Paper 184-30

Empirical Power for Higher-Order Crossover Designs in Comparative Bioavailability Clinical Trials

Jihao Zhou, Ying Yuan*, Melissa Chen**, and Bruce Coate

Pfizer Global Research and Development La Jolla Laboratories, San Diego, CA
* The University of Michigan, Ann Arbor, MI
** Prosanos Corporation, La Jolla, CA

Abstract

A standard 2x2 crossover design has been widely used in drug clinical trials, but it is not optimal because of some undesirable statistical properties such as the low power to test carry-over effect. Higher-order crossover designs address this defficiency and, therefore, have drawn more attention recently. However, exact power calculation of higher-order crossover designs poses difficulty for users. It requires sophisticated numerical integration, and it is not available in statistical software packages. This paper presents a SAS Macro to compute the empirical power for higher-order designs in comparative bioavailability clinical trials. The power curve and tables for higher-order crossover design with 2 sequences and 3 periods are presented.

Keywords: Higher-order crossover design, empirical power, comparative bioavailability

Introduction

Cross-over designs have been widely used in drug clinical trails [1, 2, 3]. Among crossover studies, the 2x2 crossover design is most popular [1, 2], but it is not optimal for 2-treatment crossover clinical trials because of some undesirable statistical properties [2, 4]. For example, the sequence effect is aliased with carryover effect. If the carryover effect is not balanced between sequences, the unbiased estimates for the direct drug effect do not exist. In addtion, the test for the carryover effect lacks power, since it is based on the between-subject comparison. The higher-order designs address these defficiencies of the 2x2 crossover design, mainly by increasing degrees of freedom, which allows more precise estimation of the carry-over effect based on the within-subject comparison. Some higher-order designs, such as Balaam's design, are optimal in the sense that they yield the minimum variance unbiased estimator of both the direct and carry-over effects [2]. The high cost of potential later-phase trial failures in pharmaceutical industry also dictates efficient deisgns in the early phase trials. Consequently, the higher-order designs are being used increasingly in drug clinical trials.

The exact power calculation for the higher-order crossover designs requires sophisticated numerical integration and it is not available in statistical software packages [4]. A Monte Carlo simulation is a useful alternative approach to the power calculation in this particular situation [5]. This paper presents SAS Macros to compute the empirical power of the higher-order designs under different scenarios in comparative bioavailability clinical studies using Monte Carlo simulations.

Background

Higher Order Cross Over Designs

A higher-order crossover design is defined as a crossover design in which either the number of periods or the number of sequences is greater than the number of treatments (or chemical formulations) to be compared [4]. For a 2-treatment comparison, the commonly used designs are a 4-sequence, 2-period design (ie, Balaam's design), a 2-sequence, 3-period design (extended period design), and a 2- or 4- sequence, 4-period design.

Comparative Bioavailability Studies

A comparative bioavailability study (sometime called a market image study) is required as the regulatory standard for demonstrating that test (T) and reference (R) drug products will offer the same therapeutic benefit and safety profile when used in the marketplace. T and R can be a to-be-marketed dosage formulation versus a clinical trial drug formulation respectively, a generic drug versus a reference-listed drug, a drug product changed after approval versus an approved drug product, etc [6]. The objective of such study is to demonstrate whether the average pharmacokinetic parameters, such as area under the drug concentration-time curve (AUC) and peak

drug concentration (Cmax), are comparable after administration of the T and R products. The preferred statistical method is the 2 One-Sided Tests (TOST) procedures [7].

Two One-Sided Tests Procedure

Let $\mu_T - \mu_R$ be the difference of the pharmacokinetic parameter means, such as AUC and Cmax, between a treatment (a new drug or a new formulation) and a reference (a generic drug or an existing/approved formulation) depicted as above, θ_L denotes the lower no-effect boundary, and θ_U denotes the upper no-effect boundary, then the study objective of the comparative bioavailability can usually be tested in the following 2 one-sided hypotheses:

H_{01}: $\mu_T - \mu_R \le \theta_L$ or
H_{02}: $\mu_T - \mu_R \ge \theta_U$

H_a: $\theta_L < \mu_T - \mu_R < \theta_U$

Assuming logarithmically transformed data would be normally distributed, then H_0 is rejected at significance level α and no effect (equivalence) is concluded if 100 (1-2α) % confidence interval (CI) $(\hat{\mu}_T - \hat{\mu}_R - t_{1-\alpha,n-1} s\hat{e},\ \hat{\mu}_T - \hat{\mu}_R + t_{1-\alpha,n-1} s\hat{e})$ of the mean $\mu_T - \mu_R$ is entirely within (θ_L, θ_U), otherwise H_0 fails to be rejected. The choices of no-effect boundaries depend on the nature of the drugs studied. The commonly used θ_L and θ_U are ($\log_e$80%, $\log_e$125%) and ($\log_e$77%, $\log_e$130%). The type-I error α of the TOST procedure is often set as 5%. There is no closed algebraic formula to calculate the statistical power for the TOST procedure [4, 8].

Methods

Statistical Simulations

The Monte Carlo simulation involves random sampling techniques to generate a series of random samples from a distribution that represents the study population of interest (eg, the population under H_a, or no drug-drug interaction exists). For each generated random sample, the TOST procedure is applied and the conclusion of rejecting or accepting H_0 is made. The empirical power of the TOST procedure is then calculated as the proportion of the replications in which H_0 is rejected. Herein, we generate *m* (eg, 1000) samples of size *n* (eg, 12). To save the computation time, we generate all these *m* samples by calling a multivariate normal random number generator (a SAS Macro) once, then analyze the samples in sequential orders. The computational process is outlined as follows:

Step 1. Generate *m* random sample of size *n* according to the following multivariate nomal distribution with pre-specified mean vector μ and covariance matrix Σ,

$$X \sim MVN(\mu, \Sigma)$$

Step 2. Run PROC MIXED on the *i*th random sample to calculate the 100 (1-2α) % CI for $\mu_T - \mu_R$, for *i*=1 to *m*.

Step 3. The empirical power is calculated as the proportion of *m* random samples in which the CI falls entirely within pre-specified (θ_L, θ_U).

SAS Program Development

I). Two-sequence, 3-Period Crossover Design (2x3 Design)

First, the macro deletes data set *empower,* which is used to store coverage results, if any exists from previous macro calls. The SAS Macro MVN() [9] is called 2 times to generate *(mn)/2* random normal trivariates for the first sequence (RTT) and *(mn)/2* random normal trivariates for the second sequence (TRR), based on the multivariate normal distribution with mean vector μ and compound symmetric covariance matrix Σ under H_a. These random trivariates are stored in a dataset called *samples*. For the *i*th iteration of total *m*, a SAS dataset step selects the *i*th sample from the dataset *samples*, then PROC TRANSPOSE and a SAS dataset step are used to rearrange the data in a format that is suitable for PROC MIXED. The 100 (1-2α) % CI for the direct treatment effect is obtained by applying PROC MIXED on the *i*th sample and saved into a dataset *empower*. PROC APPEND is then used to concatenate all *m* resulting

data sets. SAS data step is used to produce a binary (0/1) variable to indicate whether the CI from each simulation falls within the criteria (θ_L, θ_U). If it does, it means the H_0 is rejected. Finally, SAS PROC MEANS is used to obtain the proportion of rejecting H_0 for all simulations (ie, the empirical power of the TOST procedure), and keep the result in dataset *power*.

```
%Macro xover2x3(meand=0,runs=1000,R_Mean=1.521, T_mean=1.521, SD=0.338, N=24, r=0.40,
alpha=0.10, lower=0.80, upper=1.25);
  * delete dataset empower;
  proc datasets library=work nolist;
    delete empower;
  run;
  quit;

  * Store the variance-covariance matrix in a data set ;
  data varcov;
   m1=&sd*&sd;     m2=&sd*&sd*&r;  m3=&sd*&sd*&r;  output;
   m1=&sd*&sd*&r;  m2=&sd*&sd;     m3=&sd*&sd*&r;  output;
   m1=&sd*&sd*&r;  m2=&sd*&sd*&r;  m3=&sd*&sd;     output;
  run;

  *Store the mean vector in data sets Means1 for the first sequence
  *(RTT) and means2 for the second sequence (TRR);
  data means1;
   m1=&R  mean; output;
   m1=&T  mean; output;
   m1=&T_mean; output;
  run;

  data means2;
   m1=&T  mean; output;
   m1=&R  mean; output;
   m1=&R_mean; output;
  run;

  **** Monte Carlo simulation starts ***;
  * Simulate multivariate normal variables using SAS Macro MVN;
  * The MVN Macro (source: SAS Techinical Support);

  * generate multivariate normal variables for the first sequence;
  %mvn(varcov=varcov,means=means1,n=%eval(&n*&runs/2),sample=all1,
      seed=0)

  * generate multivariate normal variables for the second sequence;
  %mvn(varcov=varcov,means=means2,n=%eval(&n*&runs/2),sample=all2,
      seed=2004)
  * Combine the data for two sequences
  data all1;
    set all1;
    id=_n_;
  run;

  data all2;
    set all2;
    id=_n_;
  run;

  data samples;
    set all1 all2;
  run;

  proc sort;
    by id;
```

```
run;

proc datasets nolist;
  delete all1 all2;
run;
quit;

   * Add an iteration indicator variable (Iteration) into the generated
     data set;
data samples;
  set samples(drop=id);
  Iteration=int((_n_-1)/&n);
run;

* Perform analysis on each sample;

%do ii=1 %to &runs;
  * Assign subject id, and sequence;
  data b;
    set samples(where=(iteration=%eval(&ii-1)));
          subject= n ;
          sequence=mod(_n_,2)+1;
  run;

  * Sort the data;
  proc sort data=b;
    by sequence subject;

  * transpose the dataset b;
  proc transpose data=b out=sample(rename=col1=AUC
                                   rename=_name_=treat);
    by sequence subject;
          var col1 col2 col3;
  run;

  * Assign treatment and period variables;
  data sample;
    length sequence 3. subject 3. period 3. treat $8.;
          set sample;

          * assign period;
          if treat='COL1' then period=1;
          if treat='COL2' then period=2;
          else if treat='COL3' then period=3;

          * assign treatment code;
          if sequence=1 and period=1 then treat='T';
          else if sequence=1 then treat='R';
          if sequence=2 and period=1 then treat='R';
          else if sequence=2 then treat='T';
  run;

  * Use SAS ODS to save resources;
  ods exclude all;
  ods results=off;

  * use SAS procedure specific ODS to store the results;
  ODS OUTPUT diffs=epower(keep=Lower Upper);

  * use PROC MIXED to analyze the data to produce CI;
  proc mixed data=sample;
    class sequence subject period treat;
```

```
        model auc=sequence period treat/ddfm=kenwardroger;
        random subject(sequence);
        lsmeans treat/pdiff cl alpha=&alpha;
        estimate 'Comparative Bioavailability' treat -1 1;
 run;

 ods exclude none;
 ods results;
 ods listing;

     * Test the existence of the empower and do the appending;
     %if %sysfunc(exist(empower)) %then
   %do;
          * append the results;
          proc append base=empower data=epower force;
          run;
        %end;
 %else
   %do;
          * rename epower if empower does not exist;
          proc datasets library=work nolist;
            change epower = empower;
          run;
          quit;
        %end;
%end;

* Create an indicator variable for empirical power;
data prop;
  set empower;
  * if indi=1, reject the null hypothesis, otherwise not;
  indi=(lower > log(&lower) and upper < log(&upper));
run;

* Calculate empirical power;
proc means data=prop noprint;
  var indi;
  output out=propall(drop=_type_ _freq_) mean=power_x;
run;

* Output the Macro parameters;
data power;
  set propall;
  n=symget('N');
  SD=symget('SD');
  meand=1.0*symget('meand');
  runs=symget('runs');
  alpha=symget('alpha');
  lower=symget('lower');
  upper=symget('upper');
  r=symget('r');
run;

%Mend xover2x3;

* The end of this SAS program;
```

II). Other Higher-Order Crossover Designs

The SAS macro XOVER2X3 can be easily adopted to other higher-order crossover designs. For example, for 2x4 design, we only need to change the dimension of the random multivariate normal vector. Some dataset steps are correspondingly modified.

The SAS Macro POWER

The SAS Marco POWER is used to generate the power curve and power table for the 2x3 design. Specifically,

(1) it produces the plot of the empirical power spanning the parameter space (θ_L, θ_U) of the treament population mean difference, by setting the macro parameter PLOT=yes. Here, fifteen equally spaced points are selected from (θ_L, θ_U), including 0 and 2 boundary points. PROC GPLOT is used to present the power curve.

(2) it generates a numerical power table for a certain range of commonly used sample sizes, correlation, variance and/or no-effect boundaries, by setting the parameter TABLE=yes. PROC REPORT and ODS are used to create the table. The Macro POWER is attached at the end of the paper.

Results

Figure 1 shows that the power curve of the higher-order crossover design with 2 sequences and 3 periods. An 80% power reference line is shown, since comparative bioavailability studies generally require the power of at least 80%. The power curve is symmetric and it achieves the maximum when the mean difference is 0, consistent with the theory of Schuirmann's TOST procedure. It is somewhat risk to use the power given mean difference is 0. In reality, instead of using the power at 0 mean difference, we usually estimate the power given a small nonzero mean difference, such as 0.025, to get a relatively conservative power estimate. Table 1 presents the empirical power estimates for commonly used sample sizes, CVs and no-effect boundaries.

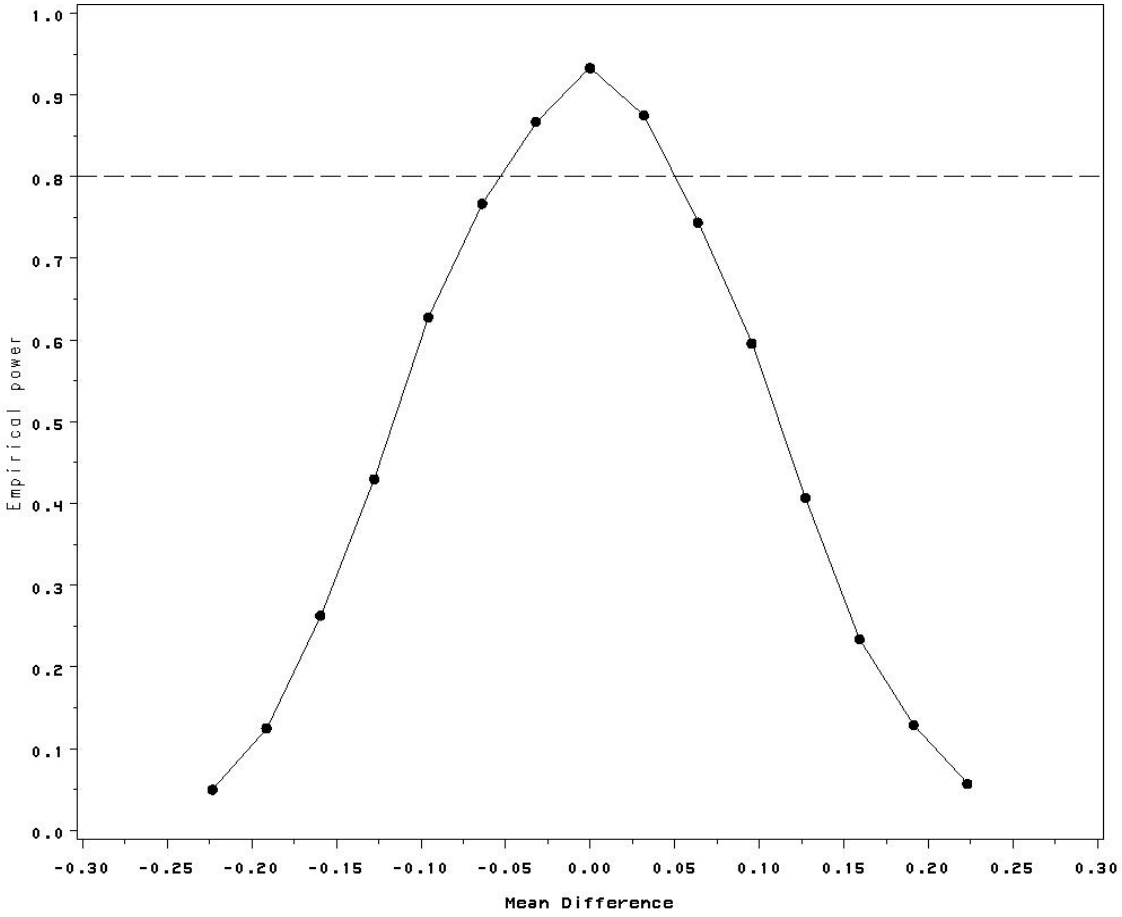
Table 1. Empirical Power for the Higher-Order Crossover Design (2x3) With $\mu_T - \mu_R$ = 0 Based on 1000 Simulations

			Standard Deviation					
Criteria	**Correlation**	**N**	**0.238**	**0.288**	**0.338**	**0.388**	**0.438**	**0.488**
(0.80, 1.25)	0.2	16	0.91	0.72	0.52	0.33	0.17	0.06
		20	0.96	0.86	0.69	0.51	0.31	0.15
		24	0.99	0.91	0.80	0.65	0.47	0.30
	0.4	16	0.98	0.88	0.71	0.55	0.36	0.21
		20	0.99	0.96	0.85	0.69	0.50	0.39
		24	1.00	0.98	0.92	0.80	0.66	0.50
	0.6	16	1.00	0.98	0.90	0.81	0.63	0.51
		20	1.00	0.99	0.97	0.90	0.80	0.64
		24	1.00	1.00	0.99	0.96	0.88	0.78

Conclusions

The Monte Carlo simulation is a useful approach for estimating the power of a variety of statistical tests [5], especially when there are no closed formulae for power calculation. SAS is a practical tool for carrying out this simulation. In this paper, we present a SAS macro to estimate the power of the Schuirmann's TOST procedure in higher-order crossover designs for comparative bioavailability clinical studies. The power curve and tables for the higher-order crossover design with 2 sequences and 3 periods are presented. In the present study, we did not consider carryover effects. The empirical power of detecting carryover effects for higher-order effect designs will be studied in future.

Figure 1. Empirical Power Curve for the Higher-Order Crossover Design (2x3) Based on 1000 Simulations



Acknowledgments

We are grateful to Alex Kuperman at Pfizer Inc for valuable suggestions in preparation of this paper.

References

1. Huang S-M, Lesko LJ, and Williams, RL. Assessment of the quality and quantity of dug-drug interaction studies in recent NDA submissions: Study design and data analysis issues. *Journal of Clin Pharmacology* 1999; 39: 1006-1014.

2. Jones B and Kenward MG. Design and Analysis of Cross-Over Trials. Chapman & Hall/CRC 2003.

3. Patterson SD and Jones B. Bioequivalence and the pharmaceutical industry. *Pharmaceutical Statistics* 2002; 1: 83-95.

4. Chow, S-C and Liu J-P. *Design and Analysis of Bioavailability and Bioequivalence Studies*. Markel Dekker, Inc., New York, NY. 2000.

5. Zhou J, Coate B, and He Y. Empirical Power Computation Using SAS for Schuirmann's Two One-Sided Tests Procedure in Clinical Pharmacokinetic Drug-Drug Interaction Studies. 2004 *SUGI 29 Proceedings*, Paper 154-29.

6. FDA Guidance for Industry. *In Vivo Drug Metabolism/Drug Interaction Studies – Study Design, Data Analysis, and Recommendations for Dosing and and labeling.*. FDA, Center for Drug Evaluation and Resaerch (CDER) and Center for Biological Evaluation and Resaerch (CBER), November 1999.

7. Schuirmann DJ. A Comparison of the Two One-Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Bioavailability. *Journal of Pharamcokinetics and Biopharmaceutics* 1987; 15(6): 657-680.

8. Phillips, K. Power of the Two One-Sided Tests Procedures in Bioequivalence. *Journal of Pharamcokinetics and Biopharmaceutics* 1990; 18(2): 137-144.

9. The MVN Macro, http://ftp.sas.com/techsup/download/stat/mvn.html. The SAS Institute Inc., Cary, NC.

Contact Information

J. Zhou and B. Coate can be reached at Pfizer Inc., 11085 Torreyana Road, San Diego, CA 92121. Y. Yuan can be reached at the Department of Biostatistics, The University of Michigan, Ann Arbor, MI 48109. M. Chen, Prosanos Corporation, La Jolla, CA 92037.

Attachment

```
*************************************************************************************************************
*   SAS Program: Power.sas                                                                                  *
*   Function: Generate empirical power curve plot and report for 2x3 cross-over design by calling the Macro*
*                     XOVER2X3.sas                                                                          *
*   Remark:         SAS V8.2                                                                                *
*************************************************************************************************************;
options sasautos=('C:\temp\') nosource nonotes  nosymbolgen nomlogic nomprint nodate nonumber
nocenter;

%Macro POWER(plot=no, table=yes, sampleN=12, alpha=0.05, runs=200, lower=0.80, upper=1.25,
npoints=15, R_Mean=1.521, SD=0.356, r=0.40);

  *Delete dataset p_power if any ***;
  %if %sysfunc(exist(p_power)) %then
   %do;
     Proc datasets library=work nolist;
      delete p_power;
           run;
           quit;
           %end;
  %if %upcase(&plot)=YES %then
   %do;
    * The Macro Call to Make Power Curve ***;
    %do number=1 %to &npoints;
     * Create meand for making plots;
           %let plot_mean=log(&lower) + (&number - 1) * (log(&upper)
                -log(&lower)) /(&npoints - 1);
            * To trick the calculation;
            data _null_;
           y=&plot_mean;
            call symput('meand',y);
           run;

           %xover2x3(meand=&meand, runs=&runs,R_Mean=&R_Mean,
           T_mean=%sysevalf(&R_mean+&meand),SD=&SD,
           N=&sampleN, r=&r, alpha=&alpha, lower=&lower,
           upper=&upper)

           * Test the existence of the power and do the appending;
           %if %sysfunc(exist(p_power)) %then
            %do;
             * append the results;
             proc append base=p_power data=power force;
             run;
```

```
        %end;
       %else
    %do;
     * rename power if p_power does not exist;
            proc datasets library=work nolist;
                  change power = p_power;
            run;
            quit;
        %end;
  %end;

  goptions reset=all;
  symbol1 interpol=line value=dot height=1 c=black;
  legend across=1
  cborder=blue
  position=(top inside right)
  mode=share;

  legend2;
  axis1 label=('Mean Difference' height=2.5 ) order=(-0.30 to
        0.30 by 0.05)
  minor=(number=1);
  axis2 label=(angle=90 'Empirical power' height=2.5)
  order=(0.0 to 1.0 by 0.1)
  minor=(number=1);
  proc gplot data=p_power;
   plot power*x * meand/ haxis=axis1 vaxis=axis2 vref=0.80
                lvref=2 ;
  run; title; quit;
 %end;

* delete p_table if any;
%if %upcase(&table)=YES %then
 %do;
  %if %sysfunc(exist(p_table)) %then
   %do;
    Proc datasets library=work nolist;
     delete p_table;
    run; quit;
   %end;

        * Set n=16 to 28 by 4;
          %do n=16 %to 24 %by 4;
           * Set standard deviation = 0.238 to 0.488 by 0.05
           %do SD1000=238 %to 488 %by 50;
            * Set correlation coefficient = 0.2, 0.4, 0.6.
            %do r10=2 %to 6 %by 2;
             * Call Macro XOVER2X3 to calculate empirical power
      %xover2x3(meand=0, runs=&runs,R_Mean=&R_Mean,
            T_mean=&R_mean,
            SD=%sysevalf(&SD1000/1000), N=&n,
            r=%sysevalf(&r10/10),alpha=&alpha,
            lower=&lower, upper=&upper)

      * Test the existence of the power and do the
        appending;
      %if %sysfunc(exist(p_table)) %then
       %do;
                        *append the results;
                        proc append base=p_table data=power force;
                        run;
       %end;
```

```
        %else
          %do;
                              * rename power if p  table does not exist;
                              proc datasets library=work nolist;
                                change power = p_table;
                              run;
                              quit;
                            %end;
        %end;
      %end;
    %end;

  * make the report;
  data power  all;
    set p  table;
    lcrit=round(lower,0.01);
    ucrit=round(upper,0.01);
    crit="("||trim(left(put(lcrit,4.2)))||",
          "||trim(left(put(ucrit,4.2)))||")";
     n=input(n,best.);
  run;

  proc sort data=power_all out=power_all;
    by lower _n r sd;
  run;

  proc transpose data=power  all
                 out=xpower(where=(_name_="power_x"));
    by lower   n r;
    var power_x;
    id sd;
    copy crit;
  run;

  ods listing close;
  ods rtf file='C:\temp\table.rtf';
  options nodate nonumber orientation=portrait;
  proc report data=xpower nowindows headline headskip spacing=2 split="@" ;
    columns crit r _n ("SD" _0D238  _0D288 _0D338 _0D388  _0D438 _0D488);
    define crit /"Criteria" width=16 order;
    define  r /"Correlation" width=16 group;
    define _n /" N " width=10 order=internal ;
    define  _0D238 /" 0.238" width=6 f=6.2;
    define  _0D288 /" 0.288" width=6 f=6.2;
    define _0D338 /" 0.338" width=6 f=6.2;
    define  _0D388 /" 0.388" width=6 f=6.2;
    define  _0D438 /" 0.438" width=6 f=6.2;
    define _0D488 /" 0.488" width=6 f=6.2;
    break after crit/skip;
    title; footnote;
  run;
  ods rtf close;
  ods listing;
%end;

%Mend POWER;
```