Chapter 8
Factorial Designs

Domenic J. Reda

Definitions and Examples of Factorial Designs

At its most basic level, a factorial trial tests the effects of two treatments in one trial in a way that examines the treatments alone or in combination with each other. In this instance, we can design a 2 by 2 factorial trial. Let us say we have two treatments, *A* and *B*, and we would like to evaluate them in a factorial design. We can randomly assign study participants as follows:

Randomization 1: treatment *A* or no treatment *A* (perhaps a placebo version of *A*)
Randomization 2: treatment *B* or no treatment *B* (perhaps a placebo version of B)

Here, each patient will undergo randomization 1 and randomization 2 simultaneously. Table 8.1 shows the formation of four groups by this factorial randomization.

However, the design is more often represented as follows (Table 8.2):

For factor 1, treatment *A* and placebo *A* are identified as the levels of factor 1. Correspondingly, treatment *B* and placebo *B* are the levels of factor 2.

Within a factor, the levels are not constrained to have one be a placebo (or perhaps no treatment). They could be differently intensities or different types of treatment within a similar class. Tables 8.3 and 8.4 show two other possible 2 by 2 designs as examples.

As mentioned earlier, a 2 by 2 factorial design is the most basic of these designs. The number of factors can increase, and the number of levels within a factor can increase. For the example in Table 8.4, if aspirin alone was incorporated into the trial as another choice for post-stent antithrombosis therapy, then the trial would have a 2 by 3 factorial design. If it were desired also to assess the effect of two types

D.J. Reda (✉)
Department of Veterans Affairs, Cooperative Studies Program Coordinating Center (151K), Hines VA Hospital, Building 1, Room B240, Hines, IL 60141, USA
e-mail: Domenic.Reda@va.gov

K.M.F. Itani and D.J. Reda (eds.), *Clinical Trials Design in Operative and Non Operative Invasive Procedures*, DOI 10.1007/978-3-319-53877-8_8

Table 8.1 Four groups formed by a factorial randomization

Group 1	Group 2	Group 3	Group 4
Treatment *A* + treatment *B*	Placebo *A* + treatment *B*	Treatment *A* + placebo *B*	Placebo *A* + placebo *B*

Table 8.2 Representation of a 2 by 2 factorial design

Factor 1	Factor 2	
	Treatment *B*	**Placebo *B***
Treatment *A*	Treatment *A* + treatment *B*	Treatment *A* + placebo *B*
Placebo *A*	Placebo *A* + treatment *B*	Placebo *A* + placebo *B*

Table 8.3 Representation of a 2 by 2 factorial design where levels of a factor are different intensities of a treatment

Factor 1	Factor 2	
	Low-dose *B*	**High-dose *B***
Low-dose *A*	Low-dose *A* + low-dose *B*	Low-dose *A* + high-dose *B*
High-dose *A*	High-dose *A* + low-dose *B*	High-dose *A* + high-dose *B*

Table 8.4 Representation of a 2 by 2 factorial design where levels of a factor are different types of treatment within a class

Stent type	Post-stent antithrombosis therapy	
	Clopidogrel	**DAPT**
BMS	BMS + clopidogrel	BMS + DAPT
DES	DES + clopidogrel	DES + DAPT

Note: *BMS* bare metal stent, *DES* drug-eluting stent, *DAPT* dual antiplatelet therapy (clopidogrel + aspirin)

Table 8.5 Representation of a 2 by 3 by 2 factorial design

Stent type	Rosuvastatin		
	Post-stent antithrombosis therapy		
	Clopidogrel	**Aspirin**	**DAPT**
BMS	BMS + clopidogrel	BMS + aspirin	BMS + DAPT
DES	DES + clopidogrel	DES + aspirin	DES + DAPT
Stent type	**Atorvastatin**		
	Post-stent antithrombosis therapy		
	Clopidogrel	**Aspirin**	**DAPT**
BMS	BMS + clopidogrel	BMS + aspirin	BMS + DAPT
DES	DES + clopidogrel	DES + aspirin	DES + DAPT

Note: *BMS* bare metal stent, *DES* drug-eluting stent, *DAPT* dual antiplatelet therapy (clopidogrel + aspirin)

of statin drug as part of the trial, then the trial would have a 2 by 3 by 2 factorial design (Table 8.5).

Advantages and Disadvantages of a Factorial Trial

The main advantage of a factorial design is that it provides the ability to answer multiple questions about treatment effects within one trial. For example, the following questions could be asked in a factorial design as shown in Table 8.4 where post-stent thrombosis is the primary outcome measure.

Is there a difference in post-stent thrombosis:

1. between bare metal and drug-eluting stents?
2. between clopidogrel given alone or with dual antiplatelet therapy (DAPT)?
3. between bare metal stents and drug-eluting stents for those given clopidogrel alone?
4. between bare metal stents and drug-eluting stents for those given DAPT?
5. between clopidogrel and DAPT for those given bare metal stents?
6. between clopidogrel and DAPT for those given drug-eluting stents?

This does not exhaust all the possible questions that can be asked regarding post-stent thrombosis with this design. The reader may wish to examine the structure of the trial and identify other possible questions that could be answered. Thus, by overlaying two different treatment modalities in one trial, a factorial design provides the opportunity to examine treatment effects of a therapy alone and in combination with other therapies.

Another advantage is that in placebo-controlled factorial designs, recruitment into the trial may be easier. Referring to the design shown in Table 8.2, let us assume that a factorial design was not considered and the investigator decided to conduct two different trials. The first would compare treatment A to placebo, and the second would compare treatment *B* to placebo. In each of these trials, a participant would have a 50% chance of receiving placebo. However, in the trial represented in Table 8.2, the participant would have a 25% chance of receiving (double) placebo.

There are three main disadvantages to a factorial design. The first is that the complexity of the trial is considerably increased, which can impede the investigator's ability to successfully complete the trial.

The second disadvantage is that the interpretation of the results can be more complicated if it turns out that a treatment level in factor *A* modifies the effect of a treatment level in factor B.

The third is that the sample size requirement will increase considerably if it is possible that the effect of one factor will modify the effect of the other factor. For example, in the list of questions above, questions 3–4 imply that the difference between bare metal and drug-eluting stents may depend on whether clopidogrel is

given alone or in combination with aspirin, while questions 5–6 imply that the difference between clopidogrel alone and clopidogrel combined with aspirin depends on the type of stent. All four of these questions reflect subgroups in the overall study. Only questions 1 and 2 are based on the entire sample size. We go a bit deeper into this issue in the next several sections of this chapter.

Main Effects

Let us return to the example in Table 8.3. If the effect of low-dose treatment *A* is the same regardless of the dose for treatment *B* and the effect of high-dose treatment *A* is the same regardless of the dose for treatment B, then we can draw a conclusion about whether the effect of low-dose treatment *A* differs from the effect of high-dose treatment *A* without needing to consider what dose of treatment *B* the participants received. It turns out we can also ask whether low-dose treatment *B* differs from high-dose treatment *B* without needing to consider the dose of treatment *A*. We refer to these as the factor 1 main effect (low-dose *A* and high-dose *A* differ) and the factor 2 main effect (low-dose *B* and high-dose *B* differ). Main effects are also described as marginal effects. The factor 1 effect can be assessed by ignoring (or combining) the low-dose *B* and high-dose *B* columns. Similarly, the factor 2 effect can be assessed by ignoring (or combining) the low-dose *A* and high-dose *A* (Table 8.6).

Interactions

When the factor 1 effect depends on the level of factor 2, or vice versa, then there is effect modification between the two factors, also known as an interaction effect. When interaction is present, we can no longer make a statement about the main effect of factor 1 or 2. Rather, in order to determine whether the effect of low-dose *A* differs

Table 8.6 Representation of a 2 by 2 factorial design and the main effect for the two factors

Factor 1	**Factor 2**		
	Low-dose *B*	**High-dose *B***	**Main effect of factor 1**
Low-dose *A*	Low-dose *A* + low-dose *B*	Low-dose *A* + high-dose *B*	Low-dose *A*
High-dose *A*	High-dose *A* + low-dose *B*	High-dose *A* + high-dose *B*	High-dose *A*
Main effect of factor 2	Low-dose *B*	High-dose *B*	

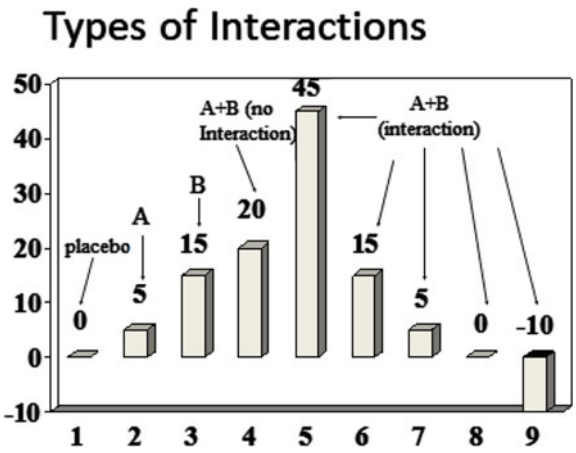

Fig. 8.1 Figure types of interactions. Column 1—there is no placebo effect. Column 2—the effect of *A* is 5. Column 3—the effect of *B* is 15. Column 4—the combined effect of *A* and *B* is additive and =20 (no interaction). Column 5—the combined effect of *A* and *B* is larger than either alone (interaction). Column 6—the combined effect of *A* and *B* is the same as *B* alone, i.e. there is no added benefit to *A* in the presence of *B* (interaction). Column 7—the combined effect of *A* and *B* is the same as *AB* alone, i.e. there is no added benefit to *B* in the presence of *A* (interaction). Column 8—the combined effect of *A* and *B* is the same as placebo, i.e. *A* and *B* each cancel the effect of the other (interaction). Column 9—the combined effect of *A* and *B* is worse than either alone and worse than placebo (interaction)

from that of high-dose *A*, we need to consider the dose level of *B*. Essentially, an interaction effect means that the comparison of low- and high-dose *A* for those who received low-dose *B* is different from the comparison of low- and high-dose *A* for those who received high-dose *B*. The reverse also holds when interaction is present. In other words, the margins of the table are ignored and the appropriate comparisons are within the 2 by 2 table, i.e. the four cells.

Another way of considering interaction is the following. When no interaction is present, the effects of factor 1 and factor 2 are additive. Any departure from an additive relationship between factor 1 and factor 2 is an interaction. Interaction effects can take many forms. Figure 8.1 shows several scenarios indicating the presence of interaction.

Statistical Analysis

When the outcome measure is continuous, analysis of variance is used to examine main and interaction effects. The statistical models for a 2 by 2 factorial design with a continuous outcome measure are:

Two-way ANOVA with interaction →

$$x_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk};$$
$$i = 1, 2; \quad j = 1, 2; \quad k = 1, 2, \ldots, n$$

$\alpha_i \rightarrow$ effect due to factor A

$\beta_j \rightarrow$ effect due to factor B

$(\alpha\beta)_{ij} \rightarrow$ interaction of factors A and B

$\varepsilon_{ijk} \rightarrow$ error term

Two-way ANOVA without interaction →

$$x_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk};$$
$$i = 1, 2; \quad j = 1, 2; \quad k = 1, 2, \ldots, n$$

$\alpha_i \rightarrow$ effect due to factor A

$\beta_j \rightarrow$ effect due to factor B

$\varepsilon_{ijk} \rightarrow$ error term

The analytic steps are:

- Perform two-way ANOVA with interaction
- If the interaction is not statistically significant
  - Perform two-way ANOVA again but with no interaction term
  - Examine the significance level of factor A and factor B
- If interaction is statistically significant
  - Transform the problem into a one-way ANOVA with each group defined by the levels of factors A and B
  - For example, a 2 × 4 factorial becomes a one-way ANOVA with 8 groups
  - Then proceed as if you have an 8 parallel group trial. Use a multiple comparison procedure to determine which pairs of groups are different

If the outcome measure is binary, then a similar approach can be followed using logistic regression. If the outcome is a time-to-event outcome (survival analysis), then Cox regression can be used. Note that for binary and time-to-event outcomes, the multiple comparisons step will use a more sophisticated process to perform multiple comparisons.

Effect of Interaction on Sample Size

During the design of a factorial trial, the investigator will need to determine whether it can be assumed there is no interaction. Here, the nature of the medical condition

being considered, and the known effects of the treatments may help inform that decision. However, absent good evidence that interaction is unlikely, the study should be designed assuming interaction is possible. Unfortunately, this will have the effect of increasing the sample size for the study, often substantially. A fourfold increase in sample size would not be unusual.

Let us assume we have designed a 2 by 2 factorial trial that will have 1000 participants in total, equally divided among the four groups. If no interaction can be assumed, then the comparison of the two levels of factor *A* (the factor *A* main effect) will involve 500 participants assigned to one level of factor *A* and 500 assigned to the other level. However, if interaction is present, then the analysis will involve a comparison of the four groups constructed by the 2 by 2 design and each pairwise comparison of any of the four groups will involve 500 total (250 per group), or half the sample size available for a test of a main effect. However, the relationship between power and sample size is quadratic rather than linear.

BARI 2D—A 2 by 2 Trial

The Bypass Angioplasty Revascularization Investigation 2 Diabetes (BARI 2D) used a 2 by 2 factorial design in a sample of participants with type 2 diabetes mellitus and angiographically documented stable coronary artery disease. BARI 2D compared revascularization combined with aggressive medical treatment versus aggressive medical treatment alone, and simultaneously, two glycemic control strategies, insulin sensitization versus insulin provision [1, 2].

The trial group randomly assigned 2368 patients with both type 2 diabetes and heart disease to undergo either prompt revascularization with intensive medical therapy or intensive medical therapy alone and to undergo either insulin-sensitization or insulin-provision therapy. There were two primary end points: the rate of death and a composite of death, myocardial infarction, or stroke (major cardiovascular events). Randomization was stratified according to the choice of percutaneous coronary intervention (PCI) or coronary artery bypass grafting (CABG) as the more appropriate intervention.

The protocol included as part of its statistical analysis plan to test for interaction effects for the two factors. The statistical interactions between the cardiac study groups and the glycemic study groups for rates of death and major cardiovascular events were tested overall and within the PCI and CABG strata at a two-sided alpha level of 0.05. The interactions were found not to be statistically significant, which allowed the group to compare the revascularization and medical-therapy groups (regardless of the diabetes treatment) and vice versa.

At 5 years, rates of survival did not differ significantly between the revascularization group (88.3%) and the medical-therapy group (87.8%, $P = 0.97$) or between the insulin-sensitization group (88.2%) and the insulin-provision group (87.9%, $P = 0.89$). The rates of freedom from major cardiovascular events also did not differ significantly among the groups: 77.2% in the revascularization group and

75.9% in the medical-treatment group ($P = 0.70$) and 77.7% in the insulin-sensitization group and 75.4% in the insulin-provision group ($P = 0.13$).

The Physician's Health Study

Another use of a factorial design when it can be assumed there is no interaction, is to create an efficient trial which is one trial overlayed on another trial, each asking different questions. The Physician's Health Study used this approach [3–5].

A total of 22,071 physicians were randomly assigned, according to a 2 by 2 factorial design, to one of four treatment groups: aspirin and beta carotene, aspirin and placebo beta carotene, placebo aspirin and beta carotene, or placebo aspirin and placebo beta carotene. There were two primary outcome measures: The study was designed to test two primary-prevention hypotheses in a population of healthy male physicians: whether aspirin in low doses (325 mg every other day) reduces mortality from cardiovascular disease, and whether beta carotene (50 mg on alternate days) decreases the incidence of cancer.

The trial design assumed no interaction between low-dose aspirin and beta carotene for either outcome measure. The study design is depicted in Fig. 8.2.

The design was somewhat controversial at the time because it made an assumption of no interaction between low-dose aspirin and beta carotene. While it appeared to be acceptable to make that assumption for the cardiovascular outcome, it was less clear that it was safe to assume no interaction for the cancer outcome.

The trial's Data and Safety Monitoring Board stopped the aspirin component of the trial early when it became clear that aspirin had a significant effect on the risk of a first myocardial infarction. At that time, there were too few strokes or deaths upon which to draw conclusions about the effect of aspirin on stroke or cardiovascular

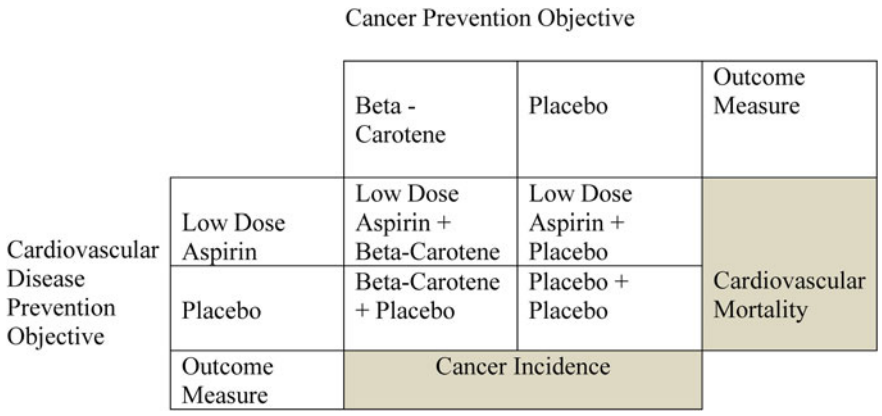

Fig. 8.2 Figure design of the Physician's Health Study

mortality, but the DSMB felt the study should not continue to provide more definitive information about those endpoints since the benefit for myocardial infarction was now established. The beta carotene component continued to completion. It was concluded that 13 years of supplementation with beta carotene produced neither benefit nor harm regarding cancer incidence.

Conclusion

Factorial designs can be very useful in large-scale trials to assess the effects of multiple treatments and how they influence each other.

References

1. Brooks MM, Frye RL, Genuth S, Detre KM, Nesto R, Sobel BE, Kelsey SF, Orchard TJ. Bypass angioplasty revascularization investigation 2 diabetes (BARI 2D) trial investigators. Hypotheses, design, and methods for the bypass angioplasty revascularization investigation 2 diabetes (BARI 2D) trial. Am J Cardiol. 2006;97(12A):9G–19G.
2. BARI 2D Study Group, Frye RL, August P, Brooks MM, Hardison RM, Kelsey SF, MacGregor JM, Orchard TJ, Chaitman BR, Genuth SM, Goldberg SH, Hlatky MA, Jones TL, Molitch ME, Nesto RW, Sako EY, Sobel BE. A randomized trial of therapies for type 2 diabetes and coronary artery disease. N Engl J Med. 2009;360(24):2503–15. doi:10.1056/NEJMoa0805796.
3. Hennekens CH, Buring JE. Methodologic considerations in the design and conduct of randomized trials: the U.S. physicians' health study. Controlled Clin Trials. 1989;10:142S–50S.
4. The Steering Committee of the Physicians' Health Study Research Group. Belanger C, Buring JE, Cook N, Eberlein K, Goldhaber SZ, Gordon D, Hennekens CH, Mayrent SL, Peto R, Rosner B, Stampfer M, Stubblefield F, Willett W. Final report on the aspirin component of the ongoing physicians' health study. N Engl J Med 1989;321:129–35.
5. Hennekens CH, Buring JE, Manson JE, Stampfer M, Rosner B, Cook NR, Belanger C, LaMotte F, Gaziano JM, Ridker PM, Willett W, Peto R. Lack of effect of long-term supplementation with beta-carotene on the incidence of malignant neoplasms and cardiovascular disease. N Eng J Med. 1996;334:1145–9.