



## Statistical Round

Korean J Anesthesiol 2021;74(4):293-299

<https://doi.org/10.4097/kja.21165>

pISSN 2005-6419 • eISSN 2005-7563

Received: April 15, 2021

Revised: June 14, 2021

Accepted: July 5, 2021

### Corresponding author:

Junyong In, M.D., Ph.D.

Department of Anesthesiology and Pain Medicine, Dongguk University Ilsan Hospital, 27 Dongguk-ro, Ilsandong-gu, Goyang 10326, Korea

Tel: +82-31-961-7875

Fax: +82-31-961-7864

Email: [dragona1@dumc.or.kr](mailto:dragona1@dumc.or.kr)

ORCID: <https://orcid.org/0000-0001-7403-4287>

# Considerations for crossover design in clinical study

Chi-Yeon Lim<sup>1</sup>, Junyong In<sup>2</sup>

<sup>1</sup>Department of Biostatistics, Dongguk University College of Medicine, <sup>2</sup>Department of Anesthesiology and Pain Medicine, Dongguk University Ilsan Hospital, Goyang, Korea

This article introduces a crossover design that is often used in clinical studies, with the advantage of comparing treatment effects within one study subject. In particular, the advantages and disadvantages of the two-period, two-sequence crossover design ( $2 \times 2$  or AB/BA crossover design), which is widely used in clinical practice, are identified, and the elements necessary for analysis are introduced. This article explains the carryover effect, period effect, sequence effect, and period-by-treatment interaction in a crossover design and examines the analysis commands of SAS along with example data. After confirming the carryover effect using a general linear model, the treatment effect is analyzed using a linear mixed effect model.

**Keywords:** Biomedical research; Carryover effect; Crossover design; Linear mixed effect model; Statistics; Washout period.

## Introduction

In a crossover study design, two or more treatments (e.g., drugs, procedures) are provided to subjects at different time periods, and the sequence of treatments is randomized for each subject. This design is applied to several fields such as bioequivalence clinical trials, and the simplest study design is a two-period, two-sequence crossover design (AB/BA design)<sup>1)</sup>. Subjects are randomly allocated to the AB or BA sequence<sup>2)</sup> [1,2]. Subjects in the AB sequence receive treatment A in the first period, and after the effects of treatment A have subsided, they receive treatment B in the second period. The remaining subjects assigned to the BA sequence receive treatment B first and then treatment A (Fig. 1). Therefore, the difference in treatments is derived from a within-subject comparison. On the other hand, in a parallel design, one subject receives only one treatment, therefore, the difference in treatments is derived from a between-subject comparison.

A crossover design has the following advantages. First, the treatment effect is compared within a subject since each subject serves as his or her own control. It removes the inter-subject variability from the comparison between groups and the effect of covariates<sup>3)</sup>

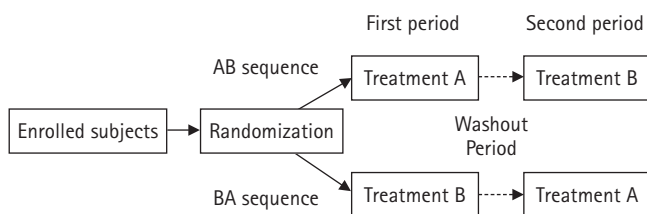
© The Korean Society of Anesthesiologists, 2021

© This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

<sup>1)</sup>A and B usually refer to test group A and control group B. It was introduced as a crossover design including the four orders AA/AB/BA/BB in consideration of the treatment type and treatment sequence; however, a simplified study design is mainly used in clinical studies, excluding AA and BB, which are the sequences in which the same drug is repeatedly administered.

<sup>2)</sup>Depending on the literature, it is also described as a sequence, a group, or an order. In this article, a group of subjects in the same order of receiving treatment is described as a sequence, and a set of data for subjects receiving the same treatment is described as a group.

<sup>3)</sup>In addition to the independent variables, a variable that the researcher wants to control as a factor that can affect the dependent variable is designated a covariate. Variables such as age and gender are often included.



**Fig. 1.** Two-period, two-sequence crossover design.

can be reduced [3]. For this reason, the imbalance of allocation often seen in randomized controlled parallel studies are seldom seen in randomized controlled crossover designs. Second, the crossover design has high power and statistical efficiency. In other words, it is possible to obtain an estimate with the same level of accuracy as a parallel design, even with a smaller number of subjects [1,2,4].

However, there are also limitations in the crossover design. First, the conditions of the subjects must be stable throughout the study. In other words, a case is inappropriate for a study if the disease status changes over time, such as an acute cure, or when symptoms disappear or are cured by treatment in the first period. Second, a washout period may be necessary until the effect of the first treatment subsides. Therefore, if a treatment drug has an extended half-life, it may be difficult to conduct a study with a crossover design. Third, there is a burden that all treatments are carried out on one subject, which often causes ethical problems. Fourth, the processing of dropped or missing data is more problematic than in a parallel design, and the statistical analysis is complex [2,4].

Based on these advantages and disadvantages, it will be helpful in planning and proceeding with an appropriate crossover design to fully understand the various factors encountered in its design and analysis process.

## Standard 2 × 2 crossover design

The standard 2 × 2 crossover design is used to assess between two groups (test group A and control group B). Each subject is randomly allocated to either an AB sequence or a BA sequence. Subjects in the AB sequence receive treatment A at the first period and treatment B at the second period.

A model can be used to describe the standard 2 × 2 crossover design as follows:

(1)

$$Y_{ijk} = \mu + S_{ik} + P_j + T_{jk} + C_{j-1,k} + e_{ijk}$$

Where

$Y_{ijk}$  is the response of the  $i^{\text{th}}$  subject in the  $k^{\text{th}}$  sequence at the  $j^{\text{th}}$  period,

subject  $i = 1, 2, \dots, n_k$  ( $i^{\text{th}}$  subject in  $k^{\text{th}}$  sequence)

period  $j = 1, 2$  (First or second)

sequence  $k = 1, 2$  (AB or BA)

The subject ( $S_{ik}$ ) and error ( $e_{ijk}$ ) are independent and identically distributed random variables and have a normal distribution with mean 0 and variance of  $\sigma_s^2$ , and mean 0 and variance  $\sigma_e^2$ , respectively. This model<sup>4)</sup> includes fixed effects such as the period effects ( $P_j$ ), the direct treatment effects ( $T_{jk}$ ), and the carryover effects ( $C_{j-1,k}$ ). For example,  $T_{1,2}$  represents the direct fixed effect of the treatment at period 1 in sequence 2 (Here, at the first period of the BA sequence), and  $C_{2-1,1}$  is the residual effect carried over from the (2-1)<sup>th</sup> period to the second period in sequence 1 (Here, the AB sequence). The carryover effect in the standard 2 × 2 crossover design can occur at the second period. The fixed effects at each period in each sequence are summarized as follows.

Sequence	First period	Second period
AB sequence	$\mu_{11} = \mu + P_1 + T_A$	$\mu_{21} = \mu + P_2 + T_B + C_A$
BA sequence	$\mu_{12} = \mu + P_1 + T_B$	$\mu_{22} = \mu + P_2 + T_A + C_B$

Here,

$$\mu_{jk} = E(Y_{ijk}), P_1 + P_2 = 0, T_A + T_B = 0, \text{ and } C_A + C_B = 0.$$

## Treatment emergent effects

In studies using a crossover design, elements such as carryover effect, period effect, sequence effect, and period-by-treatment interaction should be evaluated before testing the treatment effect. Even if effects other than the treatment effect were carefully considered and excluded in the research planning stage, it is necessary to check the carryover effect or the period effect before analyzing the treatment effect. After checking the above-mentioned effects are not absent, it is common to analyze the treatment effect. It is important in the crossover study planning stage is to design an analysis that does not become complicated and allow effects other than the treatment effect to influence the interpretation of the results [5,6].

<sup>4)</sup>Subjects are randomly assigned to AB or BA sequence so that the assumption of no sequence effect is established.

## Treatment effect

This refers to the direct effect of the treatment. In this article, it indicates the effects of A and B and  $T_{j,k}$  in Equation (1).

## Period effect

The period effect implies that the effect of the same treatment received at two different periods is different for each period and corresponds to  $P_j$  in Equation (1). Since the first and second treatments are inevitably separated in time, an effect may appear depending on the measurement period, not the treatment. Therefore, when comparing the value obtained by subtracting the first period from the second period of the AB sequence with the value obtained by subtracting the value of the first period from the second period of the BA sequence<sup>5)</sup>, there should be no difference if there is no period effect.

## Carryover effect

The carryover effect, which corresponds to  $C_{j-1,k}$  in Equation (1), refers to the effect of the previous treatment or the change caused by the first treatment continues until the next period and alters the effect of the next treatment. Rather than determining that there is no carryover effect by statistical testing, it is better to select a crossover design when the possibility of a carryover effect is medically unlikely or when the effect can be eliminated through a washout period. Analyzing for the period-by-treatment interaction is used to determine whether the two treatment effects are different in the two periods, and it is difficult to distinguish the carryover effect from the period-by-treatment interaction; therefore, the carryover effect and the period-by-treatment interaction are often treated as identical. However, depending on which parameters are included in the crossover design model, the carryover effect may be embedded in parameters other than the period-by-treatment interaction<sup>6)</sup>. In other words, it is difficult to analyze the carryover effect in the simplified  $2 \times 2$  crossover design [5]; therefore, it is important in the study planning stage to design such that the carryover effect does not occur. For example, there is a method to set a sufficient washout period until the treatment effect or change disappears. In the case of drug studies, a washout period is sometimes set at 3–4 times or more of the blood plasma elimination half-life.

<sup>5)</sup>Both values are the result of treatment B minus the result of treatment A.

<sup>6)</sup>In the example of this article, the carryover effect is evaluated through sequence effect analysis.

## Sequence effect

The fact that subjects are allocated to a particular sequence may affect the results. That is, when comparing the means of the dependent variables in the AB and BA sequences, there should be no difference if there is no sequence effect. This allows the assumption that there is no sequence effect by randomization in the AB/BA sequence. However, it should be noted that this assumption cannot be verified through statistical analysis [6].

## Randomization in crossover design

Randomization is performed to eliminate selection bias and to provide statistical evidence for quantitative evaluation. In parallel design, randomization to different treatment groups (A or B) can ensure independence between groups. However, in the crossover design, randomization is performed in sequence, that is, AB or BA sequence, so that virtually no independence between treatment groups is guaranteed. Therefore, as discussed above, it is necessary to test whether the treatment effect shown in the first period remains in the second period. The randomization method may be reviewed in the article of Lim and In [7].

## Other considerations

When planning the crossover design, to confirm the treatment effect, the carryover effect should not appear above all. The carryover effect can be found through data analysis, and when this effect is significant, it is difficult to interpret the treatment effect. Therefore, a protocol such as the above-mentioned washout period setting is an appropriate method to eliminate this effect. However, if the washout period is too long, the dropout rate may increase at the second period. In some cases, crossover studies may be impossible. For example, in the case of an acutely curable disease, if it is cured in the first period, there is no reason for the subjects to participate in the second period. Therefore, the crossover design is more suitable for chronic diseases than for acute diseases. These factors should be reflected in crossover design planning.

## Statistical model and SAS code

### Example

The example used below was published by Senn and Auclair [8]. Thirteen pediatric patients aged 7–14 years were treated sequentially with two distinct bronchodilators, the newer drug formoter-

ol (For, 12 µg) and the control drug salbutamol (Sal, 200 µg). Seven patients were randomized to the For-Sal sequence and six were randomized to the Sal-For sequence. An established wash-out period of at least 1 d was allowed between administration of the two drugs. Peak expiratory flow (pef, L/min) was measured 8 h after inhalation of each bronchodilator. The SAS code used in the following example was run using SAS 9.4 (SAS Institute Inc., USA) and SAS University Edition installed on a Microsoft Windows 10 (64-bit) operating system.

## Data input into SAS

The full data entry is presented in Fig. 1 of the Appendix. The SAS code for printing up to the first five lines (OBS = 5) of all 26 lines of data and the results are as follows:

```
title "PEF data";
PROC PRINT DATA = PEF (OBS = 5);
RUN;
```

PEF data

OBS	sequence	subject	period	treat	pef
1	For-Sal	1	1	For	310
2	For-Sal	1	2	Sal	270
3	For-Sal	4	1	For	310
4	For-Sal	4	2	Sal	260
5	For-Sal	6	1	For	370

## Analysis of variance table to test other effects

Determining whether various effects other than the treatment effect appearing in this data can be easily accomplished through

the analysis of variance table. The following code is used for the construction and execution of a general linear model to which analysis of variance is applied. Here, **sequence** represents For→Sal or Sal→For, the **subject** is a pediatric patient, and **period** represents the first period (1) or the second period (2). **treat** means cure using For or Sal. **CLASS** defines the variables to be used in the model, and **MODEL** is a code that specifies the independent and dependent variables.

```
PROC GLM DATA = PEF;
  CLASS sequence subject period treat;
  MODEL pef = sequence subject (sequence) period treat;
RUN;
```

Fig. 2 shows the execution result of the above SAS code, and Table 1 summarizes this result in an analysis of variance table. As mentioned above, it is not possible to separately confirm all effects or interactions because of the limitations of the simplified cross-over design; hence, the carryover effect is inherent in other effects or interactions. In this crossover design model, the carryover effect is inherent in the **sequence**. Therefore, the result for the sequence is shown as carryover in Table 1.

In the table of the sum of squares in Fig. 2, the total sum of squares is divided into the between-subject sum of squares and the within-subject sum of squares. The between-subject sum of squares is again divided into the sum of squares for the carryover effect and the sum of squares for the residual. The within-subject sum of squares is divided into the sum of squares for the treatment effect, the sum of squares for the period effect, and the sum of squares for the residual. Neither the carryover effect ( $[F(1,11) = 4.84 > 0.45]$ ,  $P = 0.5177$ ) nor the period effect ( $[F(1,11) = 4.84 > 2.17]$ ,  $P = 0.1683$ ) is statistically significant.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	14	130233.9973	9302.4284	12.40	< .0001
Error	11	8254.4643	750.4058		
Corrected Total	25	138488.4615			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sequence	1	335.1877	335.1877	0.45	0.5177
subject (sequence)	11	114878.2738	10443.4794	13.92	< .0001
period	1	1632.0742	1632.0742	2.17	0.1683
treat	1	14035.9203	14035.9203	18.70	0.0012

Fig. 2. Part of the output of SAS PROC GLM (General linear model procedure). DF: degrees of freedom, SS: sum of squares.

**Table 1.** Analysis of Variance Table for a Standard 2 × 2 Crossover Design (Example Data)

Source of variation	d.f.	Type III sum of squares	Mean square	F	P value
Between subjects					
Carryover	1	335	335	0.45	0.518
Residual	11	114878	10443		
Within-subjects					
Treatment	1	14036	14036	18.70	0.001
Period	1	1632	1632	2.17	0.168
Residual	11	8254	750		
Total	25	138488			

Due to the limitation of the simplified 2 × 2 crossover design, all effects or interactions cannot be estimated separately. Therefore, depending on the model to be designed, the carryover effect is inherent in other effects or interactions. Here, the carryover effect is confirmed through the sequence. That is, in this crossover design model, the carryover effect is inherent in the sequence effect. Therefore, the result for the sequence is displayed as Carryover. d.f.: degrees of freedom.

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
sequence	1	11	0.03	0.8611
period	1	11	2.17	0.1683
treat	1	11	18.70	0.0012

Estimates								
Label	Estimate	Standard Error	DF	t Value	Pr >  t	Alpha	Lower	Upper
Treatment	46.6071	10.7766	11	4.32	0.0012	0.05	22.8881	70.3262

Least Squares Means						
Effect	treat	Estimate	Standard Error	DF	t Value	Pr >  t
treat	For	341.49	20.8110	11	16.41	< .0001
treat	Sal	294.88	20.8110	11	14.17	< .0001

Differences of Least Squares Means							
Effect	treat	_treat	Estimate	Standard Error	DF	t Value	Pr >  t
treat	For	Sal	46.6071	10.7766	11	4.32	0.0012

**Fig. 3.** Part of the output of SAS PROC MIXED (Mixed effect model procedure). DF: degrees of freedom, Den DF: DF for the denominator, Num DF: DF for the numerator.

## Analysis of treatment effect

```
PROC MIXED DATA = PEF;
  CLASS sequence subject period treat;
  MODEL pef = sequence period treat;
  RANDOM subject(sequence);
  LSMEANS treat/PDIFF;
  ESTIMATE 'Treatment' treat 1 -1 /CL ALPHA = 0.05;
RUN;
```

The subject is included as a random effect through **RANDOM subject (sequence)**. **LSMEANS treat/PDIFF** is a code that describes the P value for the difference between two treatments,

and **ESTIMATE 'Treatment' treat 1 -1 /CL ALPHA = 0.05** represents the estimate of the treatment effect and 95% confidence interval.

After confirming that other effects do not appear through the analysis of variance table, the treatment effect can be analyzed using a linear mixed effect model (Fig. 3). In **pef**, the average was 341.5 ml after inhalation of **For** and 294.9 ml after inhalation of **Sal**, and the **pef** was as high as 46.6 ml after inhalation of **For** (95% CI: 22.9 – 70.3 ml, P = 0.012).

## Conclusion

The AB/BA design is a simplified form of the AA/AB/BA/BB crossover design introduced in clinical practice. It has many ad-

vantages, such as the ability of subjects themselves to reduce the influence of covariates by acting as a control group, and the higher power and statistical efficiency compared to those of the parallel design. However, there are also several limitations such as the difficulty in studying drugs with long half-lives, the requirement for subject stability even at different periods, administration of all treatments to one subject, and complicated processing of dropped or missing data. Therefore, it is necessary to understand the strengths and weaknesses of this study design and to apply an appropriate analysis method. It should also be noted that if a carry-over effect occurs when using a crossover design, only the results of the first period can be explored.

## Conflicts of Interest

No potential conflict of interest relevant to this article was reported.

## Author Contributions

Chi-Yeon Lim (Conceptualization; Resources; Supervision; Writing – review & editing)

Junyong In (Conceptualization; Software; Writing – original draft)

## ORCID

Chi-Yeon Lim, <https://orcid.org/0000-0003-0178-6976>

Junyong In, <https://orcid.org/0000-0001-7403-4287>

## References

1. Chow SC, Liu JP. Design and Analysis of Bioavailability and Bioequivalence Studies. 3rd ed. Boca Raton, Chapman & Hall/CRC. 2009, pp 31-53.
2. Senn S. Cross-over Trials in Clinical Research. 2nd ed. West Sussex, John Wiley & Sons. 2002, pp 1-16.
3. Jones B, Kenward MG. Design and Analysis of Cross-Over Trials. 2nd ed. Boca Raton, CHAPMAN & HALL/CRC. 2003, pp 13-115.
4. Wellek S, Blettner M. On the proper use of the crossover design in clinical trials: part 18 of a series on evaluation of scientific publications. *Dtsch Arztebl Int* 2012; 109: 276-81.
5. Senn S. Cross-over Trials in Clinical Research. 2nd ed. West Sussex, John Wiley & Sons. 2002, pp 35-88.
6. Ratkowsky D, Evans M, Alldredge J. Cross-over experiments. Boca Raton, CRC Press. 2019, pp 77-120.
7. Lim CY, In J. Randomization in clinical studies. *Korean J Anesthesiol* 2019; 72: 221-32.
8. Senn SJ, Auclair P. The graphical representation of clinical trials with particular reference to measurements over time. *Stat Med* 1990; 9: 1287-302.

## Appendix

**Fig. 1.** Data Input in SAS

```
DATA PEF;
  INPUT sequence $ subject period treat $ pef;
  DATALINES;
For-Sal 1 1 For 310
For-Sal 1 2 Sal 270
For-Sal 4 1 For 310
For-Sal 4 2 Sal 260
For-Sal 6 1 For 370
For-Sal 6 2 Sal 300
For-Sal 7 1 For 410
For-Sal 7 2 Sal 390
For-Sal 10 1 For 250
For-Sal 10 2 Sal 210
For-Sal 11 1 For 380
For-Sal 11 2 Sal 350
For-Sal 14 1 For 330
For-Sal 14 2 Sal 365
Sal-For 2 1 Sal 370
Sal-For 2 2 For 385
Sal-For 3 1 Sal 310
Sal-For 3 2 For 400
Sal-For 5 1 Sal 380
Sal-For 5 2 For 410
Sal-For 9 1 Sal 290
Sal-For 9 2 For 320
Sal-For 12 1 Sal 260
Sal-For 12 2 For 340
Sal-For 13 1 Sal 90
Sal-For 13 2 For 220
;
RUN;
```

Spaces are used to separate free formatted data. Data reported by Senn SJ and Auclair P (Statistics in Medicine 1990; 9: 1287-302) were permitted to reuse from the Wiley Publisher.