

## Multiplicity Controlled Analyses Using SAS/IML®

Xingxing Wu, Eli Lilly and Company, Indianapolis, IN  
Hangtao Xu, Eli Lilly and Company, Indianapolis, IN

### ABSTRACT

Multiplicity controlled analyses are very important for the clinical trials with more than one objectives. Graphical approach proposed recently is becoming popular because of many advantages. Related tools implemented in R programming language have been developed to apply this approach to clinical trials. Even though the code developed in SAS/IML® is available to apply this approach in some straightforward cases, there is still lack of a comprehensive tool to fully apply this approach into multiplicity controlled analyses in clinical trials. In this paper, the program implemented in SAS/IML is presented because of the popularity of SAS. This program can not only be applied in the case when the p-value of each endpoint is available, it can also estimate the endpoint p-values and predict the success rate (power) of each endpoint through simulation to help the early decision-making. Furthermore, the implementation of family-wise gatekeeping approach in SAS/IML is also integrated into the program for the purpose of comparison of different approaches in multiplicity controlled analyses. This paper first introduces the graphical approach briefly. Then the implementation of graphical approach in SAS/IML is introduced for the case when the p-value of each endpoint is available. In addition, the approach to estimate the p-value and predict the success rate of each endpoint is proposed. Finally, this paper presents the implementation of family-wise gatekeeping approach in SAS/IML.

### INTRODUCTION

If a clinical trial has other objectives in addition to the primary objective, a strategy should be decided beforehand in order to meet the requirements of regulatory agencies on the strong control of family-wise error rate (FWER) at a pre-specified significance level  $\alpha$ . The purpose of the strategy is to reduce the probability of erroneously rejecting the true null hypothesis while still providing great flexibility in addressing clinical trial objectives (Alosh 2014). In the last decade, many approaches have been introduced in the literature to handle the multiplicity controlled analyses in clinical trials. Bretz (2009) and Burman (2009) proposed graphical approach around 2009. This approach is becoming more and more popular because the complex multiplicity controlled analyses can be constructed in a visualized approach. In addition, this approach can iteratively recycle unspent significant levels (Alosh 2014). Over the years, some tools have been created using R graph-based Multiple Comparison Procedures [gMCP] package and R Shiny in order to apply graphical approach into the multiplicity controlled analyses in clinical trials. Even though the code developed in SAS/IML is available to apply this approach in certain cases, there is still a need to develop a comprehensive tool in order to fully apply this approach to the multiplicity controlled analyses in clinical trials. The program presented in this paper is implemented in SAS/IML because of the popularity of SAS in different application areas. This paper will first introduce the graphical approach briefly. Then the implementation of graphical approach in SAS/IML will be introduced. The program can identify the significance of each endpoint based on the provided endpoint p-values. This information is critical to decide the successfulness of the clinical trials. However, these p-values will usually become available only at the very end stage of clinical trials. In order to help the early decision-making, the paper also presents the way to estimate the endpoint p-values based on the information from other clinical trials or literature. Based on these estimated p-values, the success rate (power) of each endpoint can then be calculated through simulation. This information will be helpful when deciding whether to continue or stop the clinical trials. Even though the graphical approach has many advantages, other methods also have their own advantages and are useful in certain multiplicity controlled analyses. This paper also presents the implementation of family-wise gatekeeping approaches in SAS/IML in order to compare the analysis results using different approaches in multiplicity controlled analysis. The programs provided in this paper are very useful for the multiplicity controlled analyses and will provide valuable information to help the decision-making.

## GRAPHICAL APPROACH

In graphical approach, a set of vertices with associated local significance level  $\alpha_i$  are used to represent the endpoint null hypotheses. In addition, the weight of the directed edge between any two vertices  $i$  and  $j$  is indicated by  $g_{ij} \geq 0$  (Bretz 2009). Basically, the p-value of each endpoint will be compared with the significance level  $\alpha$  assigned to this endpoint. If there is no endpoint hypothesis to reject or there is only one endpoint left in the graph, then stop; otherwise, one of the rejected endpoint hypothesis will be removed from the graph. The  $\alpha$  values of the endpoints which are still in the graph will be re-calculated. Similarly, the weight  $g_{ij}$  between any two vertices  $i$  and  $j$  still in the graph will be re-assigned accordingly. These steps continue until there is no endpoint hypothesis to reject, or there is only one endpoint in the graph. The algorithm that ensures a sequentially rejective Bonferroni-based multiple testing procedure can be expressed as follows, where  $I = 1, \dots, k$ ,  $H_i (i = 1, \dots, k)$  is the null hypothesis, and  $|A|$  denotes the cardinality of a set  $A$  (Bretz 2009):

- (1) Select an  $i$  such that  $p_i \leq \alpha_i$ . If no such  $i$  exists, stop; otherwise, reject  $H_i$ . If  $|I|=1$ , stop.
- (2) Update the graph:
  - a.  $I \rightarrow I \setminus i$
  - b.  $\alpha_j \rightarrow \begin{cases} \alpha_j + \alpha_i g_{ij}, & \text{if } j \in I, \\ 0, & \text{otherwise} \end{cases}$

If  $|I|=1$ , go to step 1.

  - c.  $g_{jl} \rightarrow \begin{cases} \frac{g_{jl} + g_{ji}g_{il}}{1 - g_{ji}g_{il}}, & \text{if } j, l \in I, j \neq l, g_{ji}g_{ij} < 1 \\ 0, & \text{otherwise} \end{cases}$
- (3) If  $|I| > 1$ , go to step 1; otherwise, stop.

We will now use a clinical trial example to illustrate the algorithm above. Figure 1 is the initial graph:

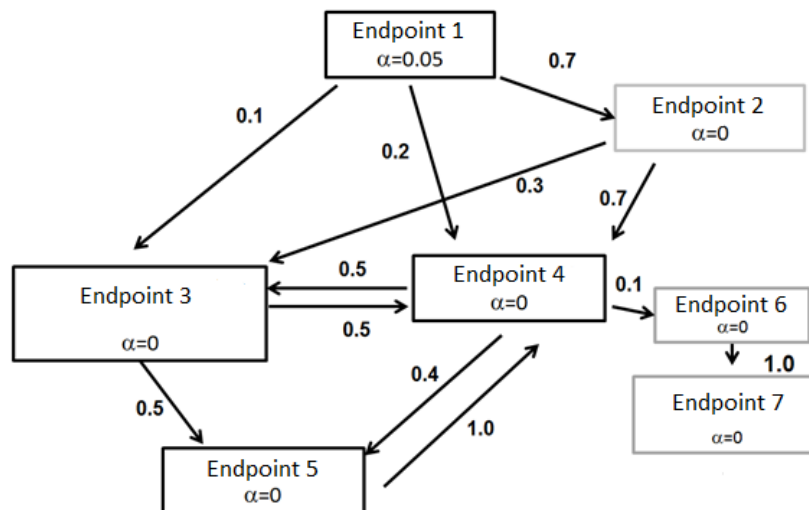


Figure 1. Graphical Approach Example: Initial Graph

Assume the null hypothesis of Endpoint 1 is rejected after comparing the p-value of Endpoint 1 with the local significance level  $\alpha = 0.05$ , the graph will be updated based on the algorithm mentioned above. Endpoint 2 will have  $\alpha = 0 + 0.05 \times 0.7 = 0.035$ , Endpoint 3 will have  $\alpha = 0 + 0.05 \times 0.1 = 0.005$ , while Endpoint 4 will have  $\alpha = 0 + 0.05 \times 0.2 = 0.01$ . In addition, the weights  $g_{12}$ ,  $g_{13}$ , and  $g_{14}$  will become 0, while other weights have no change based on this algorithm. The updated graph is shown in Figure 2.

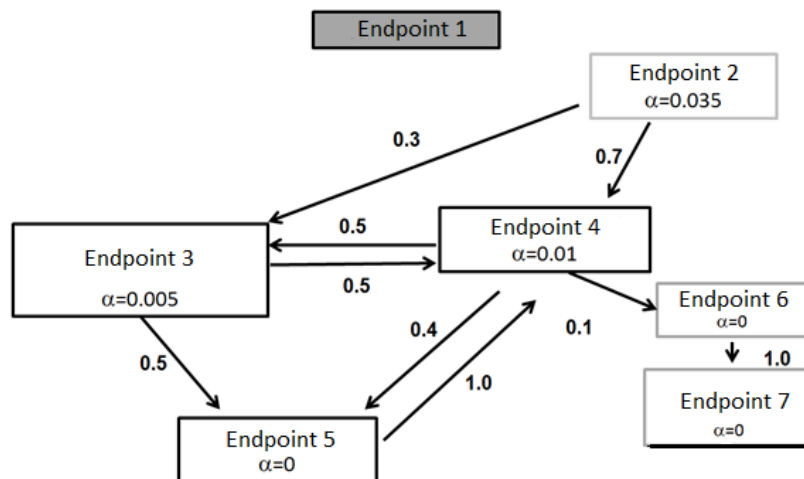


Figure 2. Graphical Approach Example: Updated Graph

The step above will continue until there is no null hypothesis to reject, or there is only one endpoint left in the graph. The graphical algorithm described above can be implemented using SAS/IML as follows (Bretz 2009):

```

start graphapp(a, w, p);
  n = ncol(a);
  rslt=j(1, n, 0);
  cond = 0;
  do until(cond = 1);
    hyporej = (p < a);
    if (any(hyporej)) then do;
      rejidx = min(loc(hyporej#(1:n)));
      rslt[rejidx] = 1;
      w1 = j(n, n, 0);
      do i = 1 to n;
        a[i] = a[i] + a[rejidx]*w[rejidx,i];
        if (w[i,rejidx]*w[rejidx,i]<1) then do j = 1 to n;
          w1[i,j] = (w[i,j] + w[i,rejidx]*w[rejidx,j])/
            (1 - w[i,rejidx]*w[rejidx,i]);
        end;
        w1[i,i] = 0;
      end;
      w = w1; w[rejidx,] = 0; w[,rejidx] = 0;
      a[rejidx] = 0;
    end;
    else cond = 1;
  end;
end;

```

```

end;
return(rslt);

finish graphapp;

```

In the code above, the parameter  $a$  is a vector to hold the local significance  $\alpha$  values,  $w$  is a matrix to hold the values of weight  $g_{ij}$ , and  $p$  is a vector to hold the p-values of endpoints. The return variable  $rslt$  is a vector to hold the rejection status of each null hypothesis. The value of 1 indicates that the null hypothesis of related endpoint is rejected.

## ESTIMATION OF P-VALUES AND PREDICTION OF SUCCESS RATE

The graphical approach mentioned above requires the availability of the p-value of each endpoint. Usually, these exact p-values can be calculated using SAS procedures, such as PROC MIXED, at the very end stage of clinical trials. In order to provide useful information to help early decision-making, this paper also presents a way to estimate the p-values of endpoints based on the information from other clinical trials or literature. Therefore, the success rate of each endpoint can then be predicted using simulation at the early stage of clinical trials. This will provide supportive information to decide whether to continue or stop the clinical trials.

In clinical trials, the effect size values or ranges can usually be estimated from the information of other clinical trials or literature. For example, for some clinical trials, the effect size of continuous endpoints can

be given by  $\theta = \frac{\mu_1 - \mu_2}{\sigma}$  where  $\mu_1$  and  $\mu_2$  are the true mean responses of the treatment groups and  $\sigma$

is the pooled standard deviation, while the effect size of binary endpoints can be given by

$\theta = \frac{p_1 - p_2}{\sqrt{p(1-p)}}$ , where  $p_1$  and  $p_2$  are the true response rates of the treatment groups and  $p$  is the

average response rate:  $p = (p_1 + p_2)/2$ .

The following code in SAS/IML is used to illustrate how to generate the p-value of each endpoint for a clinical trial, assuming the clinical trial has 8 endpoints, and the estimated effect sizes of endpoints are stored in vector *theta*.

```

* Number of tests;
ntests=&ntests;

* Effect sizes for Endpoints 1 through ntests;
theta=j(ntests,1,1);
theta={&theta1, &theta2, &theta3, &theta4, &theta5, &theta6, &theta7,
&theta8};

* Correlation matrix;
rho=j(ntests,ntests,1);

rho={ 1    &r12  &r13  &r14  &r15  &r16  &r17  &r18,
      &r12  1    &r23  &r24  &r25  &r26  &r27  &r28,
      &r13  &r23  1    &r34  &r35  &r36  &r37  &r38,
      &r14  &r24  &r34  1    &r45  &r46  &r47  &r48,
      &r15  &r25  &r35  &r45  1    &r56  &r57  &r58,
      &r16  &r26  &r36  &r46  &r56  1    &r67  &r68,
      &r17  &r27  &r37  &r47  &r57  &r67  1    &r78,
      &r18  &r28  &r38  &r48  &r58  &r68  &r78  1};

* Sample size per treatment arm of each endpoint;

```

```

nnm={&n1, &n2, &n3, &n4, &n5, &n6, &n7, &n8};
mu=j(ntests, 1, 1);

mu=theta/sqrt(2/nnm);

alpha=&alpha;

* Generate p-values;
yval=j(&nsims, ntests, 0);
sigma=j(ntests, 1, 1);
d=diag(sigma);
drd=d*rho*d`;
t=half(drd);
do i=1 to &nsims;
  z=rannor(j(ntests, 1, 1234*i));
  y=mu+t`*z;
  yval[i,]=y`;
  p0=(1-probnorm(yval));
end;

```

The estimated endpoint p-values are stored in matrix  $p0$ . In addition, this matrix totally holds  $nsims$  groups of endpoint p-values, where  $nsims$  is the number of simulation loops.

After the estimated p-values are generated, the following code can be used to predict the success rate (power) of each point:

```

testrslt=j(&nsims, ntests, 0);

do j=1 to &nsims;
  a0={&a1 &a2 &a3 &a4 &a5 &a6 &a7 &a8};
  a=a0/(1/(&alpha/2));
  w={&w11 &w12 &w13 &w14 &w15 &w16 &w17 &w18,
    &w21 &w22 &w23 &w24 &w25 &w26 &w27 &w28,
    &w31 &w32 &w33 &w34 &w35 &w36 &w37 &w38,
    &w41 &w42 &w43 &w44 &w45 &w46 &w47 &w48,
    &w51 &w52 &w53 &w54 &w55 &w56 &w57 &w58,
    &w61 &w62 &w63 &w64 &w65 &w66 &w67 &w68,
    &w71 &w72 &w73 &w74 &w75 &w76 &w77 &w78,
    &w81 &w82 &w83 &w84 &w85 &w86 &w87 &w88};
  p=p0[j, 1:ntests];
  testrslt [j, 1:ntests]=graphapp(a, w, p);
end;

* power;
power= testrslt [:,];
power=t(power);

```

The rejection status of each null hypothesis is stored in the matrix  $testrslt$ . The success rate (power) of each endpoint is stored in the vector  $power$ .

## FAMILY-WISE GATEKEEPING APPROACH

In addition to graphical approach, other approaches to perform multiplicity controlled analyses have also been proposed over the last decade. Different approaches have their own advantages and have found applications in different kinds of clinical trials. The family-wise gatekeeping approach implemented in SAS/IML is also included here for the purpose of comparing the analysis results using different approaches.

The following family-wise gatekeeping procedure is based on the methodology developed in Dmitrienko, Tamhane and Wiens (2008) and Dmitrienko and Tamhane (2011). It has been used to control the overall Type I error rate in the clinical trial designed to test high dose and low dose versus placebo. In this procedure, the endpoints are grouped into families. For example, we can put the primary endpoint in Family 1, put key secondary endpoints in Family 2, and put other secondary endpoints in Family 3.

In Family 1, the tests can be conducted using single-step Dunnett procedure (Dunnett 1955) at a two-sided  $\alpha$ . The following is the code to test high dose and low dose versus placebo (Dmitrienko 2009):

```
data family1;
  n=&n1;
  m=&numtrt;
  alpha=&alpha1;
  crit=probmcc("DUNNETT1", ., 1-alpha, (m+1)*(n-1), m);
  siglev=2*(1-probt(crit, 2*(n-1))); ***two-sided significance level;
run;
```

The tests in Family 1 will then be performed at the calculated significance level *siglev*.

The Family 2 tests are based on truncated Hommel procedure (Brechenmacher 2011) at a significance level  $\alpha$  which is determined by the significance of the tests in Family 1. The algorithm of truncated Hommel procedure can be implemented using the following SAS/IML code (Dmitrienko 2009):

```
* Compute Simes p-value for a given intersection hypothesis (HINTS);
start simesp(p,hints,gamma); *p is p-value vector;
  p1=t(p[loc(hints)]);
  m=ncol(hints);
  k=sum(hints);
  p2=p1;
  r=rank(p1); *sort the p-value;
  p1[r]=p2;
  num=p1; *numerator;
  index=(1:k);
  denom=index*gamma/k+(1-gamma)/m; *denominator;
  simesp=min(num/denom); *Simes p-value;
  return(simesp);
finish simesp;

* truncated Hommel test for the given p-values, significance level and gamma;
start hommel(p,siglev,gamma);
  nhyps=ncol(p); *number of hypotheses;
  if nhyps=1 then adjp=p;
  if nhyps>1 then do;
    nints=2**nhyps-1; *number of intersection hypotheses;
    hints=j(nints,nhyps,0); *intersection hypotheses;
    * Assign intersection hypotheses;
    do i=1 to nhyps;
      do j=0 to nints-1;
        k=floor(j/2**(nhyps-i));
```

```

        if k/2=floor(k/2) then hints[j+1,i]=1;
    end;
end;
p1=j(nints,nhyps,0);
do i=1 to nints;
    p1[i,]=hints[i,]*simesp(p,hints[i,],gamma); *Simes p-values;
end;
adjp=j(1,nhyps,0);
do i=1 to nhyps;
    adjp[i]=max(p1[,i]); *adjusted p-values;
end;
end;
rslt=(adjp<=siglev); *hypothesis tests based on adjusted p-values;
return(rslt);
finish homm1;

```

In the code above, parameter  $p$  is the vector of endpoint p-values in Family 2,  $siglev$  is the  $\alpha_2$  value of Family 2 tests, and  $gamma$  is a value between 0 and 1. The test results are stored in the return matrix named  $rslt$ .

The tests in Family 3 will be carried out using Hommel procedure (Hommel 1988) at significance level  $\alpha_3$  which will be determined by the significance of the tests in Family 2. The Hommel procedure is a special case of the truncated Hommel procedure with  $\gamma = 1$ , and the tests in Family 3 can be done by calling:

$Hommel(p, alpha3, 1)$ ; where  $p$  is the vector of the p-value of each tests in Family 3, and  $alpha3$  is the  $\alpha_3$  value used for the Family 3 tests.

Similarly, for family-wise gatekeeping approach, the endpoint p-values can be estimated using the same approach mentioned above. Also, the success rate of each endpoint can be predicted to help the early decision-making.

The results can be further processed for different kinds of applications. For example, the predicted success rate of each endpoint using different approaches can be displayed in bar charts (Figure 3).

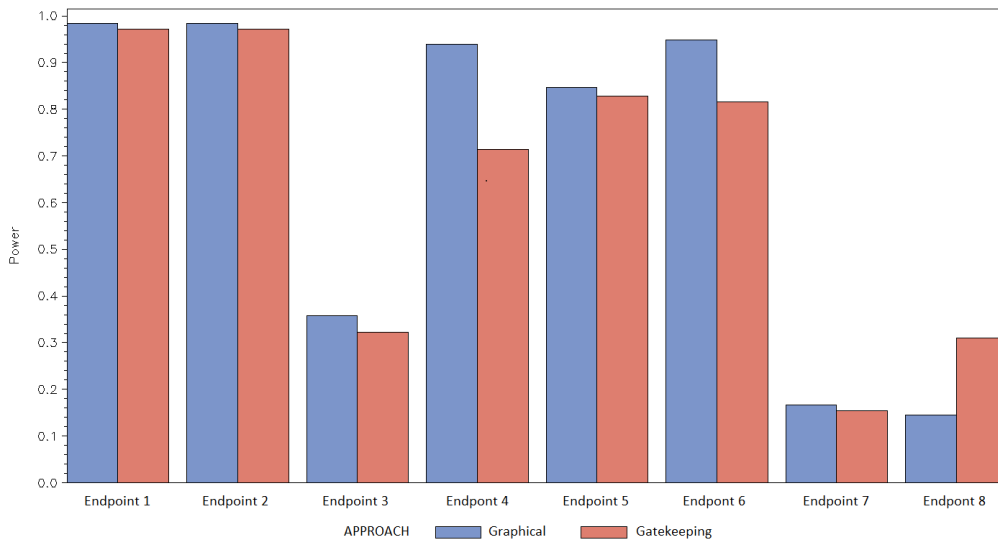


Figure 3. Comparison of Different Approaches

## CONCLUSION

The programs presented in this paper are very useful to perform multiplicity controlled analyses. They can not only be used to identify the successfulness of the clinical trials with multiple objectives, in addition, the simulation functions can also be used to predict the success rate of each endpoint in the clinical trials at the early stage of clinical trials. Furthermore, the ability to compare different approaches in multiplicity controlled analyses makes it possible to provide additional information to help the decision-making.

## REFERENCES

- Alosh, M., Bretz, F., Huque, M. 2014. "Advanced multiplicity adjustment methods in clinical trials." *Statistics in Medicine*, 33:693-713.
- Brechenmacher, T., Xu, J., Dmitrienko, A., and Tamhane, A. 2011. "A mixture gatekeeping procedure based on the hommel test for clinical trial application." *Journal of Biopharmaceutical Statistics*, 21:748-761.
- Bretz, F., Maurer, W., Brannath, W., Posch, M. 2009. "A graphical approach to sequentially rejective multiple test procedures." *Statistics in Medicine*, 28:586-604.
- Burman, C.F., Sonesson, C., Guilbaud, O. 2009. "A recycling framework for the construction of Bonferroni-based multiple tests." *Statistics in Medicine*, 28:739-761.
- Dmitrienko, A., Tamhane, A., Wiens, B. 2008. "General multistage gatekeeping procedures." *Biometrical Journal*, 50:667-677.
- Dmitrienko, A., Tamhane, A., and Bretz, F. 2009. *Multiple Testing Problems in Pharmaceutical Statistics*. Boca Raton, FL: Chapman & Hall/CRC.
- Dmitrienko, A., Tamhane, A. 2011. "Mixtures of multiple testing procedures for gatekeeping applications in clinical trials." *Statistics in Medicine*, 30:1473-1488.
- Dunnnett, C.W. 1955. "A multiple comparison procedure for comparing several treatments with a control." *Journal of the American Statistical Association*, 50:1096-1121.
- Hommel, G. 1988. "A stagewise rejective multiple test procedure based on a modified Bonferroni test." *Biometrika*, 75:383-386.

## CONTACT INFORMATION

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

Xingxing Wu  
Eli Lilly and Company  
wu\_xingxing@lilly.com

Hangtao Xu  
Eli Lilly and Company  
xu\_hangtao@lilly.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.