

A SAS® Macro to Automate the Process of Pooling Sites

Changhong Shi, Merck & Co., Inc., Rahway, NJ

Lili Chen, Merck & Co., Inc., Rahway, NJ

ABSTRACT

In an international clinical trial, the patients are located in different geographic regions, such as US region, Asia-Pacific region, and other regions. In each region, the patients can be randomized to different sites, within different treatment groups.

To minimize the difference in the number of patients between each site within the same region, we pool several small sites together to make a larger site called CENTER according to pre-specified criteria. A center, rather than individual small sites, can be used to test the stability of treatment effect of a drug using SAS Mixed model, that is commonly used in statistical analyses.

This paper describes two approaches of applying sites/center(s) to fit the Mixed model. One is using an automatic macro call that pools small sites to create centers, another approach is considering the individual sites as the random effect without pooling.

This paper is focused on the way how a macro pools small sites to create centers, by using SAS system version 8.02 or higher in windows operating system. The paper is suitable for any level of audiences, especially the medium level users who are experienced in clinical trials.

How to pool sites into a Center

By pre-specified criteria, the number of patients n (for example n=3) in each treatment group has been defined. If a site contains all treatment groups and at least n patients in each treatment group, then this site has satisfied the criteria of a Center. Otherwise it is considered as a small site.

1. Calling the Traditional Macros: %poolcntr

By calling the POOLCNTR macro, it generates a SAS data set that contains a list of small sites. All small sites on this list can be pooled into one large center.

For example, to make a Satisfied Center, there is $n \geq 3$ in each treatment group.

Site#	Trt1 (n)	Trt2 (n)	Trt3 (n)	Total (n)					
0067	4	3	3	10	- satisfied center	List of small sites			
0068	4	5	5	14	- satisfied center				
0069	3	3	6	12	- satisfied center				
....	...	...	...	...					
0071	1	2	3	6	small site =>	=> pool into one center			
0072	3	1	3	7	small site =>				
0073	1	2	1	4	small site =>				
....	...	...	...	...					

Site#	Trt1	Trt2	Trt3	Total
0071	1	2	3	6
0072	3	1	3	7
0073	1	2	1	4
Total:	5	5	7	17

The assumption for this macro is that we only have a limited number of small sites, or a limited number of patients in the small sites. Therefore, after all small sites are pooled into one center, the number of patients in the center would not be too large, and there will be only comparable sized centers after pooling.

To release these limitations, we implemented a new macro %newpoolcntr.

2. Calling a New Macro: %newpoolcntr

By calling a macro %newpoolcntr, we can pool small sites into centers which each contain a comparable number of patients. It will distribute patients to each center to avoid the extreme large sized centers that can happen easily from the previous macro poolcntr. The following example shows how we use the new macro to pool the small sites into centers.

- 1) Get a data set that contains the patents located in the same region, such as US region, Asia-Pacific region, or Other region.

e.g. Where region = 'Asia-Pacific';

- 2) In a region, calculate the number of patients (n) in each treatment group for each site. In this example, there are 3 treatment groups per site in Asia-Pacific region.

Table 1: Sites in Asia-Pacific region

Site #	trt1(n)	trt2(n)	trt3(n)	
0067	4	3	3	satisfied
0068	4	5	5	satisfied
0069	2	2	3	
0095	2	2	2	
0096	3	4	3	satisfied
0097	3	2	3	
0099	1	2	2	
0103	3	2	4	
0104	2	2	2	
0105	3	3	2	
0106	2	2	3	
0107	2	2	2	
0108	2	2	1	
0109	3	2	3	
0121	2	2	2	
0125	1	2	2	
0126	2	2	2	
0127	2	2	2	

To make a satisfied center, the pre-specified criteria is $n \geq 3$ in each treatment group. From above table 1, only 3 sites are considered as 3 satisfied centers, the rest of small sites are considered as unsatisfied. Output 3 satisfied centers into a **Satisfied** data set, and output the rest of the small sites into an **Unsatisfied** data set.

```
data satisfied unsatisfied;
  set Asia_Pacific_region;
  if trt1 >= 3 and trt2 >= 3 and trt3 >= 3 then do;
    center=trim(left("&regionnum" || studynum));
    /* assign an unique center# to the satisfied site */
    output satisfied;
  end;
  else output unsatisfied;
run;
```

- 3) Call %newpoolcntr macro to pool the Unsatisfied small sites to make the satisfied center(s). In the example, it is sorted by ascending site # in Unsatisfied data set.

Table 2: Unsatisfied data set (n<3 in at least one treatment group):

Site #	trt1(n)	trt2(n)	trt3(n)
0069	2	2	3
0095	2	2	2
0097	3	2	3
0099	1	2	2
0103	3	2	4
0104	2	2	2
0105	3	3	2
0106	2	2	3
0107	2	2	2
0108	2	2	1
0109	3	2	3
0121	2	2	2
0125	1	2	2
0126	2	2	2
0127	2	2	2

To make a satisfied center, the macro adds up n (within each treatment group) from the last site into the next site, until it meets the specified criteria: at least 3 patients exist in each treatment group.

Table 3: Illustration of Making Satisfied Centers

Site #	trt1(n)	trt2(n)	trt3(n)	center	tot_trt1(n)	tot_trt2(n)	tot_trt3(n)
0069	2	2	3	1111	2	2	3
0095	2	2	2	1111	4	4	5
0097	3	2	3	1112	3	2	3
0099	1	2	2	1112	4	4	5
0103	3	2	4	1113	3	2	4
0104	2	2	2	1113	5	4	6
0105	3	3	2	1114	3	3	2
0106	2	2	3	1114	5	5	5
0107	2	2	2	1115	2	2	2
0108	2	2	1	1115	4	4	3
0109	3	2	3	1116	3	2	3
0121	2	2	2	1116	5	4	5
0125	1	2	2	1117	1	2	2
0126	2	2	2	1117	3	4	4
0127	2	2	2	1117	5	6	6

From the example, number of patients (n) in site 0069 is added to site 0095 per treatment group: 2+2=4 in trt1, 2+2=4 in trt2, 3+2=5 in trt3, it satisfies $n \geq 3$ per treatment group after combined these 2 small sites into a center 1111. Number of patients (n) in site 0097 is added to site 0099 within each treatment group: 3+1=4 in trt1, 2+2=4 in trt2, 3+2=5 in trt3, it makes another satisfied center 1112 after combined these two small sites. An unknown/unspecified number of small sites can be combined as a satisfied center until the pre-specified criteria is met (e.g. total $n \geq 3$ in each treatment group).

From Unsatisfied data set, if there is/are small site(s) (e.g. $n < 3$ in at least one treatment group) left in this data set after it has reached the last observation, the macro adds the left over small site(s) up to the last center. From above table 3, the last site 0127 is left in the Unsatisfied data set. As we can see the last satisfied center is 1117, which is created by combining site 0125 and site 0126. The left small site 0127 will be added to the last center 1117. This way, we found that there is no small site(s) left over and all small sites are pooled into the satisfied center(s).

The following sample programming shows how the macro %newpoolcntr works:

```
data unsatisfied;
  set unsatisfied end=eof;  /*when it reaches the last observation*/
  retain cn 1;              /*keep track of the increased cn number(center#)*/

  /* adds n per trt to tot_trt to keep track of n number*/
  tot_trt1+trt1;
  tot_trt2+trt2;
  tot_trt3+trt3;

  /* assign a center # to sites */
  center=trim(left("&regionnum")) || trim(left(put(cn, 4.)));

  if eof then do;          /* when it reaches the last observation */
    if tot_trt1 < 3 or tot_trt2 < 3 or tot_trt3 < 3 then
      call symput('last_cn', put(cn-1, 2.));
      /* if it is the small site left, then decreases the cn number to combine the left site
      to the last center */
    else call symput('last_cn', put(cn, 2.));
  end;  /* assign the last cn as a macro variable for later using */

  else do;
    /* if it reaches the satisfied center, it sets tot_trt to 0, increment cn to cn+1, to keep track of the
    center# and start to count total n in each trt in the next center */
    if tot_trt1 >= 3 and tot_trt2 >= 3 and tot_trt3 >= 3 then do;
      tot_trt1=0;
      tot_trt2=0;
      tot_trt3=0;
      cn=cn+1;
    end;
  end;
  drop tot_trt1 tot_trt2 tot_trt3;
run;

%let last_cn=&last_cn;

data unsatisfied;
  set unsatisfied;
  if cn > &last_cn then do;  /* combine the left small site(s) to the last center */
    cn = &last_cn;
    center=trim(left("&regionnum")) || trim(left(put(cn, 4.)));
  end;

run;
```

Sample Analysis

To provide the evidence on the stability of treatment effect of a drug, we are using SAS Mixed model. This part describes the approach of doing the analysis using center in the model, where center are created by pooling sites by using macro %newpoolcntr.

```

proc mixed data=effglm;
  by week;
  class trt priormed protocol center;
  model basechn = trt priormed basevalu protocol center trt*center;
  lsmeans trt / pdiff cl;
  ods output LSMeans=lsmn Diffs=lsmdiff CovParms=rmse Tests3=fixedeff;
run;
quit;

```

The MODEL statement specifies the continuous response variable, change from baseline (basechn), which is placed to the left of the equal sign. The explanatory variables are listed to the right of the equal sign. The last term, **trt*center**, provides a way to test for the interaction between the two effects: **treatment and center**.

The result shows the p-value for ANCOVA effects. By calling the macro %newpoolcntr that we specified in the previous part, it pools at least 3 patients in each of 3 treatment groups as a center. Thus the p-value shows the treatment effect across a center is approximately 0.728 (>0.05), which denotes no significant difference between treatment effects across centers.

week	sort	Descrip	effect	Pvalue
24	1	Baseline Value	basevalu	<0.001
24	2	Treatment	trt	<0.001
24	2	Treatment	trt*center	0.728
24	3	Prior Anti-hyperglycemic Medication	priormed	0.176
24	4	Center	center	<0.001

When the macro %newpoolcntr pools the small sites to a center as $n \geq 6$ or $n \geq 10$ in each of 3 treatment groups, the results show the p-values are 0.358 (>0.05) and 0.403 (>0.05) respectively.

AS WE CAN SEE, THE P-VALUES SHOW THAT THERE IS NO SIGNIFICANT CHANGE FROM THE TREATMENT EFFECTS ACROSS CENTERS, EVEN WHEN DIFFERENT NUMBER OF PATIENTS ARE POOLED INTO A CENTER (E.G. $N \geq 3$, $N \geq 6$, $N \geq 10$, IN EACH TREATMENT GROUP).

CONCLUSION

To test the stability of treatment effect of a drug within a certain number of patients, the macro %newpoolcntr is a useful utility that can facilitate the pooling of small sites data to one center before fitting to MIXED model. Therefore, no huge differences between sites can cause any potential misfit of the model.

REFERENCES

Macrolib libraries in Merck

ACKNOWLEDGMENTS

We would like to thank our management team for their encouragement and review of this paper. Special thanks to Elaine B. Czarnecki and Kaifeng Lu for getting us started and continuous support.

CONTACT INFORMATION

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

Changhong Shi
Merck & Co., Inc.
RY34-A320

126 East Lincoln Ave.
Rahway, NJ 07065
Work Phone: (732) 594-1383
Fax: (732) 594-6075
Email: changhong_shi@merck.com

Lili Chen
Merck & Co., Inc.
RY34-A320
126 East Lincoln Ave.
Rahway, NJ 07065
Work Phone: (732) 594-0613
Fax: (732) 594-6075
Email: lili_chen@merck.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.