Designs for Clinical Trials
Chapter 5 Reading Instructions

• 5.1: Introduction
• 5.2: Parallel Group Designs (read)
• 5.3: Cluster Randomized Designs (less important)
• 5.4: Crossover Designs (read+copies)
• 5.5: Titration Designs (read)
• 5.6: Enrichment Designs (less important)
• 5.7: Group Sequential Designs (read include 10.6)
• 5.8: Placebo-Challenging Designs (less important)
• 5.9: Blinded Reader Designs (less important)
• 5.10: Discussion
Design issues
Statistical optimality not enough!
Use simulation models!
Parallel Group Designs

- Easy to implement
- Accepted
- Easy to analyse
- Good for acute conditions

\[ Y_{ij} = \mu_i + \varepsilon_{ij} \]
\[ i = 1, \ldots, n_j \]
\[ j = 1, \ldots, k \]
\[ \varepsilon_{ij} \sim N(0, \sigma^2) \]
Where does the variation go?

\[ Y = X\beta + \epsilon \]

Anything we can explain  Unexplained
Between and within subject variation

Baseline 8 weeks

Placebo
Drug X
Female
Male

DBP mmHg

Baseline 8 weeks
What can be done?

**Stratify:** Randomize by baseline covariate and put the covariate in the model.

**More observations per subject:** Baseline
More than 1 observation per treatment

**Run in period:** Ensure compliance, diagnosis
Parallell group design with baseline

Compare bloodpressure for three treatments, one test and two control.

Model: \[ Y_{ijk} = \mu + 1_{\{i=2\}} \tau_j + \xi_k + \varepsilon_{ijk} \]

- Treatment effect \( \tau_j \)
- Subject effect \( \xi_k \text{ iid } N(0, \sigma_s^2) \)
- Random error \( \varepsilon_{ijk} \text{ iid } N(0, \sigma^2) \)

Observation \( i = 1,2 \)
Treatment \( j = 1,2,3 \)
Subject \( k = 1, \ldots, n_j \)
Change from baseline

The variance of an 8 week value is

$$\text{Var}(Y_{2,jk}) = \text{Var}(\mu + 1_{\{i=2\}} \tau_j + \xi_k + \varepsilon_{ijk})$$

$$= \text{Var}(\xi_k + \varepsilon_{ijk}) = \text{Var}(\xi_k) + \text{Var}(\varepsilon_{ijk}) = \sigma_s^2 + \sigma^2$$

Change from baseline

$$Z_{jk} = Y_{1,jk} - Y_{2,jk} = \varepsilon_{1,jk} - (\tau_j + \varepsilon_{2,jk})$$

The variance of change from baseline is

$$2\sigma^2$$

Usually

$$\sigma^2 < \sigma_s^2 \Rightarrow 2\sigma^2 < \sigma^2 + \sigma_s^2$$
Baseline as covariate

Model: \[ Y_{ij} = \mu + \tau_j + \beta x_i + \varepsilon_{ij} \]

Subject \( i = 1, \ldots, n_j \)

Treatment \( j = 1, 2, 3 \)

Treatment effect \( \tau_j \)

Baseline value \( x_i \)

Random error \( \varepsilon_{ijk} \text{ iid } N(0, \sigma^2) \)
Crossover studies

- All subject gets more than one treatment
- Comparisons within subject

- Within subject comparison
- Reduced sample size
- Good for chronic conditions
- Good for pharmaceutical studies
Model for a cross over study

\[ Y_{ijk} = \mu + \gamma_i + \xi_{i(k)} + \pi_j + \tau_t + \delta_j \lambda_r + \epsilon_{ijk} \]

- \( \gamma_i \): effect of sequence \( i = 1, 2 \) \( \gamma_1 + \gamma_2 = 0 \)
- \( \xi_{i(k)} \): effect of subject \( k = 1, \ldots, n_i \) within sequence \( i \) iid \( N(0, \sigma_s^2) \)
- \( \pi_j \): effect of period \( j = 1, 2 \) \( \pi_1 + \pi_2 = 0 \)
- \( \tau_t \): effect of treatment \( t \in \{A, B\} \) \( \tau_A + \tau_B = 0 \)
- \( \epsilon_{ijk} \): random error iid \( N(0, \sigma^2) \)

\[ \text{Obs} = \text{Period} + \text{sequence} + \text{subject} + \text{treament} + \text{carryover} + \text{error} \]
2 by 2 Crossover design

\[ E[Y_{ijk}] = \begin{pmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \end{pmatrix} \]

\[
\begin{array}{c|c}
\gamma_1 + \pi_1 + \tau_A & \gamma_1 + \pi_2 + \tau_B + \lambda_A \\
\gamma_2 + \pi_1 + \tau_B & \gamma_2 + \pi_1 + \tau_A + \gamma_B \\
\end{array}
\]

\[
\frac{1}{2}((\mu_{11} - \mu_{12}) + (\mu_{22} - \mu_{21})) =
\]

\[
\frac{1}{2}((\pi_1 - \pi_2 + \tau_A - \tau_B - \lambda_A) + (\pi_2 - \pi_1 + \tau_A - \tau_B - \lambda_B)) =
\]

\[
\tau_A - \tau_B - \frac{1}{2}(\lambda_A + \lambda_B) \quad \text{Effect of treatment and carry over can not be separated!}
\]
Matrix formulation

Model:
\[ Y_{ijk} = \mu + \gamma_i + \xi_{i(k)} + \pi_j + \tau_t + \varepsilon_{ijk} \]

Sum to zero:
\[ \lambda_1 + \lambda_2 = 0 \]
\[ \pi_1 + \pi_2 = 0 \]
\[ \tau_1 + \tau_2 = 0 \]

Matrix formulation \( \mu = X\beta \)

\[
\begin{align*}
\beta &= (\mu \quad \gamma_1 \quad \pi_1 \quad \tau_A)^T \\
\mu &= (\mu_{11}, \mu_{12}, \mu_{21}, \mu_{22})^T
\end{align*}
\]

\[
X = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 \\
1 & 1 & -1 & -1 & 1 \\
1 & -1 & 1 & -1 & 1 \\
1 & -1 & -1 & 1 & 1
\end{bmatrix}
\]
Matrix formulation

Parameter estimate: \( \hat{\beta} = (X^T X)^{-1} X^T Y \)

\[ \hat{\sigma}^2 = (Y^T Y) - \hat{\beta}X^T T \]

\[ Cov(\hat{\beta}) = (X^T X)\hat{\sigma}^2 \]

\[
(X^T X)^{-1} = \begin{bmatrix}
0.25 & 0 & 0 & 0 & 0 \\
0 & 0.25 & 0 & 0 & 0 \\
0 & 0 & 0.25 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.25 \\
0 & 0 & 0 & 0 & 0.25 \\
\end{bmatrix}
\]

Estimates independent and

\[ Var(\hat{\tau}_A) = 0.25\hat{\sigma}^2 \]
Alternatives to 2*2

Compare

\[
\begin{array}{cc}
A & B \\
B & A \\
\end{array}
\]

to

\[
\begin{array}{ccc}
A & B & B \\
B & A & A \\
\end{array}
\]

Same model but with 3 periods and a carry over effect

\[
Y_{ijk} = \mu + \gamma_i + \xi_{i(k)} + \pi_j + \tau_t + \delta_j \lambda_r + \varepsilon_{ijk}
\]

- \(\xi_{i(k)}\) = effect of subject \(k = 1, \ldots, n_i\) within sequence \(i\) iid \(N(0, \sigma_s^2)\)
- \(\gamma_i\) = effect of sequence \(i = 1, 2\) \(\gamma_1 + \gamma_2 = 0\)
- \(\pi_j\) = effect of period \(j = 1, 2, 3\) \(\pi_1 + \pi_2 = 0\)
- \(\tau_t\) = effect of treatment \(t \in \{A, B\}\)
- \(\varepsilon_{ijk}\) = random error iid \(N(0, \sigma^2)\) \(\tau_A + \tau_B = 0\)
Parameters of the AAB, BBA design

\[ Y_{ijk} = \mu + \gamma_i + \xi_{i(k)} + \pi_j + \tau_t + \epsilon_{ijk} \quad E[Y_{ijk}] = \mu_{ijk} \]

\begin{align*}
\mu_{11} &= \mu + \gamma_1 + \pi_1 + \tau_A \\
\mu_{12} &= \mu + \gamma_1 + \pi_2 + \tau_B + \lambda_A \\
\mu_{13} &= \mu + \gamma_1 + \pi_3 + \tau_B + \lambda_B \\
\mu_{21} &= \mu + \gamma_2 + \pi_1 + \tau_B \\
\mu_{22} &= \mu + \gamma_2 + \pi_2 + \tau_A + \lambda_B \\
\mu_{23} &= \mu + \gamma_2 + \pi_3 + \tau_A + \lambda_B
\end{align*}

\[
\lambda_1 + \lambda_2 = 0 \quad \pi_1 + \pi_2 + \pi_3 = 0 \quad \tau_1 + \tau_2 = 0 \quad \lambda_A + \lambda_B = 0
\]
Matrix again

\[ \mathbf{x}\beta = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & 1 \\ 1 & -1 & 1 & 0 & -1 & 0 \\ 1 & -1 & 0 & 1 & 1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \gamma_1 \\ \pi_1 \\ \pi_2 \\ \tau_A \\ \lambda_A \end{bmatrix} \]

\[ (\mathbf{x}^T\mathbf{x})^{-1} = \begin{bmatrix} 0.17 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.19 & 0 & 0 & 0.06 & 0 \\ 0 & 0 & 0.33 & -0.17 & 0 & 0 \\ 0 & 0 & -0.17 & 0.33 & 0 & 0 \\ 0 & 0.06 & 0 & 0 & 0.19 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.25 \end{bmatrix} \]

Effect of treatment and carry over can be estimated independently!
### Other 2 sequence 3 period designs

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Var}(\hat{\tau}_A)$</td>
<td>0.19</td>
<td>0.75</td>
<td>0.25</td>
<td>N/A</td>
</tr>
<tr>
<td>$\text{Var}(\hat{\lambda}_A)$</td>
<td>0.25</td>
<td>1.00</td>
<td>1.00</td>
<td>N/A</td>
</tr>
<tr>
<td>$\text{Corr}(\hat{\tau}_A, \hat{\lambda}_A)$</td>
<td>0.0</td>
<td>0.87</td>
<td>0.50</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Comparing the AB, BA and the ABB, BBA designs

\[
\begin{array}{ccc}
A & B & \text{Var}(\hat{\tau}_A) = 0.25\sigma^2 \\
B & A & \\
\end{array} \\
\begin{array}{ccc}
A & B & B & \text{Var}(\hat{\tau}_A) = 0.19\sigma^2 \\
B & A & A & \\
\end{array}
\]

Can’t include carry over  
2 treatments per subject  
Shorter duration  

Carry over estimable  
3 treatments per subject  
Longer duration

Exercise: Find the best 2 treatment 4 period design
More than 2 treatments

Tool of the trade: Define the model

Investigate $(X^T X)^{-1}$

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Watch out for drop outs!
Titration Designs

Increasing dose panels (Phase I):

- SAD (Single Ascending Dose)
- MAD (Multiple Ascending Dose)

Primary Objective:

- Establish Safety and Tolerability
- Estimate Pharmaco Kinetic (PK) profile

Increasing dose panels (Phase II):

Dose - response
Titration Designs (SAD, MAD)

Dose: $Z_1$ mg
- X on drug
- Y on Placebo

Dose: $Z_2$ mg
- X on drug
- Y on Placebo

Dose: $Z_k$ mg
- X on drug
- Y on Placebo

Stop if any signs of safety issues

VERY careful with first group!
Titration Designs

Which dose levels?

• Start dose based on exposure in animal models.
• Stop dose based on toxdata from animal models.
• Doses often equidistant on log scale.

Which subject?

• Healthy volunteers
• Young
• Male

How many subjects?

• Rarely any formal power calculation.
• Often 2 on placebo and 6-8 on drug.
Titration Designs

Not mandatory to have new subject for each group.

• Slightly larger groups to have sufficiently many exposed.
• Dose in fourth group depends on results so far.
• Possible to estimate within subject variation.
Factorial design

Evaluation of a fixed combination of drug A and drug B

<table>
<thead>
<tr>
<th>A placebo</th>
<th>A active</th>
</tr>
</thead>
<tbody>
<tr>
<td>B placebo</td>
<td>B placebo</td>
</tr>
</tbody>
</table>

The U.S. FDA’s policy (21 CFR 300.50) regarding the use of a fixed-dose combination

The agency requires:

**Each component must make a contribution to the claimed effect of the combination.**

Implication: At specific component doses, the combination must be superior to its components at the same respective doses.
Factorial design

Usually the fixed-dose of either drug under study has been approved for an indication for treating a disease.

Nonetheless, it is desirable to include placebo (P) to examine the sensitivity of either drug given alone at that fixed-dose (comparison of AB with P may be necessary in some situations).

Assume that the same efficacy variable is used for studying both drugs (using different endpoints can be considered and needs more thoughts).
Factorial design

Sample mean $Y_i \sim N( \mu_i, \sigma^2/n )$, $i = A, B, AB \ n = \text{sample size per treatment group (balanced design is assumed for simplicity)}$.

$H_0: \mu_{AB} \leq \mu_A \ or \ \mu_{AB} \leq \mu_B$

$H_1: \mu_{AB} > \mu_A \ and \ \mu_{AB} > \mu_B \ j=A, B$

$$T_{AB:j} = \sqrt{\frac{n}{2}} \left( \frac{Y_{AB} - Y_j}{\hat{\sigma}} \right); \ j = A, B$$

Min test and critical region: $\min(T_{AB:A}, T_{AB:B}) > C$
Group sequential designs

A large study is a huge investment, $, ethics

• What if the drug doesn’t work or is much better than expected?
• Could we take an early look at data and stop the study if it looks good (or too bad)?
Repeated significance test

Let: \( Y_{ij} \sim N(\mu_j, \sigma^2) \)

Test: \( H_0 : \mu_1 = \mu_2 \)

Test statistic: \( Z_{mk} = \frac{\hat{\mu}_1 - \hat{\mu}_2}{\sqrt{2\sigma^2/mk}} \)

\( \hat{\mu}_j = \frac{1}{mk} \sum_{i=1}^{mk} Y_{ij} \)

For \( k = 1, \ldots, K - 1 \)

If \( |Z_k| \geq C_\alpha \) stop, reject \( H_0 \)
otherwise continue to group \( k + 1 \)

If \( |Z_k| \geq C_\alpha \) stop, reject \( H_0 \)
otherwise stop accept \( H_0 \)
## True type I error rate

Repeat testing until $H_0$ rejected

<table>
<thead>
<tr>
<th>Tests</th>
<th>Critical value</th>
<th>$P$(type I error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.96</td>
<td>0.05</td>
</tr>
<tr>
<td>2</td>
<td>1.96</td>
<td>0.08</td>
</tr>
<tr>
<td>3</td>
<td>1.96</td>
<td>0.11</td>
</tr>
<tr>
<td>4</td>
<td>1.96</td>
<td>0.13</td>
</tr>
<tr>
<td>5</td>
<td>1.96</td>
<td>0.24</td>
</tr>
</tbody>
</table>
Pocock’s test

Suppose we want to test the null hypothesis 5 times using the same critical value each time and keep the overall significance level at 5%.

For $k = 1, \ldots K - 1$ 

If $|Z_k| \geq C_p(\alpha, K)$ Stop, reject $H_0$

otherwise Continue to group $k + 1$

After group K

If $|Z_k| \geq C_p(\alpha, K)$ Stop, reject $H_0$

otherwise stop accept $H_0$

Choose $C_p(\alpha, K)$ Such that

$P(\text{Reject } H_0 \text{ at any analysis } k = 1 \ldots K) = \alpha$
Pocock’s test

<table>
<thead>
<tr>
<th>Tests</th>
<th>Critical value</th>
<th>P(type I error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.960</td>
<td>0.05</td>
</tr>
<tr>
<td>2</td>
<td>2.178</td>
<td>0.05</td>
</tr>
<tr>
<td>3</td>
<td>2.289</td>
<td>0.05</td>
</tr>
<tr>
<td>4</td>
<td>2.361</td>
<td>0.05</td>
</tr>
<tr>
<td>5</td>
<td>2.413</td>
<td>0.05</td>
</tr>
</tbody>
</table>

All tests has the same nominal significance level

A group sequential test with 5 interim tests has level

\[ \alpha' = 2(1 - \phi(2.413)) = 0.0158 \]
Pocock’s test

Accept $H_0$

Reject $H_0$

$Z_k$

2.413

-2.413

stage $k$
O’Brian & Flemmings test

Increasing nominal significance levels

For \( k = 1, \ldots K - 1 \) : If \( |Z_k| \geq C_p(\alpha, K)\sqrt{K/k} \) Stop, reject \( H_0 \)
otherwise Continue to group \( k + 1 \)

After group \( K \) : If \( |Z_k| \geq C_p(\alpha, K) \) Stop, reject \( H_0 \)
otherwise stop accept \( H_0 \)

Choose \( C_p(\alpha, K) \) Such that
\[ P(\text{Reject} \ H_0 \ \text{at any analysis} \ k = 1 \ldots K) = \alpha \]
O’Brian & Flemmings test

Critical values and nominal significance levels for a O’Brian Flemming test with 5 interim tests.

<table>
<thead>
<tr>
<th>Test (k)</th>
<th>( C_B(K, \alpha) )</th>
<th>( C_B(K, \alpha) \cdot \text{Sqrt}(K/k) )</th>
<th>( \alpha' )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.04</td>
<td>4.56</td>
<td>0.000005</td>
</tr>
<tr>
<td>2</td>
<td>2.04</td>
<td>3.23</td>
<td>0.0013</td>
</tr>
<tr>
<td>3</td>
<td>2.04</td>
<td>2.63</td>
<td>0.0084</td>
</tr>
<tr>
<td>4</td>
<td>2.04</td>
<td>2.28</td>
<td>0.0225</td>
</tr>
<tr>
<td>5</td>
<td>2.04</td>
<td>2.04</td>
<td><strong>0.0413</strong></td>
</tr>
</tbody>
</table>

Rather close to 5%
O’Brian & Flemmings test

Stage K

Z_k

0.000005

0.0013

0.0084

0.0225

0.0413

0.0013

0.0084

0.0225

0.0413
Comparing Pocock and O’Brian Flemming

<table>
<thead>
<tr>
<th>Test (k)</th>
<th>$C_B(K,a)\times\text{Sqrt}(K/k)$</th>
<th>$\alpha'$</th>
<th>$C_P(K,a)$</th>
<th>$\alpha'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.56</td>
<td>0.00001</td>
<td>2.413</td>
<td>0.0158</td>
</tr>
<tr>
<td>2</td>
<td>3.23</td>
<td>0.0013</td>
<td>2.413</td>
<td>0.0158</td>
</tr>
<tr>
<td>3</td>
<td>2.63</td>
<td>0.0084</td>
<td>2.413</td>
<td>0.0158</td>
</tr>
<tr>
<td>4</td>
<td>2.28</td>
<td>0.0225</td>
<td>2.413</td>
<td>0.0158</td>
</tr>
<tr>
<td>5</td>
<td>2.04</td>
<td>0.0413</td>
<td>2.413</td>
<td>0.0158</td>
</tr>
</tbody>
</table>
Comparing Pocock and O’Brian Flemming

![Graph comparing Pocock and O’Brian Flemming](image)
Group Sequential Designs

Pros:

• Efficiency Gain (Decreasing marginal benefit)
• Establish efficacy earlier
• Detect safety problems earlier

Cons:

• Smaller safety data base
• Complex to run
• Need to live up to stopping rules!
Selection of a design

The design of a clinical study is influenced by:

- Number of treatments to be compared
- Characteristics of the treatment
- Characteristics of the disease/condition
- Study objectives
- Inter and intra subject variability
- Duration of the study
- Drop out rates
Backup

Back up:

\[ Var(\bar{Y}_{21} - \bar{Y}_{22}) = Var(\bar{Y}_{21}) + Var(\bar{Y}_{22}) \]
\[ = \frac{1}{n_1^2} \left( \sigma_s^2 + \sigma^2 \right) + \frac{1}{n_2^2} \left( \sigma_s^2 + \sigma^2 \right) \]