PharmaSUG 2016 - Paper PO19

SAS Macro for Summarizing Adverse Events

Julius Kirui, Sarah Cannon Research Institute, Nashville, TN
Rakesh Mucha, Sarah Cannon Research Institute, Nashville, TN

ABSTRACT

Adverse Events (AE) summaries are very important in determining the continuity of most early phase studies. These summaries can be reported in various ways or formats as requested by study reviewers, FDA, or other regulatory agencies. The majority of the mock shells created by the biostatistician specify that an AE dataset is best summarized at patient level, System Organ Class (SOC), and Preferred Term (PT) by worst grade and a subset of relatedness or seriousness. Writing and validating SAS programs to mimic these mock shells, especially for multi-cohort or randomize studies can be challenging, time consuming, and tedious. This paper will discuss a SAS macro program (%sumAE) that can summarize raw or ADaM AE datasets by simple manipulation of macro parameters. The flexibility to output one summary table at a time or output as many summary tables as possible to one folder is accounted for. This macro can be used by individuals with minimum SAS 9.3 skills.

INTRODUCTION

Summarizing and reporting Adverse Events (AE) is a common and important type of safety output for clinical trial data. The most common summary reports include number of patients per treatment group with occurrence rate n(%). AE summaries are often grouped by System Organ Class (SOC) and Preferred Term (PT). This summaries can be further grouped by maximum CTCAE grade, relationship to study treatment, or seriousness criteria. The resulting combinations requires considerable programming effort which is repetitive and error-prone. SumAE macro provides a tool that facilitates production of multiple combinations of AE summary tables by applying different subsetting conditions. The macro reduces the need for repetitive manual programming, thus saving time and reducing possible sources of error.

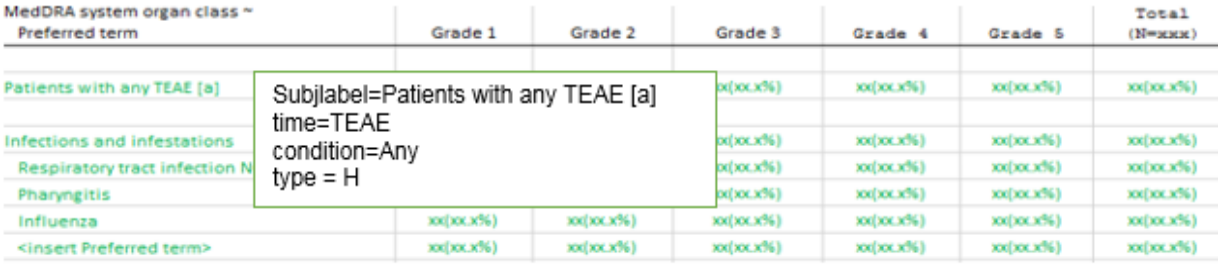
MACRO PARAMETERS

Parameters	Notes
subjlabel	Describes the number of subject who experienced an adverse event, e.g Subjects with any AE or Subjects with Any TEAE Resulting in Study Drug Discontinuation.
Table	Unique to each summary output.
Condition	Subsets safety dataset based on any AE, drug related, serious or related and serious AEs. Example, condition= RS will subset subject with drug related and serious AEs.
subset	Subsets base on one or more information, e.g subset = AEACN in('DISCONTINUED') or subset = AETOXGR in(3,4,5).
ordertotal	Parameter controls the order and display of maximum toxicity grades totals.
	ordertotal = -1 CTCAE grade total will be displayed before grade 1.
	ordertotal = 0 CTCAE grade totals will NOT be displayed in the final output.
	ordertotal = 1 CTCAE grade totals will be displayed after CTCAE grade 5.
Order	Order = 0 hide the system organ class category in the final output.
time	Controls the timing of the event ALL -all AE associated with the subject regardless of onset date TEAE -Treatment-emergent AEs (TEAE) are those with an onset on or after the initiation of therapy
grtype	Controls the display of CTCAE grades (*see shell 2*) **YES** display CTCAE grades based on macro parameter ordertotal **NO** hide CTCAE grades

Parameters	Notes
Sfdt	AE safety analysis dataset.
var2	Parameter controls stratification variable separated by '+' delimiter. The order of appearance should be maintained at all time. *Example* var2 = aaa + bbb + ccc
	Variable aaa represent study arms/cohorts, strictly in numeric (SAS character format) starting from four and incrementing by one i.e 4,5,...,n.
	Variable bbb is a text discerption of variable arm, it will be used for labeling the final summary outputs.
	Variable ccc represents arm grouping e.g Arm1, Arm2 dosed at 50mg and Arm1, Arm2 dosed at 100mg, variable ccc will host 50 mg and 100 mg.
var1	The parameter holds analysis variables. Variable names can change but the variable order and contents should be maintained at all times. This parameter will be broken into cgrade and var macro variable by scanning based on * dilimiter. *Example* var1=_ AETOXGR*Subjid+SOC+PTT
	Variable _ if the desired summary requires the **MOST** related or serious AE then _ will be replace by AE relatedness (AEREL) or seriousness (AESER) variable.
	Variable AETOXGR represents toxicity grade.
	Variable subjid represents unique subject identifier.
	Variable SOC represents system organ class.
	Variable PTT represents preferred term.
txdt	First day of study drug exposure.
aeodt	AE onset date.
type	Controls output display format. **Type = V** maximum toxicity grades will be displayed vertically. **Type = H** maximum toxicity grades will be displayed horizontally.
percent	Controls summary cut off percentage. Zero is the default cut off and the parameter is set to summary percentage greater than cut off.
pctyp	Controls the denominator used in calculating summary cut-off percentages, e.g pctyp=J will use the total number of patients in safety population as the denominator, pctyp=X will use the total number of patients per arm/cohort (a subset of safety population).
dataset	List of all dataset used e.g raw.ae raw.ex and e.t.c (applies to footnotes only).
cutoffdt	Data cutoff date (applies to footnotes only).
shell	Shell = NO allows the macro to utilize the next two parameters, cellwidth and cellwd to estimate the width of define statements in proc report. Shell=YES the program will subset the titles, footnotes, define and column information from an external metadata. Creating the metadata is beyond the scope of this paper.
path	Output destination folder
cellwidth	Controls the F1 column width style in proc report define statement i.e 'MedDRA system organ class Preferred term'
cellwd	Controls proc report column style Proc report data=v nowd style(column)=[cellwidth=&cellwd. in]

EXAMPLES OF AE SUMMARY MOCK SHELLS

TFL mock shells are designed by the statisticians as visual aids to the statistical programmers. The outputs produced by the statistical programmers should mirror the shells as much as possible. Mimicking this shells can be challenging, time consuming and tedious.

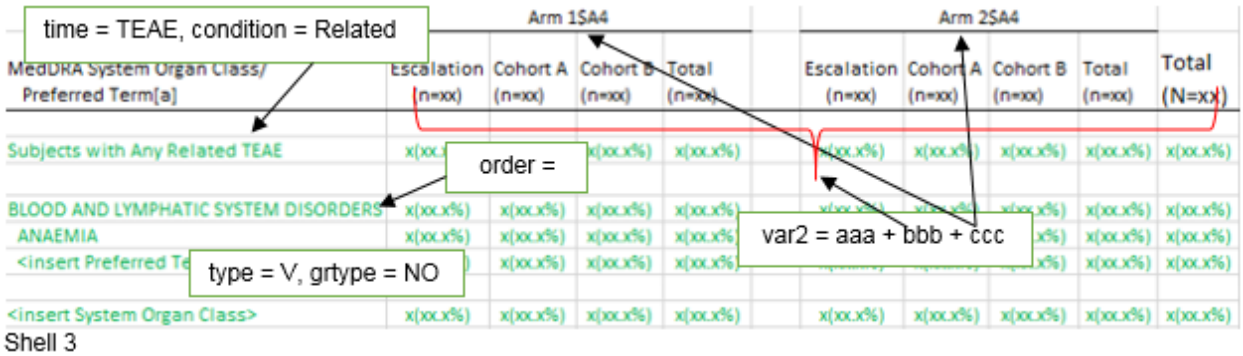


Shell 1

Shell 1 is commonly used in single arm studies and summarizes the maximum grade for each AE term for each subject. Due to page space limitations, for multi arm/cohort studies, statistical programmers may have no choice other than to subset and summarize each arm/cohort at a time, which is laborious and repetitive. This macro utilizes a *do loop* to extract CTCAE grade summaries per arm/cohort while constantly displaying category descriptors (*see RTF Extract 1*).

Shell 2

Shell 2 was extracted from a double blinded study. Most programmers anticipate double programing such studies i.e AE summaries prior and post blinding. Summaries prior to blinding require the 'Total' column only, while post blinding summaries require study drug arm and placebo columns, and 'Total' column may be optional. The chances of double programming is eliminated by this macro, by employing overall macro parameter. For example, display total column only overall = TOTAL, hide total column overall = NO and to display all study arms plus total column overall = YES.



Shell 3

Shell 3 represents an early phase study with escalation and expansion phase of cohort A and B grouped by arms. Notice that this shell requires arm totals as well as overall total making it much more complex. To accomondate for arm totals the user is required to skip one digit, and the presents of ccc in var2 macro parameter will prompt the program to complete the missing digit (*see data set up*).

SUMAE MACRO

This program requires minimal data manipulation prior to execution. The following parameter should be completed with the appropriate variables as explained above: var1, var2, aeodt, txdt and sfdt. Also, AE relatedness and seriousness variable should be named as **AEREL** and **AESER** respectively. The contents of this variables can be y/n or yes/no. We will vividly discribe this program by breaking it down into, data set up, cosmetics, counting, formating, and summary display.

Data set up

Prior to macro execution, study arms/cohorts must be set to consecutive digits starting from four to n^{th}. Arm/cohort order is majorly dictated by the shell, e.g in shell 2 study drug equal to four and placebo equals to five. The macro will introduce the total column as the maximum arm/cohort digit plus one, i.e total equals to six for shell 2. Extract below shows cohort set up for shell 3.

```
if scan(ARMCD,1,':')='ARM 1' and COHORT='DOSE ESCALATION PHASE' then  arm='4';
else if scan(ARMCD,1,':')='ARM 1' and COHORT ='Cohort A' then arm='5';
else if scan(ARMCD,1,':')='ARM 1' and COHORT ='Cohort B' then arm='6';
else if scan(ARMCD,1,':')='ARM 2' and COHORT ='DOSE ESCALATION PHASE' then arm='8';
else if scan(ARMCD,1,':')='ARM 2' and COHORT ='Cohort A' then arm='9';
else if scan(ARMCD,1,':')='ARM 2' and COHORT ='Cohort B' then arm='10';
```

Notice that arm='7' and arm='11' are missing. The presents of ccc in var2 macro parameter will prompt the program to introduce these digits as well as the overall total column arm, i.e arm='12'.
Safety dataset (&sfdt) should look and have the following variables

Prior execution data set up

Patient No	armgroup	arm	Cohort Assigned	Treatment Start Date	AE Onset date	Relationship to Study Drug	CTCAE Grade (1-5)	Serious Event?	Preferred Term	System Organ Class
1	1	4	DOSE ESCALATION PHASE	2013/05/16	2013/10/28	Y	2	N	DECREASED APPETITE	METABOLISM AND NUTRITION DISORDERS
13	2	9	Cohort A	2013/11/14	2013/11/16	Y	1	N	ABDOMINAL PAIN	GASTROINTESTINAL DISORDERS
29	1	5	Cohort A	2014/05/16	2014/05/16	N	1	N	CHILLS	GENERAL DISORDERS AND ADMINISTRATION SITE

Macro execution begins by counting unique subjects per arm and merging it back to the safety dataset. The counts will act as denominators for percentage calculations after frequency counts. The variable *ccc* in the macro parameter var2 will prompt the program to complete the skipped stratification variable by sub-setting, counting unique subjects, and setting it back to the original safety dataset. In absence of variable *ccc* and the number of arms is greater than one then the program will double the number of observations per arm to be counted and displayed as total column (see shell 2). If the number of arms equals to one then the program will pass to the next step. The program extract below illustrates how the total column is created for shell 1 and 2 and the data extract shows a completed arm for shell 3.

```
%let doarm=%eval(&doarm+1); /*&doarm = number of stratification arms*/
data sf;
       set sf;
%if &ntarm > 1 %then %do; /*&ntarm = total number of study arms*/
       output;
              %scan(&var2,1,'+')=strip(put("&doarm",10.));
              %scan(&var2,2,'+')='Total';
              count=input(&ztoth,best.); *total number of study subjects;
       output;
%let narm=%eval(&ntarm+4);%end; /*add 3+1 to account for 'Total' and first 3 columns
are reserved for other use*/
%else %if &ntarm = 1 %then %let narm=%eval(&ntarm+3);*single arm study not doubled;
run;%end;
```

Post execution arm set up

	subjid	armgroup	arm	cohort	count	ptt
1	1	1	4	DOSE ESCALATION PHASE	7	DECREASED APPETITE
2	1	1	7	Total	27	DECREASED APPETITE
3	1	3	12	Total	46	DECREASED APPETITE
4	2	2	9	Cohort A	5	ABDOMINAL PAIN
5	2	2	11	Total	19	ABDOMINAL PAIN
6	2	3	12	Total	46	ABDOMINAL PAIN

var2=arm + cohort + armgroup.
Arm 7 and 11 represents the total number of subjects per cohort, i.e escalation + cohort A + cohort B.
Arm 12 represents the total number of subjects in arm-group 1 and 2 (safety population).
Count represents the total number of subjects per arm-group per cohort.

Count column shown in the data extract above will hold and maintain the total number of subjects per arm/cohort despite and data subset.

Subset treatment Emergent Adverse Events (TEAE) or any AE regardless of occurrence time (ALL)

TEAE is defined as any adverse event (AE) that starts after the start of the first dose of study treatment. Partial dates can be handled prior to macro execution.

```
%if %scan(&time,&ti,'+')=ALL %then %do; ;%end;
%else %if %scan(&time,&ti,'+')=TEAE %then %do;where &aeodt ge &txdt ;%end;
```

Subset any, related, serious or related and serous AE

Notice that **AEREL** and **AESER** are not macro variables. These variables can be easily derived than assigning a macro parameter. Macro variables var and cgrade are derived from var1.

```
proc sort data=sf out=ds&i;by %scan(&var,1,'+') %scan(&var,&i,'+') %scan(&var2,1,'+')
&cgrade;
       %if %upcase(%scan(&condition.,&a,'+'))=ANY %then %do; ;%end;
       %else %if %upcase(%scan(&condition.,&a,'+'))=RELATED %then %do;where
       upcase(AEREL )in('Y','YES');%end;
       %else %if %upcase(%scan(&condition.,&a,'+'))=SERIOUS %then %do;where
       upcase(AESER) in('Y','YES');%end;
       %else %if %upcase(%scan(&condition.,&a,'+'))=RS %then %do;where upcase(AEREL)
       in('Y','YES') and upcase(AESER) in('Y','YES');%end; run;
```

Notice that a new dataset ds&i (i = 1,2,3) is created and the second sorting variable changes as &i changes. Variable F1 is created in each of ds&i datasets based on %scan(&var,&i,'+'). For example when i =1, F1 is assigned to the contents of subjlabel macro variable. Subjlabel can resolve to 'Subjects with any AE' or 'Subject with related TEAE' and e.t.c.

```
%if &i = 1 %then %do; *based on subject id information;
       if %scan(&var,3,'+') ^='' then do; F1="&subjlabel"; *if ptt is not missing;
                                          %scan(&var,2,'+')='1';end;
       else if %scan(&var,3,'+') ='' then do; F1=" "; *if ptt is missing;
                                             %scan(&var,2,'+')='';end;%end;
```

SAS log when i = 1

```
MLOGIC(SUMAE):  %IF condition &i = 1 is TRUE
SYMBOLGEN:  Macro variable VAR resolves to subjid+soc+ptt
MPRINT(SUMAE):  if ptt ^='' then do;
SYMBOLGEN:  Macro variable SUBJLABEL resolves to Patients with Any Treatment-Emergent
Adverse Event
MPRINT(SUMAE):  F1="Patients with Any Treatment-Emergent Adverse Event";
SYMBOLGEN:  Macro variable VAR resolves to subjid+soc+ptt
MPRINT(SUMAE):   soc='1';
MPRINT(SUMAE):  end;
```

Subset based other variables

This condition will execute when macro parameter subset is not empty. Safety dataset my not satisfy this condition thus resulting in an empty dataset. If the number of observation (&nobs) equals to zero then the program will terminate. Program extract below illustrats subset condition based on other variables.

```
%if %length(&subset) > 1 %then %do; if &subset;;%end;
%else %if %length(&subset) < 1  %then %do; ;%end;
```

Subset by arms and maximum CTCAE grade

Since we are applying a subset based on study arms/cohort, it is important to note that not all subjects in each arm/cohort will satisfy most of the previously applied conditions. In case of an empty dataset a dummy dataset will be created as, F1='XX', %scan(&var,2,'+')='XX'. The summary of dummy dataset will be deleted after combining all arm/cohort summaries.
%scan(&var,&i,'+') and &cgrade in the sort statement below will order CTCAE grade in ascending order thus allowing us to subset maximum CTCAE grade per subject per category, i.e subject level, SOC and PTT. if last.%scan(&var,&i,'+') condition is applied in the next data step.

```
proc sort data=ds&i out=ads&j;by %scan(&var,1,'+') %scan(&var,&i,'+') &cgrade;
                          where %scan(&var2,1,'+') ="&&ar&j";run;
```

arm macro variables (into :ar4 - :ar20) are pre assigned before any subsets is applied, thus allowing us to display all arms in the final summary.

Resulting data will be doubled by CTCAE grade and nine is assigned to the resulting grade. The new grade will be counted as the total events per grade.

Cosmetics

After all the major subsets, a dataset dmy1 is created per category based on SOC and PTT. Dmy1 dataset will contain variable grade with digits 1-5 and 9 per event in each category. The variable is intented to mimic CTCAE grade and will help in displaying all CTCAE grade in the final output (*see shell 2, totals are reprersented by 9*). Variable order, to be used in the final summary sort statement, is also created in this dataset, as shown.

```
if first.F1 and &i gt 2 then  order+1; *assign numeric values with increments of one
per unique PTT to order variable;
if &i = 2 then  order=0; *assigned zero to order variable per SOC;
else if &i=1 then order=-1; *assigned negative one to order variable per subject id;
```

Renaming, labelling and arm denominators

We will assign new names, labels, and summary denominators per arm. The requirement to represent study arms as digits plays comes to play in accomplishing this task. Each arm variable will be renamed to, F4, F5, ..., Fn, where n = last arm.

call symput("rename",%str("&&ar&j=**F**&&r&j")); *&&ar&j resolves to T_4 when j = 4, and &&r&j resolves to 4;
Macro variable **tarm&j** is the total number of subjects per arm *j*, acting as the desired denominator. Summary variable labelling will precede renaming.

```
select distinct strip(put(count,4.))||strip(put(count,4.))||')',orderarm,
'T_'||strip(%scan(&var2,1,'+'))||'="'||strip(%scan(&var2,2,'+'))||'~~^S={just=center}(
N='||strip(put(count,4.))||'^S={})"'
into :tarm&j, /*arm count*/       :nt&j,:dmy,:label
       from sf where orderarm=&j order by orderarm;
```

After merging all arm summaries with dmy1 dataset, variable ordertotal is created based on CTCAE grade and the contentents of ordertotal macro parameter i.e ordertotal = **9***&ordertotal. Ordertotal variable will determine the position of CTCAE grade total in the final output.
The preceding blank row per category in the mock shells above as shown below.

```
create table dummy as
       select %scan(&var,2,'+'),order-.5 as order,F1,max(cutoff) as cutoff
from vt&i where F1 ^='' group by F1
```

Maximum cutoff percentage per SOC will preserve the spaces, even when summary cutoff percentage is applied.

Counting and formating

Ads&j dataset will be counted under the *j loop*.

```
proc freq data=ads&j noprint;
       table soc*F1*&grade/ out=&&ar&j; *T_4 when j=4;
       where &grade ^='';  *&grade resolves to CTCAE grade variable;
run;
```

Resulting counts will be devided by arm denominator (&&tarm&j) to create a charater variable &&ar&j i.e T_4, T_5, …, T_&ntarm.

To satify type = H macro parameter an additional dataset _&j per arm will be created after merging dmy1 to &&ar&j dataset. _&j summary dataset will be transposed by %scan(&var,**2**,'+') F1 order, to display CTCAE grades horizontally. Each CTCAE grade summary column will be renamed as shown below (*see Footnotes*).

Summary extract 1.

F1	order	g4_1	g4_2	g4_3	g4_4	g4_5	g4_9	pctcut4	g5_1	g5_2	
Subjects with Any TEAE	-1	0	1(14.3%)	3(42.9%)	2(28.6%)	1(14.3	7(100.0%)	1	0	1(10.0%)	4(
Blood and lymphatic system disorders	0	0	2(28.6%)	1(14.3%)	2(28.6%)	0	5(71.4%)	0.7143	1(10.0%)	2(20.0%)	1(
\li400Anaemia	1	0	2(28.6%)	0	0	0	2(28.6%)	0.2857	0	1(10.0%)	0
\li400Neutropenia	2	0	0	1(14.3%)	2(28.6%)	0	3(42.9%)	0.4286	0	1(10.0%)	1(
\li400Thrombocytopenia	3	0	0	0	0	0	0	.	1(10.0%)	0	0
Cardiac disorders	0	0	1(14.3%)	0	0	0	1(14.3%)	0.1429	0	1(10.0%)	1(
\li400Cardiac arrest	4	0	0	0	0	0	0	.	0	0	0
\li400Coronary artery disease	5	0	0	0	0	0	0	.	0	0	1(

Number right after 'g' represents study arm derived from &j.
Number right after '_' represents CTCAE grade with the exception of _9 which represents grade totals.
pctcut4 is equivalent to the percentage portion of g&j_9.

We will combine individual arm summaries from each type (type = V or H) to create vt&i and ht&i respectively. At the end of *i loop*, append vt&i and ht&i per type to create vtab and htab summary datasets. Resulting dataset for each type are illustrated by summary extract 1 and 2.

Summary extract 2.

order	F1	F2	F4	F5	F6	F7	F8	F9	F10	F11	F12	cutoff
-1.5												1
-1	Subjects with Any TEAE	^S={just=left}Total	7(100.0%)	10(100.0%)	10(100.0%)	27(100.0%)	3(100.0%)	5(100.0%)	11(100.0%)	19(100.0%)	46(100.0%)	1
-1		^S={just=left}Grade 1	0	0	1(10.0%)	1(3.7%)	1(33.3%)	0	0	1(5.3%)	2(4.3%)	0.04348
-1		^S={just=left}Grade 2	1(14.3%)	1(10.0%)	0	2(7.4%)	0	0	2(18.2%)	2(10.5%)	4(8.7%)	0.08696
-1		^S={just=left}Grade 3	3(42.9%)	4(40.0%)	8(80.0%)	15(55.6%)	1(33.3%)	2(40.0%)	7(63.6%)	10(52.6%)	25(54.3%)	0.54348
-1		^S={just=left}Grade 4	2(28.6%)	4(40.0%)	0	6(22.2%)	1(33.3%)	3(60.0%)	2(18.2%)	6(31.6%)	12(26.1%)	0.26087
-1		^S={just=left}Grade 5	1(14.3%)	1(10.0%)	1(10.0%)	3(11.1%)	0	0	0	0	3(6.5%)	0.06522
-0.5												0.41304
0	Blood and lymphatic system disorders	^S={just=left}Total	5(71.4%)	5(50.0%)	2(20.0%)	12(					19(41.3%)	0.41304
0		^S={just=left}Grade 1	0	1(10.0%)	0	1(3					3(6.5%)	0.06522
0		^S={just=left}Grade 2	2(28.6%)	2(20.0%)	0	4(1					4(8.7%)	0.08696
0		^S={just=left}Grade 3	1(14.3%)	1(10.0%)	2(20.0%)	4(1					7(15.2%)	0.15217
0		^S={just=left}Grade 4	2(28.6%)	1(10.0%)	0	3(1					5(10.9%)	0.1087
0		^S={just=left}Grade 5	0	0	0	0	0	0	0	0	0	.
0.5												0.15217
1	\li300Anaemia	^S={just=left}Total	2(28.6%)	1(10.0%)	1(10.0%)	4(14.8%)	0	1(20.0%)	2(18.2%)	3(15.8%)	7(15.2%)	0.15217
1		^S={just=left}Grade 1	0	0	0	0	0	0	2(18.2%)	2(10.5%)	2(4.3%)	0.04348
1		^S={just=left}Grade 2	2(28.6%)	1(10.0%)	0	3(11.1%)	0	0	0	0	3(6.5%)	0.06522
1		^S={just=left}Grade 3	0	0	1(10.0%)	1(3.7%)	0	1(20.0%)	0	1(5.3%)	2(4.3%)	0.04348
1		^S={just=left}Grade 4	0	0	0	0	0	0	0	0	0	.
1		^S={just=left}Grade 5	0	0	0	0	0	0	0	0	0	.
1.5												0.23913
2	\li300Neutropenia	^S={just=left}Total	3(42.9%)	3(30.0%)	1(10.0%)	7(25.9%)	1(33.3%)	2(40.0%)	1(9.1%)	4(21.1%)	11(23.9%)	0.23913
2		^S={just=left}Grade 1	0	0	0	0	0	0	0	0	0	.
2		^S={just=left}Grade 2	0	1(10.0%)	0	1(3.7%)	0	0	0	0	1(2.2%)	0.02174

Max % per event to subset any % cut offs and retain a black row per category

Summary display

RTF is the primary output destination system for this macro, and it can be easily changed based on the user's desire. The changed can be made inside the macro or the user can set macro parameter path to blank to deactivate output path. Summaries can be harvested from vtab or htab summary datasets.

The uniqueness of htab and vtab summary dataset, forces us to utilize two proc report statements controlled by type macro parameter. Since RFT can decently display a certain number of columns, htab summaries will be displayed per arm. A *do while loop* starting from four will extract arm summaries while holding the descriptors, i.e columns F1-F3, at a constants. For example, if a study has two arms and shell 1 is required then arm 1 summaries will be written to RTF followed by arm2 and overall totals if needed.
The do loop above is applied to vtab summary dataset. A maximum of ten columns can be displayed per RTF page. *Do while loop* continues until the condition below resolves to false.

```
%if &j le 10 %then %let y=&j; *j is pre assigned to &narm;
%else %if &j gt 10 %then %let y=10; *maximum of 10 columns per page;
%do %while(&x <= &j and &y <=&j); *x will increment by y, y increments by y, j
increments by 3;
```

Overall macro parameters controls summary display for both vtab and htab summary dataset based on the program extract below.

```
%let q= ;
%if &overall=TOTAL %then %let x=&j;*total column only;
%if &overall=NO %then %do; %let j=%eval(&j-1);*no total column;
                    %let q=and id lt &narm;%end;
%if &overall=YES %then %let j=&j;*including total column;
```

Summary cut off percentage is applied to vtab when %upcase(&grtype)=NO, while htab applies to all formats, at arm lever or study level.

```
%if %upcase (&pctyp) =J %then %do;where pctcut&j gt %scan(&percent,&d,'+')/100; %end;
%else %if %upcase (&pctyp)=X %then %do;where pctcut&x gt %scan(&percent,&d,'+')/100;
%end;
```

At the bottom of each loop, a SAS dataset that can be harvested for proc compare and validation purposes is created. macro call to genarated rft extercts shown below.

```
%sumAE(
subjlabel=Subjects with Any TEAE,
condition=Any,                    /*Any = any AE, related, Serious, RS = related and serious*/
table=RTF.Extract1,               /*unique table ID*/
subset=,                          /*subset AE grades, PTT of interest, death due to AE and etc*/
type=H,                           /*H = CTCAE grades horizontally, V= CTCAE grade vertically*/
grtype=NO,                        /*grtype = YES display grade 1-5 and Total*/
ordertotal=-1,                    /*0= delete Totals 1=grade Totals at the bottom per PTT*/
order=,                           /*0=delete SOC*/
time=TEAE,                        /*ALL any AE regardless of time, TEAE after dose*/
sfdt=dm,                          /*safety datasets used*/
var2=arm+cohort+armgroup,         /*arm treatment coded, cohort= arm description in full, armgroup overall grouping*/
var1=_ AETOXGR*subjid+soc+ptt,    /*change variable name NOT ORDER*/
txdt=EXSTDAT,                     /*treatment start date*/
aeodt=AESTDAT,                    /*AE onset date*/
cutoffdt=,                        /*dataset cut off date, applies to footnote only*/
overall=YES,                      /*YES=all columns, NO=no Total column, TOTAL=total column only*/
percent=0,                        /*cutoff percentage, default=0 to show all*/
pctyp=J,                          /*J =use study total to calculate cutoff %, X=use group total to calculate %*/
cellwidth=2,                      /*width of column F1 in proc report*/
cellwd=.6,                        /*proc report column style */
dataset=raw.ie raw.ex raw.ae,     /*data used*/
shell=YES,                        /*Shell = YES, derive titles, footnotes, define and column info from mock shells */
path=<< output path>>)
```

RTF Extract -shell 1 output

MedDRA system organ class Preferred term	DOSE ESCALATION PHASE (N=7) Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Total
Subjects with Any TEAE	0	1(14.3%)	3(42.9%)	2(28.6%)	1(14.3%)	7(100.0%)
Blood and lymphatic system disorders	0	2(28.6%)	1(14.3%)	2(28.6%)	0	5(71.4%)
Anaemia	0	2(28.6%)	0	0	0	2(28.6%)
Neutropenia	0	0	1(14.3%)	2(28.6%)	0	3(42.9%)

MedDRA system organ class Preferred term	Cohort A (N=10) Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Total
Subjects with Any TEAE	0	1(10.0%)	4(40.0%)	4(40.0%)	1(10.0%)	10(100.0%)
Blood and lymphatic system disorders	1(10.0%)	2(20.0%)	1(10.0%)	1(10.0%)	0	5(50.0%)
Anaemia	0	1(10.0%)	0	0	0	1(10.0%)
Neutropenia	0	1(10.0%)	1(10.0%)	1(10.0%)	0	3(30.0%)
Thrombocytopenia	1(10.0%)	0	0	0	0	1(10.0%)

MedDRA system organ class Preferred term	Cohort B (N=10) Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Total
Subjects with Any TEAE	1(10.0%)	0	8(80.0%)	0	1(10.0%)	10(100.0%)
Blood and lymphatic system disorders	0	0	2(20.0%)	0	0	2(20.0%)
Anaemia	0	0	1(10.0%)	0	0	1(10.0%)
Neutropenia	0	0	1(10.0%)	0	0	1(10.0%)

MedDRA system organ class Preferred term	Total (N=27) Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Total
Subjects with Any TEAE	1(3.7%)	2(7.4%)	15(55.6%)	6(22.2%)	3(11.1%)	27(100.0%)
Blood and lymphatic system disorders	1(3.7%)	4(14.8%)	4(14.8%)	3(11.1%)	0	12(44.4%)
Anaemia	0	3(11.1%)	1(3.7%)	0	0	4(14.8%)
Neutropenia	0	1(3.7%)	3(11.1%)	3(11.1%)	0	7(25.9%)
Thrombocytopenia	1(3.7%)	0	0	0	0	1(3.7%)

RTF Extract -shell 2 output

MedDRA system organ class Preferred term	Maximum CTCAE grade	Statistic	xxx yyy zzz (N=74)	xxx yyy Placebo (N=76)	Total (N=150)
Patients with Any Treatment-Emergent Adverse Event	Total	n(%)	72(97.3)	75(98.7)	147(98.0)
	Grade 1	n(%)	3(4.1)	0	3(2.0)
	Grade 2	n(%)	9(12.2)	12(15.8)	21(14.0)
	Grade 3	n(%)	38(51.4)	36(47.4)	74(49.3)
	Grade 4	n(%)	20(27.0)	22(28.9)	42(28.0)
	Grade 5	n(%)	2(2.7)	5(6.6)	7(4.7)
Blood and lymphatic system disorders	Total	n(%)	53(71.6)	61(80.3)	114(76.0)
	Grade 1	n(%)	4(5.4)	5(6.6)	9(6.0)
	Grade 2	n(%)	14(18.9)	16(21.1)	30(20.0)
	Grade 3	n(%)	22(29.7)	26(34.2)	48(32.0)
	Grade 4	n(%)	13(17.6)	14(18.4)	27(18.0)
	Grade 5	n(%)	0	0	0
Anaemia	Total	n(%)	41(55.4)	47(61.8)	88(58.7)
	Grade 1	n(%)	5(6.8)	7(9.2)	12(8.0)
	Grade 2	n(%)	20(27.0)	21(27.6)	41(27.3)
	Grade 3	n(%)	16(21.6)	18(23.7)	34(22.7)
	Grade 4	n(%)	0	1(1.3)	1(0.7)
	Grade 5	n(%)	0	0	0
Bone marrow failure	Total	n(%)	0	1(1.3)	1(0.7)
	Grade 1	n(%)	0	0	0

RTF Extract -shell 3 output

MedDRA Preferred Term[a]	Arm 1 Escalation (n=7)	Cohort A (n=10)	Cohort B (n=10)	Total (n=27)	Arm 2 Escalation (n=3)	Cohort A (n=5)	Cohort B (n=11)	Total (n=19)	Total (N=46)
Subjects with Any TEAE	7(100.0%)	10(100.0%)	10(100.0%)	27(100.0%)	3(100.0%)	5(100.0%)	11(100.0%)	19(100.0%)	46(100.0%)
Blood and lymphatic system disorders	5(71.4%)	5(50.0%)	2(20.0%)	12(44.4%)	1(33.3%)	3(60.0%)	3(27.3%)	7(36.8%)	19(41.3%)
Anaemia	2(28.6%)	1(10.0%)	1(10.0%)	4(14.8%)	0	1(20.0%)	2(18.2%)	3(15.8%)	7(15.2%)
Neutropenia	3(42.9%)	3(30.0%)	1(10.0%)	7(25.9%)	1(33.3%)	2(40.0%)	1(9.1%)	4(21.1%)	11(23.9%)
Thrombocytopenia	0	1(10.0%)	0	1(3.7%)	1(33.3%)	0	1(9.1%)	2(10.5%)	3(6.5%)
Cardiac disorders	1(14.3%)	3(30.0%)	1(10.0%)	5(18.5%)	0	0	0	0	5(10.9%)
Cardiac arrest	0	1(10.0%)	0	1(3.7%)	0	0	0	0	1(2.2%)

CONCLUSION

After a complete set up, sumAE macro requires simple manipulations to generate as many AE summary tables as shown in appendix 1. Manual validation challenges can be eliminated by employing proc compare based on SAS summary dataset. With such simplicity, this macro is time saving and can be successfully used by individuals with minimum SAS skills.

ACKNOWLEDGEMENTS

We thank David Moorman, Bryan Lange and Zahir Karim for reviewing this paper and providing there comments.

CONTACT INFORMATION

Your comments and questions are valued.

Name: Julius Kirui
Enterprise: SCRI, Nashville, TN
Address: 3322 West End Avenue, Suite 900
City, State ZIP: Nashville, TN - 37221
Work Phone: (615) 524-4453
E-mail: Julius.kirui@scresearch.net

Name: Rakesh Mucha
Enterprise: SCRI, Nashville, TN
Address: 3322 West End Avenue, Suite 900
City, State ZIP: Nashville, TN - 37221
Work Phone: 615-329-7626
E-mail: rakesh.mucha@scri-innovations.com

APPENDIX 1.

Macro set up to generate different tables

table No.	Condition	Subjlable	Time	type	grtype	subset	order
1	ANY	subject with any AE	ALL	H			
2			ALL	V	NO		
3					YES		
4		subject with TEAE	TEAE	H			
5			TEAE	V	NO		
6			TEAE		YES		
7	RELATED	subject with any related	ALL	H			
8			ALL	V	NO		
9					YES		
10		subject with related TEAE	TEAE	H			
11			TEAE	V	NO		
12					YES		
13	SERIOUS	subject with serious AE	ALL	H			
14			ALL	V	NO		
15					YES		
16		subject with serious TEAE	TEAE	H			
17			TEAE	V	NO		
18					YES		
19	RS	subject with related and serious AE	ALL	H			
20			ALL	V	NO		
21					YES		
22		subject with related and serious TEAE	TEAE	H			
23			TEAE	V	NO		
24					YES		
25		subjects with AETOXGR => 3				AETOXGR in(3,4,5)	
26		AEs leading to dose reduction				AEACN in('REDUCED')	
27		AEs leading to treatment interruption				AEACN in('INTERUPTED')	
28		AEs leading to treatment discontinuation				AEACN in('DISCONTINUED')	
29		AEs resulting in subject death				AEOUT in('FATAL')	
30		AEs of special interest				AETERM in('ANEMIA',…)	0