PAPER SP02

SAS@ APPLICATION IN 2 * 2 CROSSOVER CLINICAL TRIAL

Wuwei Wayne Feng, Eli Lilly & Company, Indianapolis, IN
Dong Ding, MedFocus Inc. Des Plaines, IL

ABSTRACT

Crossover clinical trial design attracts a lot of attention in today's drug development environment because of its unique characteristics. This paper will start with an introduction of a 2 * 2 crossover clinical trial design, then a review of advantages and disadvantages of cross over studies compared with parallel studies. Different types of SAS@ procedure, such as PROC TTEST, PROC GLM and PROC MIXED will be used as demonstration on how to analyze the data if response variable is continuous. McNemar test, Mainland-Gart test, Prescott test will also be reviewed if the outcome measurement is a binary variable. The paper closes with an example using simulated data.

INTRODUCTION TO CROSSOVER DESIGN

Crossover study designs are applied in pharmaceutical industry as an alternative to parallel designs on certain disease types. A crossover trial is one in which subjects are given sequences of treatments with the objective of studying differences between individual treatments (Senn, 2002). In crossover design, a patient receives treatments sequentially, and he/she can be considered as his/her own control. Compared to parallel design, crossover design eliminates the between-patient variation, hence is theoretically more powerful than similar sized parallel study. To estimate the treatment effect with similar level of precision, theoretically crossover design requires a reduced sample size, hence is more cost-effective. However, crossover design is only suitable for chronic, stable diseases, where the treatment is not believed to "correct" the cause of the illness, but is considered to alleviate the symptoms, such as diabetes, asthma, hypertension (Wood, et al., 1989). Inappropriate application of crossover design will complicate the statistical analysis and make the interpretation of the result difficult.
One of the major problems in crossover design is the period by treatment interaction, the carry-over effect. Carry-over is the persistence (whether physically or in terms of effect) of a treatment applied in one period in a subsequent period of treatment (Senn, 2002). To minimize any possible carry-over effect, usually between the treatments there are washout periods during which the effect of a treatment given previously is believed to disappear. If it is believed that carry-over effect cannot be fully eliminated, the study would not be a good candidate for crossover design.

THE GRIZZLE MODEL

Since 2 * 2 crossover design is the simplest and most commonly used crossover design, we will focus on 2 * 2 crossover design and the application of SAS@ software in data analyses in this paper. In 2 * 2 crossover design, patients are randomly allocated to two study sequence, AB and BA, where the patients in the sequence AB receive treatment A first, followed by treatment B, and vice versa in the BA sequence, as illustrated in Figure 1.

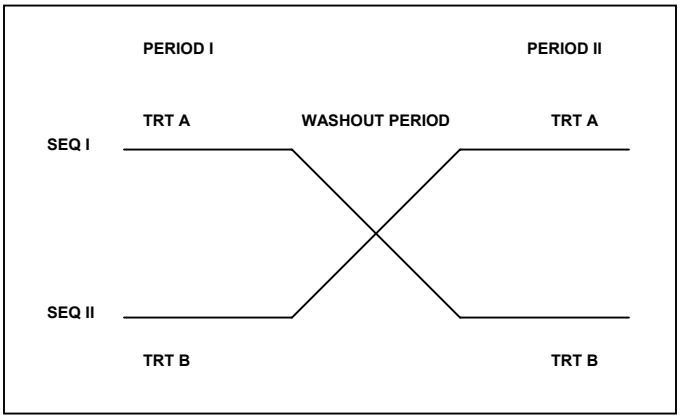

Figure 1. Illustration of a 2 * 2 crossover study design

An adequate model from Grizzle (Grizzle, 1965) is

$$y_{ijk} = \mu + b_{ij} + \pi_k + \phi_l + \lambda_{l'} + e_{ijk}$$

$$j = 1,2,...,n_i; i = 1,2; k = 1,2; l,l' = 1,2 \qquad (1)$$

where μ is a general mean, b_{ij} is the random effect for the *j*-th patient within the *i*-th sequence, π_k is the effect of the *k*-th period, ϕ_l is the direct effect of the *l*-th drug, $\lambda_{l'}$ is the residual effect of the l'-th drug, and e_{ijk} reflects random error in the measurement of the response. When the b_{ij} and the e_{ijk} are each normally distributed as $N(0,\sigma_b^2)$ and $N(0,\sigma_e^2)$, respectively, and are mutually independent, Grizzle discusses tests of hypotheses pertaining to the direct effects, residual effects, and period effects (Grizzle, 1965).

From (1), the sum of the two observations on the same patient is given by

$$y_{ij1} + y_{ij2} = 2(\mu + b_{ij}) + (\pi_1 + \pi_2) + (\phi_1 + \phi_2) + \lambda_i + e_{ij1} + e_{ij2} \qquad (2)$$

where λ_i represents the residual effect of the *i*-th drug in the sequence *ii'* . Hence the sum of the two observations can be used for the test of the hypothesis of no residual effects ($H_{0\lambda} : \lambda_1 = \lambda_2$) .

If there are no residual effects in the above test, then the differences between the two observations on the same patient reduce to

$$y_{ij1} - y_{ij2} = (\pi_1 - \pi_2) + (-1)^{i+1}(\phi_1 - \phi_2) + (e_{ij1} - e_{ij2}) \qquad (3)$$

where $(-1)^{i+1}$ =1 for the sequence AB and $(-1)^{i+1}$ =-1 for the sequence BA. Hence the difference between the two observations on the same patient can be used for the test of the hypothesis of no direct effects ($H_{0\phi} : \phi_1 = \phi_2$) .

ANALYSIS OF CONTINUOUS RESPONSE VARIABLE WITH SAS@ PROCEDURES

For studies where no carry-over effect exists, the differences between the two treatment period within patients can be applied to the testing of treatment effect. When the response variable is a continuous variable and the error terms (the b_{ij} and the e_{ijk}) are mutually independent and normally distributed as specified above, the response variable can be analyzed with linear regression models.

With the SAS@ PROC TTEST procedure, the comparison between two study arms can be used to test for treatment effect:

```
PROC TTEST;
       VAR DIFF;
              CLASS SEQ;
RUN;
```

Where DIFF reflects within patient difference of the two observations, and SEQ is the identifier of treatment sequence (AB or BA). The data needs to be pre-processed to obtain one-patient-per-record structure with the within patient differences calculated.

The Grizzle's model can also be fitted with the ANOVA model with the PROC GLM procedure:

```
PROC GLM:
       CLASS SEQ PATIENT PERIOD TRT;
              MODEL Y=SEQ PATIENT(SEQ) PERIOD TRT /SS3 ;
                     RANDOM PATIENT(SEQ);
                            TEST H=SEQ E=PATIENT(SEQ);
RUN;
```

Where Y is the value of the response variable measured at the end of either period 1 or period 2. The data structure used in fitting the GLM model is similar to the raw data collected from the case report forms, having one record at each treatment period for each patient. The F-test of treatment effect in the GLM model is essentially the same as the t-test of within patient difference by SEQ (a test for direct treatment effect). And in the GLM model specified above, the sequence effect confounds with the carry-over effect. Hence a hypothesis testing of no sequence effect is equivalent to a t-test on the sum of response variable (the test of carry-over effect). Other effects such as investigator site can also be adjusted by adding them into the base model.

The same model can also be fitted with the PROC MIXED procedure:

```
PROC MIXED;
       CLASS SEQ PATIENT PERIOD TRT;
```

```
            MODEL Y=SEQ PERIOD TRT;
                  REPEATED / TYPE=CS SUB=PATIENT(SEQ);
RUN;
```

The model fitting would be the same for the two procedures above, when there's only one measured value for the response variable at the end of each treatment period. When the response was measured repeatedly within a same treatment period, the data can be analyzed with the mixed model repeated measures by slight modification on the above SAS@ procedure. Beside compound symmetry structure, other variance-covariance structures may also be tested and adapted in the model fitting as well.

If the assumption of normal distribution of the error terms is not met, non-parametric analyses can be applied in the analysis (Koch, 1972). The NONPAR1WAY procedure in SAS@ can be used for Wilcoxon rank test, and the PROC GLM procedure can be applied to the ranked differences.

ANALYSIS OF BINARY RESPONSE VARIABLE WITH SAS@ PROCEDURES

In some clinical studies, the response variables are categorized into binary levels, for example, "Improved" or "Not Improved", "Success" or "Failure" or, more generally, "Yes" or "No". Table 1 illustrates the counts of number patients in all combinations of responses in different sequence groups from a computer-generated example trial.

Sequence	Failure in Period 1, Failure in Period 2 (0,0)	Failure in Period 1, Success in Period 2 (0,1)	Success in Period 1, Failure in Period 2 (1,0)	Success in Period 1, Success in Period 2 (1,1)	Total
AB	11	2	7	0	20
BA	6	8	2	4	20
Total	17	10	9	4	40

Table 1. Computer-generated example data

The simplest analysis on binary responses variable for 2 * 2 crossover design is the McNemar's test, when the assumption of no carry-over effect and no period effect holds (Nagelkerke, 1986). If there is neither carry-over effect nor period effect, the patients who either responded to both treatments or failed to both treatments do not provide any information about the superiority of either treatment. So only those patients who respond to one treatment and fail to another treatment are counted in McNemar's test (Table 2).

		Treatment B		
		Failure	Success	Total
Treatment A	Failure	17 (a)	4 (b)	21
	Success	15 (c)	4 (d)	19
Total		32	8	40

Table 2. Example table setting for McNemar's test

Under the null hypothesis of no difference between the two treatments, the numbers of patients dropped into the two categories (Failure, Success) and (Success, Failure) follow a binominal distribution of (b+c,0.5), i.e. they have equal probabilities of showing successful outcome with one treatment while fail in another one. A rejection to the null hypothesis suggests the superiority of the treatment with more patients who succeeded.
When performing the analysis with SAS@, the data need to be pre-processed into one-record-per-patient structure, with variables indicating the responses at the end of each treatment period. The PROC FREQ procedure with the AGREE option can be used for McNemar's test:

```
PROC FREQ;
      TABLE RESPONSEA*RESPONSEB / AGREE;
RUN;
```

In the example above, 19 of 40 patients show different responses to the two treatments, among which 15 succeed in treatment A and fail in treatment B, while only 4 succeed in treatment B and fail in treatment A. McNemar's test rejects the null hypothesis of no difference between A and B with a 2-sided p-value of 0.012. The test result is significantly in favor of treatment A.

In the actual crossover studies, the assumption of no period effect does not always hold. To adjust the period effect, Mainland & Gart designed an alternative classification of the results where patients are classified into 3 categories (Senn, 2002):

preferred the first period, no preference in both periods, preferred the second period (Table 3). Similar to McNemar's test, patients showing no preference are considered as non-informative.

Sequence	Preferred Period 1	Preferred Period 2	Total
AB	7	2	9
BA	2	8	10
Total	9	10	19

Table 3. Example table settings for Mainland-Gart's test

Under the null hypothesis of no treatment effect, the distributions of the patients' preference within each sequence are the same as the marginal distribution across the sequence. A test showing association between the preference and the sequence of treatment suggests the treatment effects of the two treatments are different. The data need to be summarized for within patient preference and the test can be performed with the PROC FREQ procedure with the EXACT or CHISQ option:

```
PROC FREQ;
       TABLE SEQ*PREFERENCE / CHISQ;
RUN;
```

Where SEQ is the sequence indicator and PREFERENCE takes values 1 or 2 for period 1 or 2, respectively. When applying Mainland-Gart's test on the same example data used in McNemar's test, there were 19 patients who contributed to the test and the expected cell counts were smaller than 5. Hence we substituted the "CHISQ" option with the "EXACT" option and the test rejects the hypothesis of no treatment difference with a p-value of 0.023.

In both McNemar's test and Mainland-Gart's test, patients showing no preference between the two periods are excluded from the analysis. The test may be under-power a large proportion of patients are excluded. To make use of the full data, Prescott proposed an alternative method to the Mainland-Gart's test which includes the part of data from patients with tied preferences as a column in the contingency table (Senn, 2002, Table 4). The numbers of patients in each cell of the contingency table follow a hypergeometric distribution. The probability of observing a table with equal or extreme values can be summed up. Rejection of the null hypothesis for no association suggests there is a difference between two treatments. The PROC FREQ procedure with the FISHER EXACT option can be used in analysis.

Sequence	Preferred Period 1	No Preference	Preferred Period 2	Total
AB	7	11	2	20
BA	2	10	8	20
Total	9	21	10	40

Table 4. Example table setting for Prescott's test

In the example data above, similar to Mainland-Gart's test, Prescott's test rejects the null hypothesis of no treatment effect with in favor of treatment A. The test p-value is 0.0426.

CONCLUSION

In this paper, we reviewed the 2 * 2 crossover clinical trial design, advantages and disadvantages of cross over studies compared with parallel studies. Different types of SAS® procedures, including PROC TTEST, PROC GLM, and PROC MIXED are demonstrated to analyze the data if response variable is continuous. McNemar's test, Mainland-Gart's test, Prescott's test are also reviewed when the outcome measurement is a binary variable with computer-generated data as examples.

SAS CODE SUPPORTING THIS PAPER

```
options nocenter nodate mprint pageno=1 nospool papersize=Letter orientation=landscape
topmargin=1 bottommargin=1 leftmargin=1 rightmargin=1;

*** Macro to generate random variables following normal distribution;
proc format;
       value  fsf    0='Failure'
                      1='Success';
       value  seq    1="AB"
                      2="BA";
       value  pref   -1="Preferred Period 1"
                      0="No Preferrence"
                      1="Preferred Period 2";
run;
```

```
title "SAS@ Application in 2 * 2 Crossover Clinical Trial";
footnote1;
footnote2  "Simulated Data";
footnote3 "\\please specify your path here\aaaaaaaa.sas";

*************************************************************************************;
*Macro Random generates a data set with &nobs observations where each number in &varname
*follows N(&mean, &std), user can specify seed for randomization or take the default one;
*************************************************************************************;
%macro random(outdsn,varname,mean,std,size1,size2,seed=,seed2=);
%put NOTE: The Macro Random is developed based on the macro random
           from http://probabilitynet.com/macro1.htm;
%put NOTE: Now Executing Macro RANDOM;

*** if user did not specify seed, use system time;

%if %length(&seed)=0 %then %let seed1=%sysfunc(floor(%sysfunc(datetime())));
%if %length(&seed2)=0 %then %let seed2b=%sysfunc(floor(%sysfunc(datetime())));

data dsn1;
       retain _seed %if &seed^=%str() %then &seed; %else &seed1;;
              do _i = 1 to &size1;
                     &varname = &mean + &std * rannor(_seed);
                            output;
              end;
       drop _seed _i;

*** Below is rearranging the data set 'filename' by using a uniform random number generator;

proc sql;
       create view norsam as
              select *, ranuni(%if &seed2^=%str() %then &seed2; %else &seed2b;)
                     as _ran_ from dsn1
                            order by calculated _ran_;
quit;

*** Below creates new data set of a different sample size ;

data &outdsn;
       set norsam(obs = &size2);
       drop _ran_;

*** rearrange the data set by using a uniform random number generator ;

proc datasets lib=work nolist;
    delete dsn1 norsam;
run;

quit;

%put NOTE: Now Ending Macro RANDOM;
%mend random;

*** Macro to create output tables to fit rtf format;

%macro tabulate(indsn,var1,var2,lab1,lab2,fmt,boxlab,styles);
%put NOTE: Now Executing Macro Tabulate;

proc tabulate data=&indsn style={just=center};
       class &var1 &var2;
              classlev &var1 &var2 /
              &styles;
              table (&var1=&lab1 all="Total"),
                     (&var2=&lab2 all="Total")*(n=''*f=10.)
                     / rts=15 box={label="&boxlab" &styles} misstext="0";
```

```
                    keyword all / &styles;
                    &fmt
run;
%put NOTE: Now Ending Macro Tabulate;
%mend tabulate;

*************************************************************************************;
*Macro Simulate take parameters and generate a dataset for Grizzle's model with two
*treatment-arms having patient effect, random effect, treatment effects, period effect and
*carry-over effect. It runs proc t-tests, PROC GLM, and PROC MIXED model test carry over
*effect and treatment effect, it also run categorical analyses, including McNemar's test,
*Mainland-Gart's test, and Prescott's test, after binary variable is created. At last it
*creates a rtf file for summary tables to be inserted to the paper.
*************************************************************************************;

%macro simulate
       (size=,          /* Sample Size */
       sigmab=,        /* Std of patient effect */
       sigmae=,        /* Std of random effect */
       seedb=,         /* Seed for patient effect */
       seede1=,        /* Seed for error 1 */
       seede2=,        /* Seed for error 2 */
       phy1=,          /* treatment effect of A */
       phy2=,          /* treatment effect of B */
       pai1=,          /* period 1 effect */
       pai2=,          /* period 2 effect */
       lamda=,         /* carry-over effect */
       thresol=,       /* thresold for categorical analysis */
       outlocat=       /* Location of the output tables */);

%put NOTE: Now Executing Macro Simulate;

*** call the macro to generate patient effect terms;

%random(b,b,0,&sigmab,&size,&size,seed=&seedb,seed2=1);

*** call the macro to generate error terms;

random(e1,e1,0,&sigmae,&size,&size,seed=&seede1,seed2=1);
random(e2,e2,0,&sigmae,&size,&size,seed=&seede2,seed2=1);

*** Create the simulated dataset;

data simulate;
       merge b e1 e2;

       patient=_n_;
              if 1<=patient<=%sysevalf(&size/2) then seq=1;
                     else seq=2;
              if seq=1 then do;
                     phy1=%sysevalf(&phy1);
                     phy2=%sysevalf(&phy2);
              end;
              if seq=2 then do;
                     phy2=%sysevalf(&phy1);
                   phy1=%sysevalf(&phy2);
              end;
       pai1=%sysevalf(&pai1);
       pai2=%sysevalf(&pai2);

       if seq=1 then lamda=0;
              else lamda=%sysevalf(&lamda);

       y1=phy1+pai1+b+e1;
       y2=phy2+pai2+lamda+b+e2;
       sum=y1+y2;
```

```
	diff=y1-y2;
	format y1 y2 4.2 ;
run;

*** Analysis for continuous variable;

title2 'T-test for Carry-over Effect';

proc ttest data=simulate ;
	var sum;
		class seq;
run;

title2 'T-test for Treatment Effect';

proc ttest data=simulate ;
	var diff;
		class seq;
run;

data sim2;
	set simulate;
	period=1;
	if seq=1 then trt='A';
		else trt='B';
		y=y1;
	output;

	period=2;
	if seq=1 then trt='B';
		else trt='A';
		y=y2;
	output;
run;

title2 "Model Fitting with the GLM Procedure";

proc glm data=sim2;
	class seq patient period trt ;
		model y= seq patient(seq) period trt / ss3;
		random patient(seq) ;
		test h=seq e=patient(seq);
run;
quit;
title2 "Model Fitting with the MIXED Procedure";

proc mixed;
	class seq patient period trt;
		model y=seq period trt;
		repeated / type=cs sub=patient(seq);
run;

*** Create Binary Outcome Variable;

data sim3;
	set simulate;

	_y1=(y1>&thresol);
	_y2=(y2>&thresol);
		if seq=1 then _a=_y1;
		if seq=1 then _b=_y2;
		if seq=2 then _b=_y1;
		if seq=2 then _a=_y2;
run;

title2 "McNemar's Test";
```

```
proc freq data = sim3;
       table _a*_b / AGREE norow nocol nopercent;
       output out=pmcne mcnemar;
       format _a _b fsf.;
run;

data _null_;
       set pmcne;
       if _n_=1 then call symput('p_mcnem',compress(put(p_mcnem,pvalue7.3)));

data sim4;
       set sim3;
       if _y1=_y2 then prefer= 0;
       else if _y1>_y2 then prefer= -1;
       else if _y1<_y2 then prefer= 1;
       format prefer pref.;
run;

title2 "Mainland-Gart Test";

proc freq data=sim4;
       table seq*prefer / exact expected norow nocol nopercent;
       where prefer ne 0;
       output out=pmgt exact;
run;

data _null_;
       set pmgt;
       if _n_=1 then call symput('p_mgt',compress(put(xp2_fish,pvalue7.3)));

title2 "Prescott Test";

proc freq data=sim4;
       table seq*prefer / fisher exact norow nocol nopercent;
              output out=ppres exact;
run;

data _null_;
       set ppres;
              call symput('p_pres',compress(put(xp2_fish,pvalue7.3)));

data summtbl;
       set sim3;
       result=put(_a,fsf.)!!' in Period 1, '!!put(_b,fsf.)!!' in Period 2     '!!
         "("!!compress(_a)!!","!!compress(_b)!!")";
run;

*** Output ODS tables;

*** Set style;
%let styles=%str(style={just=center cellwidth=2.5cm
     background=white foreground=black});

ods rtf body="&outlocat" NOGTITLE NOGFOOTNOTE;

title2 'Table 1. Computer-generated example data.';

%tabulate(summtbl,seq,result,"","",%str(format seq seq.;),Sequence,&styles)
title2 "Table 2. Example summary table for McNemar's test.";
footnote "p-Value of McNemar's Test is &p_mcnem";
footnote2  "Simulated Data";
footnote3 "\\ please specify path here\aaaaaaaa.sas ";

%tabulate(summtbl,_a,_b,%str({label="Treatment A" &styles}),
          %str({label="Treatment B" &styles}),%str(format _a _b fsf.;),,&styles)
```

```
title2 "Table 3. Example table settings for Mainland-Gart test.";
footnote "p-Value of Mainland-Gart test is &p_mgt";
footnote2 "Simulated Data";
footnote3 "\\ please specify path here\aaaaaaaa.sas ";

%tabulate(sim4,seq,prefer,"","",
     %str(where prefer ne 0; format seq seq.;),Sequence,&styles)
title2 "Table 4. Summary table for Prescott test.";
footnote "p-Value of Prescott Test is &p_pres";
footnote2  "Simulated Data";
footnote3 "\\ please specify path here\aaaaaaaa.sas ";

%tabulate(sim4,seq,prefer,"","",
     %str(format seq seq.;),Sequence,&styles)
footnote;
ods rtf close;

%put NOTE: Now ending Macro Simulate;

%mend simulate;

*** Main program, macro call;
%simulate(size=40,
       sigmab=1.04,
       igmae=1,
       eedb=20031125,
       eede1=11252003,
       eede2=25112003,
       hy1=10.4,
       hy2=9.6,
       ai1=.4,
       ai2=0,
       lamda=.05,
       thresol=11,
       outlocat=%str(\\please specify path here\aaaaaaaa.rtf));
```

REFERENCE

Senn, S, Cross-over Trials in Clinical Trial Research, Second Edition, 2002.
Grizzle, JE, The two-period change-over design and its use in clinical trials. *Biometrics, 1965 (21), 461-480*.
Koch, G, The use of non-parametric methods in the statistical analysis of the two period change-over design. *Biometrics, 1972 (June), 577-578.*
Nagelkerke, NJD et al. The two period binary response cross-over trial. *Biometrics Journal, 1986 (28), 863-869.*
Wallenstein, S & Fisher, AC, The analysis of the two period repeated measurements crossover design with application to clinical trials. *Biometrics, 1977 (March, 33), 261-269.*
Wood, JR, et al. The two period crossover design in medical research. *Annals of Internal Medicine, Volume(110), Number 7, April, 1989.*

ACKNOWLEDGEMENT

I would like to thank Virginia Sutton and Jyoti Rayamajhi for their advices on this paper/presentation, Kory Schuh and Marjorie Gordon for their helps to get through company publication process.

CONTACT INFORMATION

Wuwei (Wayne) Feng MD MS
Lilly Corporate Center, DC 4135
Eli Lilly & Company
Indianapolis, IN, 46285
Phone: 317-651-6570
Fax: 317-651-6208
Email: feng_wuwei@lilly.com

Dong Ding MS
MedFocus, Inc.
2800 River Road, Suite 290
Des Plaines, IL 60018
Phone: 800-256-4625
Email: dding2000@yahoo.com

TRADEMARK