Adaptive Design

Mark Chang, Ph.D.

Director, Biostatistics
Millennium Pharmaceuticals, Inc. USA
Mark.chang@statisticians.org

www.statisticians.org

June 3, 2007, NC, USA
What Is An Adaptive Design

• Compared to a classic trial design with static features, an adaptive design allows for changing or modifying the characteristics of a trial based on cumulative information
  – Increase the probability of success
  – Reduced the cost
  – Reduce the time to market
  – Deliver the right drug to the right patient
  – Preserve the integrity and validity
Classic Group Sequential Design

Group sequential design allows for decision based on interim results to claim efficacy, futility, or continue to next stage with a fixed sample size increment.
Sample Size Reestimation

Active

PLA

Interim results indicates additional patients required to preserve the power
Drop-Loser Design

Interim results indicate that some arms are inferior and dropped from the study.
Response Adaptive Randomization

Add certain type of balls into the urn according to the treatment response

Assign treatment according to the type of ball selected at random.
Dose-Escalation for Maximum Tolerated Dose

Group of 3 patients treated; toxicity measured by DLTs
Interim results indicate people with Gen-x are much more responsive to the drug, therefore the trial continues for the subpopulation with Gen-x.
Adaptive Treatment Switching

Treatments are switched when the disease is progressed due to ethic considerations.
Clinical Trial Simulation Model

Adaptations

Design model

Execution model

Response model

IDMC decision, Market Dynamics

Evaluation model
Bayesian Decision Theory for Competing Constraints
Bayesian Decision Approach

Trial objectives => utility function

Expected utility function

Prior knowledge => prior probability

Posterior probability => probability of outcome

Incremental trial data => likelihood function

Adaptations to maximize the expected utility

Stop based on stopping criteria
Characteristics of Adaptive Design

- Flexible
- Dynamic
- Sequential learning
- Cost-efficient
- Robust
- Real-time
- Simulation
- Bayesian
- Integrity
- Validity
- Decision-oriented
- Systematic
- Data-driven
- Optimized
- Streamlined

Adaptive Design
History of Adaptive Design

- Simon Two-Stage Design (Simon, 1987)
- Group Sequential Method (Pocock, 1977; O'Brien, 1979)
- Bayesian Designs (Berry, Louis, Goodman, Spiegelhalter, et al.)

- B-Value Monitoring Tool (Lan & Wittes, 1988)
- Error Spending Approach (Lan & DeMets, 1983)

- Hypothesis Adaptive Design (Hommel, 2001)
- Finsher’s Product of P-values (Bauer & Kohne, 1994)
- Drop-Loser Adaptive Design (Bauer & Kieser, 1999)

- Uniformly Most Powerful Design (Tsiatis and Mehta, 2003)
- Conditional Error Function (Proschan & Hunsberger, 1995)
- Sample-Size and Inference (Liu & Chi, 2001)

- Impact of Protocol Amendment (Chow, Chang & Pong, 2005)
- Fixed Weighting Method (Cui, Hung & Wang, 1999)
- Recursive Combination Tests (Brannath, Posch & Bauer, 2002)

- Bayesian Adaptive Design (Chang & Chow, 2005)
- Inverse-Normal Method (Lehmacher & Wassmer, 1999)
- Sample-Size Modification (Jennison & Turnbull, 2003)

- Bayesian Adaptive Design (Chang & Chow, 2005)
- Conditional Error Approach (Muller-Scheffä, 2001)
- Decision-Function Method (Muller & Schafer, 2004)

- Sum of P-values Approach (Chang, 2006)
- Pick-winner Adaptive Design (Sampson & Sill, 2005)

- A Regulatory View (Hung, O’Neill, Wang, Lawrence, 2006)
- Recursive Two-Stage Adaptive (Chang, 2006)
- Multi-Endpoint Adaptive Design (Chang, et al., 2007)
How to Design
Adaptive Trial
General Approach to Adaptive Designs

Stopping Rules:

\[ T_k = \text{Function of } p_1, p_2, \ldots, p_k; \quad \alpha_K = \beta_K \text{ for the final stage } K. \]

\[ p_i = \text{stagewise p-value based on subsamples at stage } i. \]
## Two-Stage Adaptive Design Methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Test Statistic at Stage 2</th>
<th>Stopping Boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSP</td>
<td>$p_1 + p_2$</td>
<td>$\alpha_1 + \frac{1}{2} (\alpha_2 - \alpha_1)^2 = \alpha$</td>
</tr>
<tr>
<td>MPP</td>
<td>$p_1 p_2$</td>
<td>$\alpha_1 + \alpha_2 \ln \frac{1}{\alpha_1} = \alpha$</td>
</tr>
<tr>
<td>MINP</td>
<td>$\Phi \left( w_1 Z_{1-p_1} + w_2 Z_{1-p_2} \right)$, where predetermined weights $w_1^2 + w_2^2 = 1$.</td>
<td>Numerical integrations</td>
</tr>
</tbody>
</table>

Test statistic at stage 1: $T_1 = p_1$.

$\alpha =$ overall type-I error rate.
Example of Stopping Boundaries (non-binding futility rule)

Ex: \( \alpha_1 = 0.0025, \beta_1 = 0.5, \) then

MSP: \( \alpha_2 = \sqrt{2(\alpha - \alpha_1)} + \alpha_1 = \sqrt{2(0.025 - 0.0025)} + 0.0025 = 0.2146 \)

MPP: \( \alpha_2 = \frac{\alpha_1 - \alpha}{\ln \alpha_1} = \frac{0.0025 - 0.025}{\ln 0.0025} = 0.0037553 \)

MINP: \( \alpha_2 = 0.0240 \) (simulation/numerical integration)
Myocardial Infarction Prevention Trial (1) - Design

- Event rate 22% to 11%
- The stopping rules:
  - MSP: $\alpha_1 = 0$, $\beta_1 = 0.2$, $\alpha_2 = 0.225$; one-sided $\alpha = 0.025$. $N_{\text{max}} = 500$ per group.
  - Stage 1: Stop for futility if $p_1 > \beta_1$, and stop for efficacy if $p_1 \leq \alpha_1$; otherwise, continue.
  - Final stage: If $p_1 + p_2 \leq \alpha_2$, claim efficacy; otherwise claim futility.
Comparison of Classic, Group Sequential, and Adaptive Designs (Ref. Slide 73)

<table>
<thead>
<tr>
<th>Design</th>
<th>Event Rate in the Test Group $P_T$</th>
<th>0.110</th>
<th>0.14</th>
<th>0.22</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Na</td>
<td>Power (%)</td>
<td>Na</td>
<td>Power (%)</td>
</tr>
<tr>
<td>Classic</td>
<td>600</td>
<td>94.2</td>
<td>600</td>
<td>72.3</td>
</tr>
<tr>
<td>GSD</td>
<td>588</td>
<td>92.0</td>
<td>550</td>
<td>67.0</td>
</tr>
<tr>
<td>SSR</td>
<td>928</td>
<td>95.8</td>
<td>874</td>
<td>80.0</td>
</tr>
</tbody>
</table>

Note: $\alpha_1 = 0$, $\beta_1 = 0.2$, $N_{\text{max}} = 500$/group, $cP = 0.95$. $\bar{N}_o (\bar{N}_a) = \text{average total sample-size under } H_o (H_a)$. 
How to Monitor

Adaptive Trial
Sample-Size Re-Estimation Based on Conditional Power

<table>
<thead>
<tr>
<th>Method</th>
<th>Sample-Size Per Group at Stage 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSP</td>
<td>[n_2 = \frac{2\sigma^2}{\delta^2} \left[ Z_{1-\max(0,\alpha_2-p_1)} - Z_{1-cP} \right] ]</td>
</tr>
<tr>
<td>MPP</td>
<td>[n_2 = \frac{2\sigma^2}{\delta^2} \left[ Z_{1-\frac{\alpha_2}{p_1}} - Z_{1-cP} \right]^2 ]</td>
</tr>
<tr>
<td>MINP</td>
<td>[n_2 = \frac{2\sigma^2}{\delta^2} \left[ \frac{1}{w_2} \left( Z_{1-\alpha_2} - w_1 Z_{1-p_1} \right) - Z_{1-cP} \right]^2 ]</td>
</tr>
</tbody>
</table>

Note: \( \delta = \) treatment effect, \( cP = \) target conditional power
Examples of Sample-Size Re-Estimation (Ref. Slide 71)

<table>
<thead>
<tr>
<th>Model</th>
<th>$\alpha_2$</th>
<th>1st Stage P-value $p_1$</th>
<th>Target Cond. Power</th>
<th>N at Stage 2 Per Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSP</td>
<td>0.2050</td>
<td>0.1</td>
<td>0.9</td>
<td>191</td>
</tr>
<tr>
<td>MPP</td>
<td>0.0043</td>
<td>0.1</td>
<td>0.9</td>
<td>408</td>
</tr>
<tr>
<td>MINP</td>
<td>0.0226</td>
<td>0.1</td>
<td>0.9</td>
<td>364</td>
</tr>
</tbody>
</table>

Note: effect size (eSize) = 0.21, eSize = $\delta/\sigma$.

For binary endpoint, $\sigma^2 = P_c(1 - P_c)/2 + P_t(1 - P_t)/2$.
$P_c$ and $P_t$ = event rate for the test and control groups.
Conditional Power Formula

<table>
<thead>
<tr>
<th>Method</th>
<th>Conditional Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSP</td>
<td>[1 - \Phi \left( Z_{1 - \max(0, \alpha_2 - p_1)} - \frac{\delta}{\sigma} \sqrt{\frac{n_2}{2}} \right)]</td>
</tr>
<tr>
<td>MPP</td>
<td>[1 - \Phi \left( Z_{1 - \frac{\alpha_2}{p_1}} - \frac{\delta}{\sigma} \sqrt{\frac{n_2}{2}} \right)]</td>
</tr>
<tr>
<td>MINP</td>
<td>[1 - \Phi \left( \frac{1}{w_2} \left[ Z_{1 - \alpha_2} - w_1 Z_{1 - p_1} \right] - \frac{\delta}{\sigma} \sqrt{\frac{n_2}{2}} \right)]</td>
</tr>
</tbody>
</table>

Assume \( \alpha_1 < p_1 \leq \beta_1 \).

\( n_2 = \) sample-size per group at stage 2.
Examples of Conditional Power (Ref. Slide 75)

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>MSP</th>
<th>MPP</th>
<th>LM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_2$</td>
<td>0.2050</td>
<td>0.0043</td>
<td>0.0226</td>
</tr>
<tr>
<td>Binary</td>
<td>Cond. Power</td>
<td>0.967</td>
<td>0.915</td>
</tr>
<tr>
<td>Normal</td>
<td>Cond. Power</td>
<td>0.772</td>
<td>0.611</td>
</tr>
</tbody>
</table>

Note: $\alpha_1 = 0.005$, $\beta_1 = 0.25$, and $p_1 = 0.01$.

Binary: event rate $P_t = 0.2$, $P_c = 0.4$; $n_2 = 100$

Normal: mean $u_t = 0.2$, $u_c = 0.4$, $\sigma = 1$; $n_2 = 200$
Projection of Number of Events for Survival Trials

Assumption: Exponential distribution

Number of Deaths:
\[ D = \begin{cases} 
R \left( T - \frac{1}{\lambda} + \frac{1}{\lambda} e^{-\lambda T} \right) & \text{if } T \leq T_0, \\
R \left[ T_0 - \frac{1}{\lambda} \left( e^{\lambda T_0} - 1 \right) e^{-\lambda T} \right] & \text{if } T > T_0,
\end{cases} \]

Time of interesting:
\[ T = \begin{cases} 
\frac{1}{\lambda} \ln \left( \frac{\lambda D}{R} - \lambda T + 1 \right) & \text{if } T \leq T_0, \\
\frac{1}{\lambda} \ln \left[ \lambda \left( T_0 - \frac{D}{R} \right) \left( e^{\lambda T_0} - 1 \right)^{-1} \right] & \text{if } T > T_0
\end{cases} \]

\( T_0 = \) enrollment duration, \( R = \) enrollment rate, \( \lambda = \) hazard rate, \( T_{\text{median}} = \frac{\ln 2}{\lambda} = T_{\text{mean}} \ln 2. \)
How to Analyze Adaptive Trial
### Adjusted (Stagewise-Ordering) P-values

<table>
<thead>
<tr>
<th>Method</th>
<th>Adjusted P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSP</td>
<td>$\alpha_1 + \frac{1}{2} \left( p_1 + p_2 - \alpha_1 \right)^2$</td>
</tr>
<tr>
<td>MPP</td>
<td>$\alpha_1 + p_1 p_2 \ln \frac{1}{\alpha_1}$</td>
</tr>
<tr>
<td>MINP</td>
<td>Simulations</td>
</tr>
</tbody>
</table>

If the trial stops at stage 2, p-value adjustment is based on the above table; if trial stops at stage 1, p-value = $p_1$. 
Meaning of Stagewise-Ordering P-values

- When the Ho is rejected, then the later the rejection is, the bigger the stagewise-ordering p-value is.
- The order of the test statistic is
  - $T_1 > T_2 > \ldots, > T_K$.
  - For a given stage, the order of the test statistic is based on their values.
- When Ho is marginally rejected, the stagewise-ordering p-value is equal to the alpha spent thus far.
Confidence Interval Concept in Adaptive Designs

• From the duality of the confidence and hypothesis test, a $100(1-\alpha)\%$ CI consists all $\delta_0$ such that the null hypothesis $H_0: \delta \leq \delta_0$ is not rejected.

• The stagewise p-value at stage $i$:

$$p_i = 1 - \Phi \left( \frac{\hat{\delta}_i - \delta_0}{\sigma} \sqrt{\frac{n_i}{2}} \right)$$

$\hat{\delta}_i$ = naive treatment difference based on subsample at stage $i$. 
### Confidence Interval Formulation

<table>
<thead>
<tr>
<th>Method</th>
<th>Equation for Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSP</td>
<td>[ \sum_{i=1}^{k} \left[ 1 - \Phi \left( \frac{\hat{\delta}<em>i - \delta</em>{\alpha_k}}{\sigma} \sqrt{\frac{n_i}{2}} \right) \right] = \alpha_k, \ k = 1, 2 ]</td>
</tr>
<tr>
<td>MPP</td>
<td>[ \prod_{i=1}^{k} \left[ 1 - \Phi \left( \frac{\hat{\delta}<em>i - \delta</em>{\alpha_k}}{\sigma} \sqrt{\frac{n_i}{2}} \right) \right] = \alpha_k, \ k = 1, 2 ]</td>
</tr>
<tr>
<td>MINP</td>
<td>[ \sum_{i=1}^{k} w_{ki} \left( \frac{\hat{\delta}<em>i - \delta</em>{\alpha_k}}{\sigma} \sqrt{\frac{n_i}{2}} \right) = \alpha_k, \ k = 1, 2 ]</td>
</tr>
</tbody>
</table>

100(1 − α)% confidence limit: \[ \delta_{0_{\text{min}}} = \max \{ \delta_{\alpha_1}, \delta_{\alpha_2} \} \]
Point Estimate

Naive: \[ \hat{\delta} = \frac{1}{n_1+n_2} \left( n_1 \hat{\delta}_1 + n_2 \hat{\delta}_2 \right) \]

Fixed weights: \[ \hat{\delta}_u = \omega_1 \hat{\delta}_1 + \omega_2 \hat{\delta}_2 \]

Symmetric boundaries: \[ \hat{\delta}_u = \frac{\sqrt{\omega_1 n_1 \hat{\delta}_1} + \sqrt{\omega_2 n_2 \hat{\delta}_2}}{\sqrt{\omega_1 n_1} + \sqrt{\omega_2 n_2}} \]

Predetermined weights satisfying: \( \omega_1 + \omega_2 = 1 \)
Myocardial Infarction Prevention Trial (2) - Interim Analysis

- Interim analysis: \( n_1 = 150 \) patients/group

- Observed event rates = 0.22 and 0.165 for the two groups

- Stagewise p-value \( p_1 = 1 - \Phi \left( \frac{(0.22 - 0.165) \sqrt{150}}{\sqrt{0.31}} \right) = 1 - \Phi (1.2111) = 0.1129 < \beta_1 = 0.2 \).  
  \( \Rightarrow \) proceed to the second stage

- New sample-size \( n_2 = 650 \) based on:
  
  target conditional power = 90%

  effect-size = \( \frac{\delta}{\sigma} = \frac{0.22 - 0.165}{\sqrt{0.31/2}} = 0.14 \).
Myocardial Infarction Prevention Trial (3)  
- Final Analysis

- **Financial constraint:** \( n_2 = 400 \) is actually used, which provides about 78% conditional power.

- **Observed event rates:** 0.22 and 0.175 for the two groups

- **Stagewise p-value:** \( p_2 = 1 - \Phi \left( \frac{(0.22 - 0.175) \sqrt{400}}{\sqrt{0.316}} \right) \)

  \[
  = 1 - \Phi(1.6011) = 0.0547.
  \]

- **Test statistic:** \( t = p_1 + p_2 = 0.1676 < \alpha_2 = 0.225 \)

  \( \Rightarrow \) The null hypothesis is rejected.
Myocardial Infarction Prevention Trial (4) - Adjusted P-value and Confidence Interval

Adjusted p-value (Ref: slides 20):

\[ p_{adj} = \alpha_1 + \frac{1}{2}(t - \alpha_1)^2 = 0 + \frac{1}{2}(0.2^2) = 0.02. \]

Confidence Interval (Ref: slide xx):

\[ \delta_{\alpha_1} = -\infty. \]

\[ 2 - \Phi \left( \frac{(0.055 - \delta_{\alpha_2})\sqrt{150}}{\sqrt{0.31}} \right) - \Phi \left( \frac{(0.045 - \delta_{\alpha_2})\sqrt{400}}{\sqrt{0.316}} \right) = 0.225 \]

Using trial-error method \( \Rightarrow \delta_{\alpha_2} = 0.007. \)

Confidence limit: \( \delta_{0\min} = \max \{ \delta_{\alpha_1}, \delta_{\alpha_2} \} = 0.007. \)
Myocardial Infarction Prevention Trial (5) - Point Estimates

- **Naive estimate:**

\[
\hat{\delta} = \frac{150(0.055)}{550} + \frac{400(0.045)}{550} = 0.0477
\]

- **Fixed weight method** \((w_i = \sqrt{0.5})\):

\[
\tilde{\delta}_u = \frac{0.055\sqrt{0.5}\sqrt{150} + 0.045\sqrt{0.5}\sqrt{400}}{\sqrt{0.5}\sqrt{150} + \sqrt{0.5}\sqrt{400}} = 0.0488.
\]

- **Estimate with symmetric boundaries:**

\[(\omega_1 = \frac{2(150)}{300 + 550} = \frac{6}{17} \text{ and } \omega_2 = 11/17)\):

\[
\delta_u = \frac{6(0.055)}{17} + \frac{11(0.045)}{17} = 0.0485.
\]
Summary Results of Myocardial Infarction Prevention Trial

The test drug has about a 4.8% event reduction with a one-sided 97.5% confidence limit of 0.7%. The adjusted p-value (one-sided) is 0.02.
How to Optimize
Adaptive Design
Evaluation of Trial Designs

- Power, Cost, Time
- Power is not a good criterion for comparison
  - Two design (classic vs. adaptive) with same power, which one you choose?
  - Classic design with 90% power versus adaptive design 88% power, but later give a reasonable chance to deliver drug much early, which is better?
Evaluation of Adaptive Design
- Operating Characteristics

- Efficacy stopping probabilities (ESP)
- Futility stopping probabilities (FSP)
- Expected sample sizes
  - under Ho, Ha, Hs
- Nmax
- Expected Trial Durations
  - under Ho, Ha, Hs
- IA Times, Study duration
Examples of Operating Characteristics

<table>
<thead>
<tr>
<th>Condition</th>
<th>FSP</th>
<th>ESP</th>
<th>$\tilde{N}$</th>
<th>$N_{max}$</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_o$</td>
<td>0.849</td>
<td>0.010</td>
<td>177</td>
<td>310</td>
<td>(0.025)</td>
</tr>
<tr>
<td>$H_a$</td>
<td>0.039</td>
<td>0.682</td>
<td>198</td>
<td>310</td>
<td>0.949</td>
</tr>
<tr>
<td>$H_s$</td>
<td>0.167</td>
<td>0.373</td>
<td>226</td>
<td>310</td>
<td>0.743</td>
</tr>
</tbody>
</table>
Calculating n based on $\delta$

- Power is dependent on $\delta_{\text{true}}$
- Sample-size calculation based on $\delta_{\text{min}}$ is not appreciated:
  - P-value is much smaller than $\alpha$ when $\delta_{\text{obs}} = \delta_{\text{min}}$
  - P-value could be $< \alpha$ even when $\delta_{\text{obs}} < \delta_{\text{min}}$
What If a Trial Is Powered Based on $\delta_{\text{min}}$?

Statistical significance is declared even when $\delta_{\text{obs}} = 0.61\delta_{\text{min}}$.

90% power-line based on $\delta_{\text{min}}$

$p = 0.0006$ when $\delta_{\text{obs}} = \delta_{\text{min}}$
Power Is Not Probability of Success

- Increasing sample size does not make a drug better, but it improves the precision of the estimate.
- Statistical significance only proves treatment effect $\delta > 0$
- Power = $\Pr (p < \alpha)$
- $\text{Pe} = \Pr (p < \alpha & \delta_{\text{obs}} > \delta_{\text{min}})$
  \[ > \Pr (\text{Lower-CI-Limit} > \delta_{\text{min}}) \]

$\delta_{\text{min}}$ = the minimal treatment difference with both clinical and commercial values.
How to Select
Design Method
### Conditional Error Function
- Stopping Boundary Based on $p_2$

<table>
<thead>
<tr>
<th>Method</th>
<th>Conditional Error Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSP</td>
<td>$p_1 + p_2 = \alpha_2$ $\Rightarrow$ $p_2 = \alpha_2 - p_1$</td>
</tr>
<tr>
<td>MPP</td>
<td>$p_1p_2 = \alpha_2$ $\Rightarrow$ $p_2 = \frac{\alpha_2}{p_1}$</td>
</tr>
<tr>
<td>MINP</td>
<td>$\ldots \Rightarrow p_2 = 1 - \Phi \left( \frac{Z_{1-\alpha_2} - w_1 Z_{1-p_1}}{w_2} \right)$</td>
</tr>
</tbody>
</table>
Stopping Boundaries at Stage 2

A: Rejected by all 3 methods
B: Rejected by MPP, MINP, not MSP
C: Rejected by MSP, not MPP or MINP

MPP: \( p_1p_2 = \alpha_2 \)
MINP: \( \Phi[2.771-\Phi^{-1}(p_1)] \)
MSP: \( p_1 + p_2 = \alpha_2 \)
Comparisons of Conditional Powers

![Graph comparing conditional powers](image)

- MSP
- MPP
- MINP

Conditional Power vs. P-value $p_1$ from Stage 1
Other Adaptive Design Methods

- Conditional Error Function Method
- Error-Spending Method
- Recursive Adaptive Design
- Relationships among different methods
IT Infrastructure

• Software and services:
  – CDISC, EDC
  – ExpDesign Studio, East, Clinical Data Workbench, IVRS.

• Source code: SAS Macros and R-functions
Some Tips for Design Optimization

• The efficiency of a design is sensitive to the sample-size fraction or information time at interim analysis. Therefore simulations should be done to determine the best information time for the interim analysis.

• Practically, the adaptive trials require early efficacy readouts. This early efficacy assessment can be the primary endpoint for the trial, biomarker, or surrogate endpoint.

• When re-estimate sample-size, use the new sample-size and assume that the observed effect = \( \delta_{\text{min}} \), the minimum clinically meaningful effect to calculate the p-value. If p-value is much smaller than what needs to reject the Ho, then the trial is likely overpowered (Why?).
Drop-Loser Design
Seamless Design
Adaptive Dose-Finding
Seamless Adaptive Design

Randomize patients with five initial arms

Interim Analysis

Drop inferior arms & early stop trial or adjust sample size

Frequentist & Bayesian results

Active

PLA

Response

Dose

P-values &

Probability

Treatment effect

Learning Phase

Confirmatory Phase
Modified Bauer-Kieser Method

• M treatment groups
• $m_1$ comparisons at stage 1
• $m_2$ comparisons at stage 2
• $p_1 \Rightarrow m_1 p_1$
• $p_2 \Rightarrow m_2 p_2$
Test Statistic and Stopping Boundaries for Drop-Loser Design

<table>
<thead>
<tr>
<th>Method</th>
<th>Test Statistic</th>
<th>Stopping Boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSP</td>
<td>( m_1 p_1 + m_2 p_2 )</td>
<td>( \alpha_1 + \frac{1}{2} (\alpha_2 - \alpha_1)^2 = \alpha )</td>
</tr>
<tr>
<td>MPP</td>
<td>( m_1 m_2 p_1 p_2 )</td>
<td>( \alpha_1 + \alpha_2 \ln \frac{1}{\alpha_1} = \alpha )</td>
</tr>
</tbody>
</table>
| MINP   | \( \Phi \left( w_1 Z_{1-p_1} + w_2 Z_{1-p_2} \right) \),  
          where \( w_1^2 + w_2^2 = 1 \). | Numerical integrations |

Test statistic at stage 1: \( T_1 = m_1 p_1 \).
An Asthma Trial

Five-Arm Trial: Placebo & 4 active arms

Primary Endpoint: FEV1 change from baseline

<table>
<thead>
<tr>
<th>Arms</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEV1 change</td>
<td>0.06</td>
<td>0.12</td>
<td>0.13</td>
<td>0.14</td>
<td>0.15</td>
</tr>
</tbody>
</table>
Seamless Design of Asthma Trial

- IA on 50% patients picking a single winner
- Early stopping trial for efficacy or futility.
- The winner and placebo continue at stage 2.
- Boundaries: $\alpha_1 = 0.01$, $\beta_1 = 0.15$, $\alpha_2 = 0.1871$.
- Adjusted p-value at stage 1: $\tilde{p}_1 = 4p_{1\min}$.
- Decision rules: if $\tilde{p}_1 \leq \alpha_1$, stop for efficacy;
  
  if $\tilde{p}_1 > \beta_1$, stop for futility;
  
  if $\alpha_1 < \tilde{p}_1 \leq \beta_1$, continue.
- At stage 2, if $\tilde{p}_1 + p_2 \leq \alpha_2$, claim efficacy.
## Simulation Results – Seamless Design for Asthma Trial (Ref: Slides 77)

<table>
<thead>
<tr>
<th>Seamless Design under Global Ho</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>Nmax</td>
<td>FSP</td>
<td>ESP</td>
<td>AveN</td>
<td>Power</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>100</td>
<td>.97</td>
<td>.002</td>
<td>252</td>
<td>.002</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Seamless Design Under Ha</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>Nmax</td>
<td>FSP</td>
<td>ESP</td>
<td>AveN</td>
<td>Power</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>100</td>
<td>.009</td>
<td>.552</td>
<td>285</td>
<td>0.903</td>
<td></td>
</tr>
</tbody>
</table>
Normal Approximation Method

\[ Z^w = I_{AB}Z^A + (1 - I_{AB})Z^B, \]

where \( I_{AB} = 1 \) if arm \( B \) is dropped; otherwise \( I_{AB} = 0 \); \( Z^A = \Phi(1 - p_A) \) and \( Z^B = \Phi(1 - p_B) \).

Shun, Soo, and Lan (2007) found that under the global null hypothesis, \( Z^w \) is approximately normal distributed with mean \( E(Z^w) = \sqrt{\frac{\tau}{2\pi}} \), and \( \text{var}(Z^w) = 1 - \frac{\tau}{2\pi} \), where information time for the interim analysis \( \tau = \frac{n_1}{n} \) (the sample-size fraction at the interim analysis). Therefore they proposed a test statistic as

\[ Z^{ws} = \frac{Z^w - \sqrt{\frac{\tau}{2\pi}}}{\sqrt{1 - \frac{\tau}{2\pi}}}, \]

which has approximately the standard normal distribution.
Practical Issues
Issues and Challenges in Adaptive Design (1)

- Consistency requirement between phases
- Smaller $p_1$ at IA => don’t reject Ho; at FA => reject Ho.
  - OB-F boundary for GSD:
    - $p_1 = 0.00251$ => failed to reject Ho IA; one-sided $p_2 = 0.51$ (wrong direction) => reject Ho.
  - MPP for AD:
    - When $p_1 = 0.00251$, any $p_2$ will lead to rejection of Ho.
  - MSP for AD:
    - $p_1 = 0.00251$ => failed to reject, $p_2 = 0.22$ => rejection of Ho.
  - Classic designs:
    - We don’t even bother to split data for checking the consistency.

- AD with MSP has built-in consistency constraints.
- Classic design, GSD and AD with MINP or MPP has virtually no built-in consistency requirements.
Issues and Challenges in Adaptive Design (2)

• Information mask in sample-size adjustment
  – GSD: $\Delta N = 0$ or $N2$
  – AD: $\Delta N = 0, 10, 20, 35, \ldots, N2, \ldots$

• Multiple-endpoint issues
  – Correlation between endpoints
  – Adaptation based on one endpoint may affect others
  – Closed test procedure available based on p-combinations

• Operational Bias
  – time-dependent factors in response-adaptive randomization
  – Roles of IDMC
Regulatory Aspects
FDA Perspectives (Hung, etc. 2006).

• Adaptive designs received by FDA have increased in the past five years.
  – extension of sample size, termination of a treatment arm
  – change of the primary endpoint, change of statistical tests
  – change of the study objective such as from superiority to non-inferiority or vice versa
  – selection of a subgroup based upon externally available studies.
FDA Expectations
(Sue-Jane Wang, 2006)

1. Is prospectively planned;
2. Has valid statistical approaches on modification of design elements that have alpha control and can be defined in terms of ICH E-9 standard;
3. Has valid point estimates and confidence interval estimates;
4. Builds on experience from external trials;
5. Takes a "learn" and "confirm" approach;
6. Has standard operating procedures and infrastructure for adaptive process monitoring to avoid bias;
7. Has SOPs on adaptive design decisions; and
8. Includes documentation of actual monitoring process, extent of compliance and potential effect on study results.
Summary

- Many successful stories from different companies
- Type-I error control – Methods available
- Non-binding Futility rule – Currently adopted by regulatory authorities
- Adjusted p-values – Several controversial versions available
- Confidence intervals - Several controversial definitions, some are available.
- Point estimation – remain challenging, available only for some very simple adaptive designs
- Several simulation software packages, SAS Macros and R-functions are available
Recommendations

• Learning by doing
• 3-Rs
  – the Right objectives
  – the Right evaluation criteria
  – the Right team members
• Interacting with FDA earlier,
• Thinking about execution challenges
References

Books:
References

SAS Macro: Sample-Size at Stage 2 Based on Conditional Power

%Macro nByCPowerUB(Model, a2, eSize, cPower, p1, w1, w2, n2New);

If &Model="MIP" Then BFun=Probit(1-&a2);

If &Model="MSP" Then BFun=Probit(1-max(0.0000001,&a2-&p1));

If &Model="MPP" Then BFun=Probit(1-&a2/&p1);

If &Model="LW" Then BFun=(Probit(1-&a2)- &w1*Probit(1-&p1))/&w2;

&n2New=2*((BFun-Probit(1-&cPower))/&eSize)**2;

%Mend nByCPowerUB;
Example of Sample-Size Re-Estimation by Conditional Power

Data cPow; keep n2New;
%nByCPowerUB("MSP", 0.205, 0.21, 0.9, 0.1, 0.707, 0.707, n2New);
Run; Proc Print; Run;

Data cPow; keep n2New;
%nByCPowerUB("MPP", 0.0043, 0.21, 0.9, 0.1, 0.707, 0.707, n2New);
Run; Proc Print; Run;

Data cPow; keep n2New;
%nByCPowerUB("LW", 0.0226, 0.21, 0.9, 0.1, 0.707, 0.707, n2New);
Run; Proc Print; Run;
SAS Macro: Adaptive Design with Sample-Size Re-estimation

%Macro TwoArmNStgAdpDsg(nSims=1000000, nStgs=2, ux=0, uy=1, NId=0, NItype="FIXED", EP="normal", Model="MSP", Nadj="Y", cPower=0.9, DuHa=1, Nmax=300, N2min = 150, nMinIcr=1, sigma=3, tAcr=10, tStd=24, nRatio=1);

DATA NStgAdpDsg; Set dInput;

KE Model power Aveux Aveuy AveTotalN FSP1-FSP&nStgs ESP1-ESP&nStgs;
Array alpha{&nStgs}; Array beta{&nStgs}; Array FSP{&nStgs}; Array ESP{&nStgs};
Model=&Model; cPower=&cPower; nRatio=&nRatio; NId=&NId; NItype=&NItype; N2min=&N2min; nStgs=&nStgs; sigma=&sigma;
power=0; AveTotalN=0; Aveux=0; Aveuy=0; ux=&ux; uy=&uy;
cumN=0; Do i=1 To nStgs; cumN=cumN+Ns{i}; End;
sumWs{k}=sqrt(sumWs{k});
* Calculate the standard deviation, sigma for different endpoints *;
If &EP="normal" Then Do sigmax=&sigma; sigmay=&sigma; End;
If &EP="binary" Then Do sigmax=Sqrt(&ux*(1-&ux)); sigmay=Sqrt(&uy*(1-&uy)); End;
If &EP="survival" Then Do sigmax=ux*sqrt(1+exp(-ux*&tStd)*(1-exp(ux*&tAcr))/(&tAcr*ux)); sigmay=uy*sqrt(1+exp(-uy*&tStd)*(1-exp(uy*&tAcr))/(&tAcr*uy)); End;
If NItype="PCT" Then sigmax=(1-NId)*sigmax;
Do i=1 To nStgs; FSP{i}=0; ESP{i}=0; End;
Do iSim=1 to &nSims;
        ThisNx=0; ThisNy=0; Thisux=0; Thisuy=0;
        Do i=1 To nStgs; TSc{i}=0; End;
        TS=0; If &Model="MPP" Then TS=1;
        If &Model="UWZ" Then Do; eSize=&DuHa/(abs(uyObs-uxObs)+0.0000001); nFinal=min(&Nmax, max(N0,eSize*Abs(eSize)*N0)); %nByCPowerUB(Model, alpha{2}, eSize, cPower, TS, ws{1}, ws{2}, n2New);
        End; If TS>beta{i} Then Do; FSP{i}=FSP{i}+1/&nSims; Goto Jump; End;
        Else If TS<=alpha{i} then do; Power=Power+1/&nSims; ESP{i}=ESP{i}+1/&nSims; Goto Jump; End;
        Else If TS==alpha{i} then do; Power=Power+1/&nSims; ESP{i}=ESP{i}+1/&nSims; Goto Jump; End;
End;
        End; End;
        Jump: Aveux=Aveux+Thisux/ThisNx/&nSims; Aveuy=Aveuy+Thisuy/ThisNy/&nSims; AveTotalN=AveTotalN+(ThisNx+ThisNy)/&nSims; End;
        Output; Run; Proc Print; Run;
%Mend TwoArmNStgAdpDsg;
SAS Macro Calls for MI Trial

Data dInput;
Array Ns{2} (150, 150); Array alpha{2} (0,0.225); Array beta{2} (0.2,0.225); Array Ws{2} (1,1);

%TwoArmNSTgAdpDsg(nStgs=2, ux=0.11, uy=0.22, EP="binary", Model="MSP", Nadj="N");
%TwoArmNSTgAdpDsg(nStgs=2, ux=0.11, uy=0.22, EP="binary", Model="MSP", Nadj="Y", cPower=0.95, DuHa=0.11, Nmax=500, N2min=150, nMinICR=50);
%TwoArmNSTgAdpDsg(nStgs=2, ux=0.14, uy=0.22, EP="binary", Model="MSP", Nadj="N");
%TwoArmNSTgAdpDsg(nStgs=2, ux=0.14, uy=0.22, EP="binary", Model="MSP", Nadj="Y", cPower=0.95, DuHa=0.11, Nmax=500, N2min=150, nMinICR=50);
%TwoArmNSTgAdpDsg(nStgs=2, ux=0.22, uy=0.22, EP="binary", Model="MSP", Nadj="N");
%TwoArmNSTgAdpDsg(nStgs=2, ux=0.22, uy=0.22, EP="binary", Model="MSP", Nadj="Y", cPower=0.95, DuHa=0.11, Nmax=500, N2min=150, nMinICR=50);
Run;
SAS Macro for Conditional Power

%Macro ConPower(EP="normal", Model="MSP", alpha2=0.205, n2=100, p1=0.8, w1=1, w2=1);

  data cPower;
  a2=&alpha2; Model=&Model;
  u=(&ux+&uy)/2;
  w1=&w1/sqrt(&w1**2+&w2**2);
  w2=&w2/sqrt(&w1**2+&w2**2);
  If &EP="normal" Then sigma=σ;
  If &EP="binary" Then sigma=\((u^2-u)^0.5\);
  eSize=(\&uy-\&ux)/sigma;
  If Model="MIP" Then BFun=Probit(1-a2);
  If Model="MSP" Then BFun=Probit(1-max(0.0000001,a2-&p1));
  If Model="MPP" Then BFun=Probit(1-a2/&p1);
  If Model="LW" Then BFun=(Probit(1-a2)- w1*Probit(1-&p1))/w2;
  cPower=1-ProbNorm(BFun-eSize*sqrt(&n2/2));
  Run;
  Proc Print data=cPower; Run;

%Mend ConPower;
SAS Macro Calls for Conditional Power

%ConPower(EP="binary", Model="MSP", alpha2=0.2050, ux=0.2, uy=0.4, n2=100, p1=0.1);
%ConPower(EP="binary", Model="MIP", alpha2=0.0201, ux=0.2, uy=0.4, n2=100, p1=0.1);
%ConPower(EP="binary", Model="MPP", alpha2=0.0043, ux=0.2, uy=0.4, n2=100, p1=0.1);
%ConPower(EP="binary", Model="LW", alpha2=0.0226, ux=0.2, uy=0.4, n2=100, p1=0.1, w1=1, w2=1);
%ConPower(EP="normal", Model="MSP", alpha2=0.2050, ux=0.2, uy=0.4, sigma=1, n2=200, p1=0.1);
%ConPower(EP="normal", Model="MIP", alpha2=0.0201, ux=0.2, uy=0.4, sigma=1, n2=200, p1=0.1);
%ConPower(EP="normal", Model="MPP", alpha2=0.0043, ux=0.2, uy=0.4, sigma=1, n2=200, p1=0.1);
%ConPower(EP="normal", Model="LW", alpha2=0.0226, ux=0.2, uy=0.4, sigma=1, n2=200, p1=0.1, w1=1, w2=1);
SAS Macro: Two-Stage Drop-Loser Adaptive Design

%Macro DrpLsrNRst(nSims=100000, CntlType="strong", nArms=5, alpha=0.025, beta=0.2, NId=0, cPower=0.9, nInterim=50, Nmax=150, nAdj="Y", alpha1=0.01, beta1=0.15, alpha2=0.1871, EP="normal", sigma=1, tStd=24, tAcr=10);
Data DrpLsrNRst; Set dInput;    Keep FSP ESP AveN Power cPower Nmax;
Array us{&nArms}; Array u1{&nArms};    Array u2{&nArms}; Array cs{&nArms};
seedx=1736; seedy=6214;    alpha=&alpha; NId=&NId; Nmax=&Nmax; nArms=&nArms;
if &EP="mean" Then sigma=\sigma
If &EP="binary" Then sigma=(us1*(1-us{1}))*0.5;
If &EP="survival" Then sigma=us{1}*(1-exp(us{1}*\tStd)*(1-exp(us{1})*\tAcr))/(\tAcr*us{1}))**(0.5); Do isim=1 to &nSims;
TotalN=nArms*n1; uMax=us{1}; Cntrst=0;
SumSqc=0;
Do i=1 To nArms;
u1{i}=Rannor(seedx)*sigma/Sqrt(n1)+us{i};
If u1{i}>uMax Then Do uMax=u1{i}; iMax=i; End;
Cntrst=Cntrst+cs{i}*u1{i}; SumSqc=SumSqc+cs{i}*cs{i}; End;
Z1 = Cntrst*Sqrt(n1)/Sqrt(SumSqc)/sigma; p1=1-ProbNorm(Z1);
If &CntlType="strong" Then p1=(nArms-1)*(1-ProbNorm((uMax-us{1})*sigma/Sqrt(n1/2)));
If p1>&beta1 Then FSP=FSP+1/&nSims;
If p1<=&alpha1 Then do; Power=Power+1/&nSims; ESP=ESP+1/&nSims; End;
if iMax=1 Then Goto myJump; If p1>&alpha1 and p1<=&beta1 Then do;
BF=Prob(1-max(0, &alpha2-p1)-Prob(1-cPower); n2=2*(sigma/(u1{iMax}-u1{1}))*BF**2; nFinal=min(n1+n2, Nmax);
If &nAdj="N" Then nFinal=Nmax; If nFinal>n1 Then do;
TotalN=2*(nFinal-n1)+nArms*n1;
u2{1}=Rannor(seedx)*sigma/Sqrt(nFinal-n1)+us{1};
u2{iMax}=Rannor(seedy)*sigma/Sqrt(nFinal-n1)+us{Max};
T2=(u2{iMax}-u2{1}+NId)*Sqrt(nFinal-n1)/2**0.5/sigma; p2=1-ProbNorm(T2); TS2=p1+p2;
If <TS2<=&alpha2 Then Power=Power+1/&nSims; End; End;
myJump:
AveN=AveN+TotalN/&nSims; End; Output; Run;
Proc Print Data=DrpLsrNRst (obs=1); Run;
%Mend DrpLsrNRst;
SAS Macro Calls for Seamless Design

Title "Simulations of Drop-loser Design under Ho";
Data dlnPut;
Array us{5} (.06, .06, .06, .06, .06); Array cs{5} (-0.54, .12, .13, .14, .15);

%DrpLsrNRst(CntlType="strong", nArms=5, alpha=0.025, beta=0.2, NId=0, cPower=0.9, nInterim=50, Nmax=100, nAdj="Y", alpha1=0.01, beta1=0.15, alpha2=0.1871, EP="normal", sigma=0.18);

Run;

Title "Simulations of Drop-loser Design under Ha";
Data dlnPut;
Array us{5} (.06, .12, .13, .14, .15); Array cs{5} (-0.54, .12, .13, .14, .15);

%DrpLsrNRst(CntlType="strong", nArms=5, alpha=0.025, beta=0.2, NId=0, cPower=0.9, nInterim=50, Nmax=100, nAdj="Y", alpha1=0.01, beta1=0.15, alpha2=0.1871, EP="normal", sigma=0.18);

Run;
Breakthrough science. Breakthrough medicine.℠