormat.sas";
   retain val;
   if _n_ = 1 then do;
     put "libname outdir 'C:\SAS\Formats\';";
     put;
     put "PROC FORMAT library = outdir cntlout = outdir.Formats;";
   end;
   if first.fmtname and last.fmtname then do;
    len=length(fmtname); val=len+11;
    put/@3 "value " @9 fmtname @val value @val+5 "= '" @val+8 label +(-1) "';";
   end;
   else if first.fmtname then do;
     len=length(fmtname); val=len+11;
     put/ @3 "value " @9 fmtname @val value @val+5 "= '" @val+8 label +(-1) "'";
   end;
   else if last.fmtname then do;
     put @val value @val+5 "= '" @val+8 label +(-1) "';";
   end;
   else put @val value @val+5 "= '" @val+8 label +(-1) "'";
   if eof then put "RUN;";
RUN;
```

After the proc format program is generated, run it to create a centralized format library. A format will now be assigned to each codelist and checkbox variable.

CREATING TEST DATA

Once the format library is created, run the macro TestData_CreateEmpty (see appendix C) which will output a new SAS program for each case report form called EmptyFormxx.sas located in the directory folder specified in the %let testdata statement. Run the following code to call the TestData_CreateEmpty macro.

```
%let blank = 999; *Blank value for checkboxes;
%let testdata = C:\SAS\TestData;

%TestData_CreateEmpty(frm_num=xx);
%TestData_CreateEmpty(frm_num=xx); *Continue for each CRF form as needed;
```

Execute the EmptyFormxx.sas programs that were created when calling the TestData_CreateEmpty macro to create an empty dataset for each form. Then run the TestData_CreateCode macro (see appendix D). This macro will output another SAS program for each case report form called TestData-Formxx.sas located in the directory folder specified in the %let testdata statement. Run the following code to call the TestData_CreateCode macro.

```
%let line1="seed=2794383;";
%TestData_CreateCode(frm_num=xx);

%let line1="seed=2694383;";
%TestData_CreateCode(frm_num=xx); *Continue for each CRF form as needed;
```

The programs TestData-Formxx.sas created from the code above can be run to populate the empty datasets with test data.

Producing a Format Library and Test Data for Case Report Forms using a Data Definition Table

Suad El Burai Félix, MPH[1], Ellen J. DeMatt, MA[1], Rebecca A. Horney, BA[1]

[1]Cooperative Studies Program Coordination Center, Office of Research & Development,

U.S. Department of Veterans Affairs, Perry Point, MD

APPENDIX A

FORM	FIELD NAME	VALID RESPONSES	VALID RESPONSE CODES	DATA ELEMENT TYPE	LENGTH	FIELD LABEL	MIN	MAX	DEPENDENT ON	IF VALUE	FIELD #
1	CENTER			integer	3	Site #	1	3			1
1	PARTID			integer	4	Participant ID	1001	1200			2
1	ALPHA			Char	4	Alpha Code					3
1	VISITDATE			Date	10	Date Form Completed	7/1/2016	7/30/2018			4
1	DOB			Date	10	1. Date of Birth	7/1/1920	7/1/1995			5
1	HEIGHT			integer	2	2. Height (inches)	62	76			6
1	WEIGHT			integer	3	3. Weight (pounds)	100	230			7
1	GENDER	1	Male	Codelist	1	4. Gender					8
		2	Female								
1	RACE	1	American Indian or Alaskan Native	Codelist	1	5a. Race					9
		2	Asian or Pacific Islander								
		3	Black, not of Hispanic Origin								
		4	Hispanic								
		5	White, not of Hispanic Origin								
		6	Other, specify								
1	RACE_OTHERSP			Char	200	5b. Other Race, Specify			RACE	6	10
1	ARMY	1	Army	Checkbox		6a. Army					11
		999	Blank								
1	AIRFORCE	1	Airforce	Checkbox		6b. Airforce					12
		999	Blank								
1	NAVY	1	Navy	Checkbox		6c. Navy					13
		999	Blank								
1	MARINES	1	Marines	Checkbox		6d. Marines					14
		999	Blank								

APPENDIX B

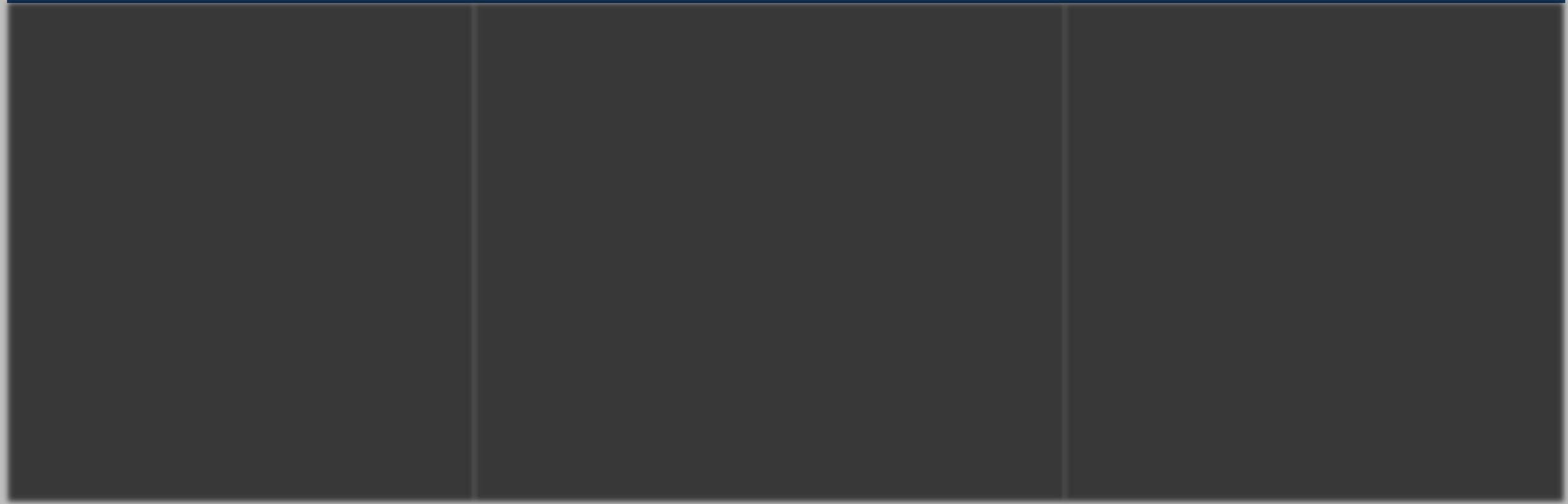
Producing a Format Library and Test Data for Case Report Forms using a Data Definition Table

Suad El Burai Félix, MPH[1], Ellen J. DeMatt, MA[1], Rebecca A. Horney, BA[1]

[1]Cooperative Studies Program Coordination Center, Office of Research & Development,

U.S. Department of Veterans Affairs, Perry Point, MD

APPENDIX C



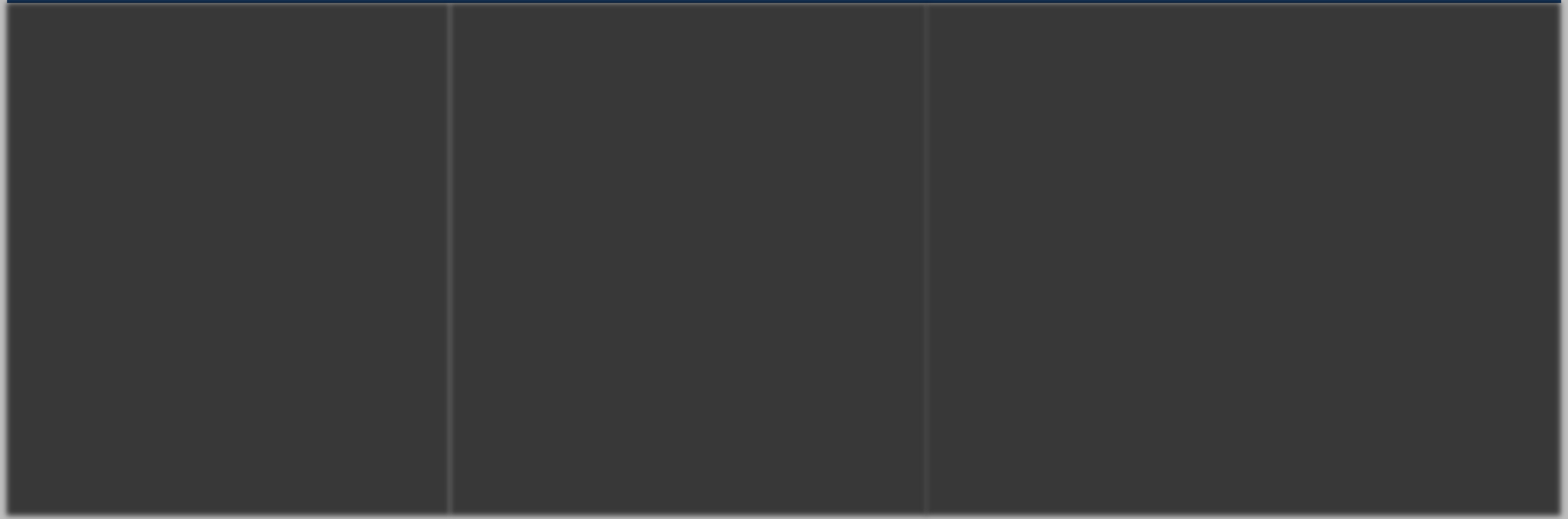
Producing a Format Library and Test Data for Case Report Forms using a Data Definition Table

Suad El Burai Félix, MPH[1], Ellen J. DeMatt, MA[1], Rebecca A. Horney, BA[1]

[1]Cooperative Studies Program Coordination Center, Office of Research & Development,

U.S. Department of Veterans Affairs, Perry Point, MD

APPENDIX D



SAS® GLOBAL FORUM 2018

April 8 – 11 | Denver, CO
Colorado Convention Center

#SASGF