9. Crossover Trials

9.1 Motivation for crossover trials

When we first consider statistical methods for clinical trial the notion of potential outcome was introduced (See section 2.1) A patient had two *potential outcomes*, say $Y_i(T)$ and $Y_i(C)$. For the i^{th} patient the treatment effect was defined as $\tau_i = Y_i(T) - Y_i(C)$. The expected treatment effect is therefore, $\tau = E[\tau_i] = E[Y_i(T) - Y_i(C)]$. For many conditions a single course of treatment may cure some if not all patients, so they would no longer be eligible for the comparator treatments. This applies to most treatments for acute (short-term) conditions such as antibiotic treatment for an infection, trauma, or surgery. For these conditions, it is only possible to measure one of the two potential outcomes in the same patient - one outcome is said to be counter factual. As previously discussed, the expected treatment effect is therefore estimated as the difference in average outcome between patients receiving the intervention (T) and those receiving the control treatment (C). The treatment effect is estimated between groups and so the precision of the estimate of the average treatment effects depends on in the variance of the outcome $Var[Y_i]$.

In some for chronic (long-term) diseases such as arthritis, asthma, diabetes, or high blood pressure, the condition is not cured by treatment. Instead symptoms may be reduced or the disease progression slowed by continuing treatment. It may be possible to measure both potential outcomes in the same patient, and so the

treatment effect τ_i can be estimated in each patient. The average treatment effect could then be estimated with greater precision for a given sample size of patients. Trial size would then be reduced making clinical trials easier to conduct.

Suppose one is able to compare two treatments, say A and B, in the same patient. One option would be to give treatment A followed by treatment B. Such a design is potentially biased, because the patient's condition may deteriorate or improve over time irrespective of treatment. To overcome this, patients should be randomly allocated to receive treatments in either order. This trial design is the called the AB-BA Crossover Design. In such a design:

- The trial is divided into two *periods* with one treatment given during each.
- Patients are randomly allocated to two groups, one receiving treatment sequence *A then B* and the other receiving *B then A*.

Random allocation is important for two reasons:

- To prevent bias due to possible change in the patient over time.
- To maintain concealment prior to treatment allocation and post treatment allocation where the trial is double blind.

Use of a Crossover design

A crossover trial designs may be suitable where:

- The condition being treated is a chronic disease (e.g. chronic diabetes).
- The condition is stable and so unlikely to change greatly from one period to the next.
- The intervention has a rapid effect.

A crossover designs are not suitable where:

- The conditions may resolve quickly, that is an acute condition, making the second period treatment unnecessary. (e.g. infections, trauma, and rehabilitation).
- Patients are likely to withdraw from treatment or be lost to follow-up.
- The effect of the first treatment could plausibly contaminate the effect of second. Note for some treatments contamination can be prevented by having a "wash-out" period between treatments.

9.2 Analysis of an *AB-BA* Crossover Design

Suppose patients are randomly allocated to either treatment *A*, followed by treatment B (Sequence *AB*) or treatment *B* then treatment *A* (Sequence *BA*). Suppose that there are n_{AB} and n_{BA} patients in each sequence with total sample size $N = n_{AB} + n_{BA}$. Even though we have stated that patients should be stable, a statistical model for a crossover design needs to include parameters called a period effect, which is the difference between period 2 and period 1 irrespective of treatment order as patients health will change over time. Two sources of variation can be identified, variation *between-patients*, and variation *within-patient*.

If Y_{ij} is the response for the i^{th} subject during period $(j = 1,2)$, a model for an *AB-BA* crossover trial can be define as follows:

$y_{ij} = \mu + \xi_i + \varepsilon_{ij}$ Sequence *AB* Period 1

$y_{ij} = \mu + \tau + \phi + \xi_i + \varepsilon_{ij}$ Sequence *AB* Period 2

$y_{ij} = \mu + \tau + \xi_i + \varepsilon_{ij}$ Sequence *BA* Period 1

$y_{ij} = \mu + \phi + \xi_i + \varepsilon_{ij}$ Sequence *BA* Period 2

μ = mean in period 1 for the Sequence *AB*.

τ = treatment effects of *B* compared to *A*.

ϕ = period effect.

ξ_i = random variable for patient i with $E[\xi_i] = 0$ and variance σ_B^2.

ε_{ij} = random variable for patient i in period j with $E[\varepsilon_{ij}] = 0$ and variance σ_ε^2 assumed to be normally distributed, $N[0, \sigma_\varepsilon^2]$.

σ_B^2 is called the between-patient variance and σ_ε^2 is called the within-patient variance.

Using a Single Sample t-Test to Analyse Crossover Trials

Sometime crossover trials are analysed by using a single sample t-test applied to the difference between the two treatments, also called the paired t-test.

Define $c_i = y_{i2} - y_{i1}$ for $i \in AB$ and $c_i = y_{i1} - y_{i2}$ for $i \in BA$, which are the difference in outcome between when a patient receives treatment B and treatment A. The treatment effect is then estimated by $\hat{\tau}_C = \bar{c} = \frac{\sum c_i}{N}$ with $N = n_{AB} + n_{BA}$. The hypothesis $H_0 : \tau = 0$ is then tested using $T_C = \frac{\bar{c}}{SE[\bar{c}]}$, where $SE[\bar{c}] = \sqrt{\frac{s_C^2}{N}}$ with s_C^2 being the sample standard deviation of the differences, c_i. If c_i can be assumed to be normally distributed, the test statistic T_C has a t-distribution with $\nu = N - 1$ degrees of freedom.

Unfortunately, $\hat{\tau}_C$ can give a biased estimate of τ.

Suppose the data generation model for an AB-BA crossover trial defined above applies. For the treatment effect estimator $\hat{\tau}_C = \bar{c}$

$E[\hat{\tau}_C] = \tau + \frac{(n_{AB} - n_{BA})\phi}{N}$ where $N = n_{AB} + n_{BA}$. If the period effect $\phi \neq 0$ and $n_{AB} \neq n_{BA}$, $\hat{\tau}_C$ will be a biased estimator of the treatment effect, τ.

Substitution from the model above gives

$$E[c_i | i \in AB] = E[(\mu + \tau + \phi + \xi_i + \varepsilon_{ij}) - (\mu + \xi_i + \varepsilon_{ij})] .$$

$$= E[(\tau + \phi + \varepsilon_{ij}) + (\varepsilon_{ij})].$$

Since $E[\varepsilon_{ij}] = 0$, $E[c_i | i \in AB] = \tau + \phi$.

Similarly, $E[c_i | i \in BA] = E[(\tau + \varepsilon_{i1}) - (\phi + \varepsilon_{i2})] = \tau - \phi$.

Now $E[\hat{\tau}_C] = E[\bar{c}] = E\left[\frac{\sum c_i}{N}\right] = \frac{\sum E[c_i]}{N}$.

Hence,

$$E[\hat{\tau}_C] = \frac{n_{AB}(\tau + \phi) + n_{BA}(\tau - \phi)}{N} = \tau + \frac{(n_{AB} - n_{BA})\phi}{N} \bullet$$

It is rarely possible to rule-out a period effect (ϕ) completely. Even where n_{AB} and n_{AB} are planned to be equal, the may differ due to imbalance arising from randomisation, such as incomplete blocks in block randomisation, or patients dropout from the trial. The treatment effect estimate $\hat{\tau}_C$ may therefore be biased and is therefore not recommended.

An Unbiased Estimator of the Treatment Effect

Notation

For subject i, define $d_i = y_{i2} - y_{i1}$, that is the difference between period 2 and period 1 irrespective of treatment order. Let

$\bar{d}_{AB} = \frac{\sum_{i \in AB} d_i}{n_{AB}}$ and $\bar{d}_{BA} = \frac{\sum_{i \in BA} d_i}{n_{BA}}$ be the sample means for sequences AB and BA respectively.

Suppose the data generating model for an AB-BA crossover trial defined above applies, then $\hat{\tau} = \frac{\bar{d}_{AB} - \bar{d}_{BA}}{2}$ is an unbiased estimator of τ, that is $E[\hat{\tau}] = E\left[\frac{\bar{d}_{AB} - \bar{d}_{BA}}{2}\right] = \tau$.

Note τ is the treatment effect of B compared to A

Substitution from the model above gives

$$E[d_i | i \in AB] = E[(\mu + \tau + \phi + \xi_i + \varepsilon_{ij}) - (\mu + \xi_i + \varepsilon_{ij})]$$

The terms ξ_i on the RHS cancel so

$$E[d_i | i \in AB] = E[(\tau + \phi + \varepsilon_{ij}) + (\varepsilon_{ij})].$$

Since $E[\varepsilon_{ij}] = 0$, $E[d_i | i \in AB] = \tau + \phi$.

Hence $E[\bar{d}_{AB}] = E\left[\frac{\sum_{i \in AB} d_i}{n_{AB}}\right] = \frac{n_{AB} E[d_i | i \in AB]}{n_{AB}} = \phi + \tau$.

Similarly $E[d_i | i \in BA] = E[(\phi + \varepsilon_{ij}) - (\tau + \varepsilon_{ij})] = \phi - \tau$.

and $E\left[\bar{d}_{BA}\right] = E\left[\frac{\sum_{i\in BA} d_i}{n_{BA}}\right] = \phi - \tau$.

Hence $E\left[\frac{\bar{d}_{AB} - \bar{d}_{BA}}{2}\right] = \frac{E\left[\bar{d}_{AB}\right] - E\left[\bar{d}_{BA}\right]}{2} = \frac{(\tau+\phi)-(\phi-\tau)}{2} = \tau \bullet$

Being the difference of two sample means of the two sequences, statistical inference on the estimator $\hat{\tau} = \frac{\bar{d}_{AB} - \bar{d}_{BA}}{2}$ can be based on a two sample t-test provided the assumptions of the test are satisfied. Since the patients in sequences AB and BA are independent, the assumption of independence follows from the design. The assumption of normality follows from $\epsilon_{ij} \sim N[0, \sigma_\epsilon^2]$.

The Variance of the Difference d_i

Define $Var\left[d_i \middle| i \in AB\right] = \sigma^2_{d_{AB}}$ and $Var\left[d_i \middle| i \in BA\right] = \sigma^2_{d_{BA}}$.

From the model for a cross-over trial (in 9.1), for sequence AB

$$\sigma^2_{d_{AB}} = Var\left[d_i\right] = Var\left[y_{2i} - y_{1i}\right] = Var\left[(\mu+\phi+\xi_i+\varepsilon_{i2}) - (\mu+\tau+\xi_i+\varepsilon_{i1})\right]$$

$$= Var\left[\varepsilon_{i2} - \varepsilon_{i1}\right] = Var\left[\varepsilon_{i1}\right] + Var\left[\varepsilon_{i2}\right] = 2\sigma^2_\varepsilon \text{ , since } \operatorname{cov}\left[\varepsilon_{i1}, \varepsilon_{i2}\right] = 0.$$

Similarly, for sequence BA, $\sigma^2_{d_{BA}} = 2\sigma^2_\varepsilon$.

Hence, the variances of the two sequences are the same for this model, that is $\sigma^2_{d_{AB}} = \sigma^2_{d_{BA}} = \sigma^2_d = 2\sigma^2_\varepsilon$.

Analysis of an *AB-BA* Crossover Trial using a Two Sample t-test

Hypothesis Testing

The hypothesis $H_0 : \tau = 0$ vs $H_1 : \tau \neq 0$ can be tested using a two-sample t-test of the means of the differences. The test statistic *T* is defined as

$T = \frac{\bar{d}_{AB} - \bar{d}_{BA}}{\hat{SE}\left[\bar{d}_{AB} - \bar{d}_{BA}\right]}$ where $\hat{SE}\left[\bar{d}_{AB} - \bar{d}_{BA}\right] = s_d\sqrt{\frac{1}{n_{AB}} + \frac{1}{n_{BA}}}$ and

$s_d = \sqrt{\frac{(n_{AB}-1)s^2_{d_{AB}} + (n_{BA}-1)s^2_{d_{BA}}}{n_{AB}+n_{BA}-2}}$ with $s^2_{d_{AB}}, s^2_{d_{BA}}$ be the sample variances of the differences for the two sequences. Under assumptions of normality and equality of variance $\left(\sigma^2_{d_{AB}} = \sigma^2_{d_{BA}} = \sigma^2_d\right)$, the test statistic *T* will have a t-distribution with $n_{AB}+n_{BA}-2$ degrees of freedom.

Confidence Interval of the Treatment Effect

A $(1-\alpha)$ size confidence interval for the treatment effect τ is defined by

$$\frac{1}{2}\left(\bar{d}_{AB} - \bar{d}_{BA}\right) \pm \frac{1}{2} t_{\alpha/2}\left(n_{AB}+n_{BA}-2\right)\hat{SE}\left[\bar{d}_{AB} - \bar{d}_{BA}\right].$$

Example Bronchodilators Crossover Trial. The data in the table below is from a two-period *AB-BA* randomised crossover trial in which patients are treated with two bronchodilators, salbutamol (S) and formoterol (F). The outcome is the peak expiratory flow (PEF). Patients were randomised to receive F then S or S then F. Increased PEF is a benefit to patients.

Patient	*PEF period 1* y_{c1}	*PEF period 2* y_{c2}	$d_c = y_{c2} - y_{c1}$
F then S	Drug F	Drug S	
1	310	270	-40
4	310	260	-50
6	370	300	-70
7	410	390	-20
9	250	210	-40
10	380	350	-30
13	330	365	35
N_{FS}	7	$\bar{d}_{FS}$	-30.7
		$s_{d_{FS}}$	33.0
S then F	Drug S	Drug F	
2	370	385	15
3	310	400	90
5	380	410	30
8	290	320	30
11	260	340	80
12	90	220	130
N_{SF}	6	$\bar{d}_{SF}$	62.5
		$s_{d_{SF}}$	44.7

Ex 9.1 Test the null hypothesis $H_0 : \tau = 0$ vs $H_1 : \tau \neq 0$

$\bar{d}_{FS} - \bar{d}_{SF} = -30.7 - (+62.5) = -93.2$

$s_d = \sqrt{\dfrac{(n_{FS}-1)s^2_{d_{FS}} + (n_{SF}-1)s^2_{d_{SF}}}{n_{FS}+n_{SF}-2}} = \sqrt{\dfrac{6\times 33.0^2 + 5\times 44.7^2}{11}} = 38.76$

$\hat{SE}\left[\bar{d}_{FS} - \bar{d}_{SF}\right] = s_d\sqrt{\dfrac{1}{n_{FS}} + \dfrac{1}{n_{SF}}} = 38.76\times\sqrt{\dfrac{1}{7}+\dfrac{1}{6}} = 21.56$

$T = \dfrac{\bar{d}_{FS} - \bar{d}_{SF}}{\hat{SE}\left[\bar{d}_{FS} - \bar{d}_{SF}\right]} = \dfrac{-93.2}{21.56} = -4.32$

Degrees of freedom $= n_{FS} + n_{SF} - 2 = 11$

For a 5% level test $t_{\alpha/2}(n_{FS}+n_{SF}-2) = t_{0.025}(11) = 2.201$

The null hypothesis is rejected at a 5% level.
Note $t_{0.005}(11) = 3.1058$, so we can also reject the hypothesis at a 1% level

Ex 9.2 Calculate the point estimate and 95% confidence interval of the treatment effect.

The point estimate of the treatment effect of salbutamol (S) compared to formoterol (F) is $\tau = \dfrac{-93.2}{2} = -46.6$

The 95% confidence interval of τ is

$$\frac{1}{2}\left(\bar{d}_{FS} - \bar{d}_{SF}\right) \pm \frac{1}{2} t_{0.025}(n_1+n_2-2)\,\hat{SE}\left[\bar{d}_{FS} - \bar{d}_{SF}\right]$$

$-46.6 \pm \frac{1}{2}\times 2.201 \times 21.56$,
which is
-70.3 to -22.9.

Summary
The treatment effect of salbutamol compared to control is -46.6 (95% C.I. -70.3 to -22.9, $p < 0.01$)

Figure 9.1 Stata Output for Bronchodilators Crossover Trial

```
Summary statistics: mean, sd, N
  by categories of: Sequence

Sequence |      PEF1       PEF2
---------+--------------------
F then S |  337.1429   306.4286
         |  53.76315   64.72469
         |          7          7
---------+--------------------
S then F |  283.3333   345.8333
         |  105.3882   70.88136
         |          6          6
---------+--------------------
```

Two Sample t-test Analysis Applied to the Differences between Period (d_i)

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
F then S |       7   -30.71429    12.46082    32.96824   -61.20482   -.2237503
S then F |       6        62.5      18.246     44.6934    15.59715    109.4028
---------+--------------------------------------------------------------------
    diff |           -93.21429    21.55312               -140.6524   -45.77619
------------------------------------------------------------------------------
    diff = mean(F then S) - mean(S then F)                        t =  -4.3249
Ho: diff = 0                                     degrees of freedom =       11

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0006         Pr(|T| > |t|) = 0.0012          Pr(T > t) = 0.9994
```

The treatment effect and its confidence interval are obtained by halving the values in the printout. From the print-out

`diff` equals `mean(F then S) - mean(S then F)`

The treatment effect for salbutamol compared to formoterol is

$$\hat{\tau} = \frac{-93.21}{2} = -46.6$$

Conclusion There is evidence that formoterol gives improved outcome compared to salbutamol with an increase in PEF equal to 46.6 (95% c.i.22.9 to 70.3 , p=0.0012).

Note that the p-values should NOT be halved!!!

Figure 9.2 Stata Output for the Single Sample t-test Applied to the Differences between treatments (c_i)

```
One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
SALB-FORM|      13   -45.38462    11.25835    40.59257   -69.91446   -20.85477
------------------------------------------------------------------------------
    mean = mean(SALB-FORM)                                        t =  -4.0312
Ho: mean = 0                                     degrees of freedom =       12

    Ha: mean < 0                 Ha: mean != 0                 Ha: mean > 0
 Pr(T < t) = 0.0008         Pr(|T| > |t|) = 0.0017          Pr(T > t) = 0.9992
```

By this method effect for salbutamol compared to formoterol

$$\hat{\tau}_C = -45.4$$

This has a slight bias compared to the unbiased method ($\hat{\tau}$ =-46.6).

It is worth noting also that the standard error for the single sample method is slightly larger. For then unbiased analysis

$$SE[\hat{\tau}] = SE\left[\frac{\bar{d}_{FS} - \bar{d}_{SF}}{2}\right] = \frac{21.55312}{2} = 10.77656$$

whereas $SE[\hat{\tau}_C] = 11.25835$, so the unbiased method is also appears to be more precise.

It can be shown that the period effect increases the standard error of the single sample methods compared to the two-sample method.

9.3 Sample Size for *AB-BA* Cross-over Trials

Given that the analysis of an *AB-BA* crossover trial should be based on a two sample t-test of the differences, we can estimate sample size for a crossover trial using the formula previously derived in the notes for sample size for a parallel group trial for a continuous outcome (see section 4.3). Consider the variance of the difference, σ_d^2.

For an *AB-BA* crossover trial with equal numbers randomised to each sequence, the total sample size *N* required to have power $(1-\beta)$ to detect a treatment effect τ using a two-sided α-size test of the hypothesis of superiority, $H_0:\tau=0$ vs $H_1:\tau\neq 0$ is $N=\frac{\sigma_d^2}{\tau^2}\left(z_{\alpha/2}+z_\beta\right)^2$, where σ_d^2 is the variance of the differences.

If analysis of an AB-BA crossover trial is based on a two sample t-test, the total sample size assuming two equal size groups can be estimated from the formula in section 4.3 for sample size per group.

Total sample size is $N=\frac{4\sigma^2}{\tau^2}\left(z_{\alpha/2}+z_\beta\right)^2$.

If σ^2 is replaced σ_d^2 in the above formula, τ should be replace by 2τ, because $\bar{d}_{AB}-\bar{d}_{BA}$ estimates 2τ.

9.4 Analysis of the Period Effect ϕ

A question of secondary interest is ' Is there a period effect?'

How does one test $H_0:\phi=0$ vs $H_1:\phi\neq 0$?

Consider the differences between treatment *B* and *A* for each subject define by $c_i=y_{i2}-y_{i1}$ for $i\in AB$ and $c_i=y_{i1}-y_{i2}$ for $i\in BA$. These are the difference used in the biased single sample method. It can be shown that $E\left[\frac{\bar{c}_{AB}-\bar{c}_{BA}}{2}\right]=\phi$ where $\bar{c}_{AB}=\frac{\sum_{i\in AB} c_i}{n_{AB}}=\bar{d}_{AB}$ and $\bar{c}_{BA}=\frac{\sum_{i\in BA} c_i}{n_{BA}}=-\bar{d}_{BA}$, suggesting $\hat{\phi}=\frac{\bar{c}_{AB}-\bar{c}_{BA}}{2}$. The details are an exercise. A test of the hypothesis $H_0:\phi=0$ vs $H_1:\phi\neq 0$ can be carried out as a two sample t-test between the two sequences.

Figure 8.3 Stata Output for Bronchodilators Crossover Trial

Two Sample t-test Analysis Applied to the Differences between treatments (c_i)

```
------------------------------------------------------------------------------
   Group |      Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
F then S |        7   -30.71429    12.46082    32.96824   -61.20482   -.2237503
S then F |        6       -62.5      18.246     44.6934   -109.4028   -15.59715
---------+--------------------------------------------------------------------
    diff |            31.78571    21.55312               -15.65238    79.22381
------------------------------------------------------------------------------
    diff = mean(F then S) - mean(S then F)                        t =   1.4748
Ho: diff = 0                                     degrees of freedom =       11

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9158         Pr(|T| > |t|) = 0.1683          Pr(T > t) = 0.0842
```

Ex 9.3 Estimate the period effect from the print-out

The period effect $\phi=\frac{31.78571}{2}=15.89$

9.5 The Carry-Over Effect

Above, we considered how to test for a period effect. This assumed the same change ϕ for each sequence. There is the possibility that the change, after accounting for the treatment effect, differ between the two sequences *AB* and *BA*. This can occur if one treatment has greater persistent into the second period than the other and is called the *carry-over effect*. One can incorporate such a differential effect into the model of an *AB-BA* trial by adding a term γ to represent the difference in the period effect for the two sequences, *BA* and *AB*. For each group and period the models are now:

$y_{ij} = \mu + \xi_i + \varepsilon_{ij}$ Sequence *AB* Period 1

$y_{ij} = \mu + \tau + \phi + \xi_i + \varepsilon_{ij}$ Sequence *AB* Period 2

$y_{ij} = \mu + \tau + \xi_i + \varepsilon_{ij}$ Sequence *BA* Period 1

$y_{ij} = \mu + \phi + \gamma + \xi_i + \varepsilon_{ij}$ Sequence *BA* Period 2

↑ carryover effect

With $d_i = y_{i2} - y_{i1}$, it follows that

$$E\left[\bar{d}_{AB}\right] = E\left[\frac{\sum d_i}{n_{AB}}\right] = \tau + \phi \qquad \text{for } AB \text{ and}$$

$$E\left[\bar{d}_{BA}\right] = E\left[\frac{\sum d_i}{n_{BA}}\right] = \phi - \tau + \gamma \qquad \text{for BA.}$$

Hence $E[\hat{\tau}] = \left[\frac{\bar{d}_{AB} - \bar{d}_{BA}}{2}\right] = \tau + \frac{\gamma}{2}$ instead of τ.

We have shown that the treatment effect will be biased, if there is a carry-over effect.

A Flawed Statistical Analysis involving Carry-Over Effect

To deal with carry-over effect in crossover trials the following analysis procedure has been suggested.

1. Test whether there is a carry-over effect i.e. $H_0 : \gamma = 0$ vs $H_1 : \gamma \neq 0$
2. If $H_0 : \gamma = 0$ is not rejected, the analysis proceeds as described above.
3. If $H_0 : \gamma = 0$ is rejected, the analysis should just be based on the period 1 data reducing the study to a parallel group trial and the period 2 data is then discarded.

The following test of for a carry-over effect has been proposed.

Define $a_i = y_{i2} + y_{i1}$ so that

$a_i = 2\mu + \phi + \tau + 2\xi_i + \varepsilon_{i2} + \varepsilon_{i1}$ $\quad i \in AB$

$a_i = 2\mu + \phi + \tau + \gamma + 2\xi_i + \varepsilon_{i2} + \varepsilon_{i1}$ $\quad i \in BA$

Defining the sample mean for each sequence as $\bar{a}_{AB}$ and $\bar{a}_{BA}$, with corresponding population means μ^a_{AB} and μ^a_{BA}

$E[\bar{a}_{AB}] = 2\mu + \tau + \phi$ and $E[\bar{a}_{BA}] = 2\mu + \tau + \phi + \gamma$

Therefore $E[\bar{a}_{BA} - \bar{a}_{AB}] = \gamma$.

This suggests the hypothesis H_0:γ =0 is equivalent to H_0: $\mu^a_{AB} = \mu^a_{BA}$, which could be tested by $T_a = \frac{\bar{a}_{BA} - \bar{a}_{AB}}{\hat{SE}[\bar{a}_{BA} - \bar{a}_{AB}]}$ using a two sample t-test.

Consider now

$$Var[a_i] = Var[2\xi_i + \varepsilon_{i2} + \varepsilon_{i1}] = 4Var[\xi_i] + Var[\varepsilon_{i2}] + Var[\varepsilon_{i1}] = 4\sigma_B^2 + 2\sigma_\varepsilon^2$$

Therefore $Var[\bar{a}_{AB}] = \dfrac{4\sigma_B^2 + 2\sigma_\varepsilon^2}{n_{AB}}$ and $Var[\bar{a}_{BA}] = \dfrac{4\sigma_B^2 + 2\sigma_\varepsilon^2}{n_{BA}}$.

Hence $Var[\bar{a}_{BA} - \bar{a}_{AB}] = \left(4\sigma_B^2 + 2\sigma_\varepsilon^2\right)\left(\dfrac{1}{n_{BA}} + \dfrac{1}{n_{AB}}\right)$,

giving $SE[\bar{a}_{BA} - \bar{a}_{AB}] = \sqrt{\left(4\sigma_B^2 + 2\sigma_\varepsilon^2\right)\left(\dfrac{1}{n_{BA}} + \dfrac{1}{n_{AB}}\right)}$.

This contains both *within-patient* and *between-patient* variation. We can therefore draw the following conclusion:

> Unless σ_B^2 is small relative to σ_ε^2 the test statistic T_a will have low power to reject the hypothesis of no carry-over effect $H_0 : \gamma = 0$ compared to a test of the treatment effect τ.

The advantage of crossover trials is that they remove *between-patient* variance σ_B^2, and so there is only going to be power to detect a carry-over effect where there is little benefit in using a crossover design. To increase the sample size so that a carry-over effect could be detected would remove the main advantage of the crossover design. Note, also that the inferential reasoning of the test is faulty as in practice one would want to show that γ is zero to justify the crossover analysis. As with an equivalence trial one would need to have $H_0 : \gamma \neq 0$ and $H_1 : \gamma = 0$. For these reasons the procedure for testing for carry-over effect in a crossover trial is no longer recommended.

Implications of Carry-Over Effect and Crossover Trials

When planning a crossover trial, it is important to consider whether a carry-over effect may occur, as estimation is only unbiased where it is absent. This assumption has to be based on scientific arguments regarding the way treatments work rather than statistical tests. For example, in a particular situation carry-over may not be plausible so it can be safely ignored. Alternatively, it may be eliminated by a washout period between the two treatments. In such circumstances a crossover trial may be legitimately conducted and the advantages of crossover trial exploited.

Unfortunately, in many situations a carry-over effect may be plausible and lengthy washouts periods may not be feasible as it would be unethical for a patients to go without treatment. There are therefore many circumstances where a crossover trial is not feasible, when testing treatments for chronic conditions.

This is an example of where clinical science is important for determining the choice of design. In summary, if there are good scientific reasons to believe a carry-over effect may occur, the crossover design is not recommended and a parallel group design should be used instead.

9.6 Summary: Comparison of Parallel and Crossover Trials

Parallel Group Design

- Comparison of treatments is between groups of patients.
- Power and sample size depends on between-subject variation.

Crossover Design

- Comparison of treatments is within patients so that each patient acts as their own control.

Advantage of a Crossover Design compared to a Parallel Group Design

- Within patient estimation of treatment effects means that variation between patients is removed from the analysis, hence sample size may be substantially smaller.

Disadvantages of a Crossover Design compared to a Parallel Group Design

- Only applicable to certain types of condition such as stable diseases.
- More complicated to organize.
- Patients withdrawing during the second period mean that their data cannot be included in the statistical analysis.
- Requires the assumption of no Carry-over effect to give unbiased estimates of the treatment effect.