Adaptive clinical trial designs incorporating mid-trial sample size adjustment

Jack Bowden and Adrian Mander

*MRC Biostatistics Unit, Cambridge*

Kings College London, March 2013
I will sketch an adaptive trial design proposed as an alternative to a ‘traditional’ (fixed sample size) trial

An adaptive, unblinded two stage trial incorporating sample size adjustment
- Use all of the data collected at the interim period to decide how many additional patients needed
- Crucially use the treatment effect estimate (hence ‘unblinded’)

There has been a very low uptake of these methods in practice

Fears include: scientific validity, type I error inflation, use of non-sufficient statistics, statistical complexity etc...

We are trying to change this!
The traditional fixed sample size approach

- Assume observations in experimental treatment group ‘X’ and standard therapy group ‘Y’ are normally distributed with means $\mu_X$ and $\mu_Y$ and have a common variance of $\sigma^2$.

- **Parameter** of interest is $\delta = \frac{\mu_X - \mu_Y}{\sigma}$. $H_0 : \delta \leq 0$.

- **Fixed design**: $n$ patients per arm

- Choose $n$ via

  $$n = \frac{2}{\delta^2} (Z_\alpha + Z_\beta)^2, \quad Z_u = (1 - u)'th \text{ normal quantile}$$

- e.g. If $\delta = 0.35$, $\alpha = 0.025$ and $\beta = 0.2$: $n = 129$ patients per arm.

- Estimation and inference for $\delta$ at **end of trial** via:

  $$\hat{\delta} = \frac{\bar{x} - \bar{y}}{\sigma} \text{ and } z = \frac{\hat{\delta}}{\sqrt{2/n}}$$
Sample size required for 80% power as a function of $\delta$.

- If $\delta \ll 0.35$ then substantially more than 129 people needed.
- If $\delta \gg 0.35$ then trial a waste of resources.
A general ‘adaptive’ two-stage design strategy

- Suppose instead \( n_1 (\ll n) \) subjects \textit{initially} recruited, giving:
  \[
  \hat{\delta}_1 = \frac{\bar{x} - \bar{y}}{\sigma}, \quad \& \quad \text{test statistic} \quad z_1
  \]
ad at the \textit{interim analysis}. Then, if:

\[
\begin{align*}
  & z_1 > k : \text{Stop the trial for \textbf{efficacy}} \\
  & z_1 < h : \text{Stop the trial for \textbf{futility}} \\
  & h \leq z_1 \leq k : \text{Recruit further} \ n_2 \ \text{patients using} \ z_1 \ \text{as a guide}
\end{align*}
\]

- Declare efficacy at stage 2 if final test statistic, \( z \) is \( \geq C \)

- Adaptive design must still control type I and II errors
- Dictates choice of design parameters and additional sample size rule \( n_2(z_1) ? \)
Standard approach  (Proschan and Hunsberger, 1995)

- Find $h$, $k$, $n_2(z_1)$ and $C(z_1)$ such that:
  1. $Pr(\text{declare efficacy at stage 1 or stage 2}|\delta = 0) = \alpha$
  2. $Pr(\text{declare efficacy at stage 2}|\delta = \hat{\delta}_1) = 1 - \beta_1$

- Complicated optimisation problem
- Little freedom to choose stage 1 thresholds $h$ and $k$
- Critical value $C(z_1)$ dependent on $z_1$
- Design can’t be pre-specified
- This includes interim sample size adjustment formula
Choosing $h$, $k$, $C$ via the LSW method: Li et al. (2002, 2005)

- Choose $h$ & $k$ almost freely, (e.g. in terms of $\hat{\delta}_1$ or p-value of $z_1$)
- Choose overall type I error $\alpha$ and conditional power $1-\beta_1$
- Find $C$ such that:
  
  1. $Pr(\text{declare efficacy at stage 1 or stage 2}|\delta = 0) = \alpha$
  2. $Pr(\text{declare efficacy at stage 2}|\delta = \hat{\delta}_1) \geq 1 - \beta_1$

Given that, if trial goes to stage 2:

$$n_1 + n_2(z_1) = \frac{2(C + Z_{\beta_1})^2}{\hat{\delta}_1^2}$$

- A very simple method! (one parameter optimisation)
- Critical value $C$ independent of $z_1$.
  
  - Whole design & analysis can be specified in advance
Example

- If \( n_1 = 50 \), \( \alpha = 2.5\% \) and conditional power = 80%.
- \((h=1,k=2.74): = (0.16,0.003) \) p-value scale, \((0.20,0.55) \) \( \hat{\delta}_1 \) scale.
- \( C \) is found to be \( \approx 1.92 \) (less than \( Z_{\alpha} = 1.96 \))
A minor modification of Li et al’s approach

- Maximum sample size $\approx 3 \times$ Fixed sample size
- Too large to be attractive?

- Want the trial to continue if (say) $0.2 < \hat{\delta}_1 < 0.55$ but with realistic maximum sample size
- That is, require $n_2(z_1) \leq n_{\max}$, for some $n_{\max}$
- Want $C$ still to be independent of trial data:

- Call this modified LSW approach
Example contd.: \( n_{\text{max}} = 90, \ C = 1.93, \ Z_{\beta} = 0.8. \)
Comparison of fixed and adaptive designs

- How do the fixed design and the two adaptive designs compare as a function of the true treatment effect $\delta$?
  - Expected sample size?
  - Overall power to reject $H_0$?

**Fixed design:** $n = 129$, $\alpha = 2.5\%$ $1 - \beta = 80\%$, $\delta_{H_1} = 0.35$

**Adaptive LSW designs:** $n_1 = 50$, $\alpha = 2.5\%$ $1 - \beta_1 = 80\%$
  - $h = 1$, $k = 2.74$

- Additionally for modified LSW: $n_{max} = 90$. 
**Expected sample size**

- **Modified LSW design** has the smallest expected sample size
Overall power to declare treatment effective

- **Fixed design**: $80\%$ power at $\delta=0.35$
- **Standard LSW design**: $72\%$ power at $\delta=0.35$
- **Modified LSW design**: $68\%$ power at $\delta=0.35$
Criticisms of the LSW method

- Allows user to specify **conditional** power level desired at stage 2 and leaves choice of $n_1$, $h$ and $k$ open
- Can identify designs with small expected sample size compared to fixed design
- However, it generally has a lower **overall** power than fixed design
- **Non-standard** critical threshold $C$ at stage 2 which is different to nominal $\alpha$ uncomfortable for trialists

  e.g. LSW design could suggest to reject null as $Z > 1.92$ but if $Z < 1.96$ standard $\alpha$-level analysis would not

- To address these concerns, propose a ‘**reverse implementation**’
Reverse implementation of the standard LSW design

1. Identify a fixed sample size design with type I error $\alpha$ and power $1-\beta$ at $\delta = \delta_{H_1}$

2. Find all possible adaptive designs $(h, k, n_1)$ where:
   1. Critical threshold $C = Z_{\alpha}$
   2. $n_1$ is the minimum stage 1 sample size that sets the unconditional power at $\delta = \delta_{H_1}$ equal to $1-\beta$

Special software written for this purpose
Possible design space: \( \alpha = 0.025 \) and \( 1 - \beta = 0.8 \)

- **Black Line** = all designs satisfying criteria
- **Red point**: minimum conditional power \( 1 - \beta_1 = \) overall power \( 1 - \beta \)
  - Occurs at (approximately) \( h = 1.14, k = 2.24 \) and \( n_1 = 70 \)
- Use a slightly different algorithm/rationale for reverse implementation of the modified LSW design
- Don’t go into details, results fairly similar
- 4 designs to compare with fixed design
  - $\alpha = 0.025, 1 - \beta = 0.8, n = 129, \delta_H_1 = 0.35$

<table>
<thead>
<tr>
<th>design</th>
<th>$h$</th>
<th>$k$</th>
<th>$1 - \beta_1$</th>
<th>$C$</th>
<th>$\alpha$</th>
<th>$n_1$</th>
<th>$n_{max}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard implementation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. LSW</td>
<td>1</td>
<td>2.74</td>
<td>0.8</td>
<td>1.92</td>
<td>0.025</td>
<td>50</td>
<td>-</td>
</tr>
<tr>
<td>2. modified LSW</td>
<td>1</td>
<td>2.74</td>
<td>0.8</td>
<td>1.93</td>
<td>0.025</td>
<td>50</td>
<td>90</td>
</tr>
<tr>
<td><strong>Reverse implementation (red dot designs)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. LSW</td>
<td>1.14</td>
<td>2.24</td>
<td>0.8</td>
<td>1.96</td>
<td>0.025</td>
<td>70</td>
<td>-</td>
</tr>
<tr>
<td>4. modified LSW</td>
<td>1.08</td>
<td>2.32</td>
<td>0.8</td>
<td>1.96</td>
<td>0.025</td>
<td>71</td>
<td>121</td>
</tr>
</tbody>
</table>
Expected sample size of four designs

- Reverse implementation designs have higher expected sample size...
Overall power of four designs

- LSW approach
- Modified LSW approach
- Fixed design
- Standard implementation
- Reverse implementation

...but comparable power to fixed design
Summary

- The LSW method to unblinded sample size adjustment:
  - Can be fully specified **before** any recruitment begins;
  - Can be implemented by an independent, *non expert*, data monitoring committee
  - *Is* motivated by clear **decision framework** linking interim effect size with future sample size via a **simple**, familiar **formula**

- Our modifications mean that:
  - The trial’s data to be consistently analysed at the end using **standard methods**
  - The resulting adaptive trial is a **more transparent alternative** to a specific fixed design under consideration.
P-values for tests using a repeated significance test design. 

Mid-course sample size modification in clinical trials based on the observed treatment effect. 

A sample size adjustment procedure for clinical trials based on conditional power. 

Adaptive increase in sample size when interim results are promising: A practical guide with examples. 

Designed extension of studies based on conditional power. 

Estimation and confidence intervals for two-stage sample size flexible design with LSW likelihood approach. 
Further details......
Probability of conducting a ‘large’ trial (Adaptive designs only)

\[ P(\text{Efficacy stopping at stage 1}) \approx 20\% \text{ chance of stopping for } \textbf{futility}, \] 
\[ 10\% \text{ chance of stopping for } \textbf{efficacy} \text{ at stage 1 at } \delta = \delta_{H_1} = 0.35 \]

\[ \approx 25\% \text{ chance of requiring } \geq n_{max} \text{ for original LSW design.} \]
Conditional power function defined as:

\[ P(z > C|z_1, n_1, \delta) = 1 - hi \left( \frac{C \sqrt{n_1 + n_2(z_1)} - z_1 \sqrt{n_1} - n_2(z_1)\delta/\sqrt{2}}{\sqrt{n_2(z_1)}} \right) \]

Li et al’s method: Solve

\[ 1 - hi(h) - \alpha = \int_{h}^{k} hi \left[ \frac{C(C + Z_\beta) - u^2}{\sqrt{(C + Z_\beta)^2 - u^2}} \right] hi(u)du. \]

for \( C \) then find \( n_2(z_1) \) via

\[ n_2(z_1) = \left( \frac{(C + Z_\beta)^2}{z_1^2} - 1 \right) n_1, \quad \text{for } z_1 \in (h, k) \]
Modified LSW design **loses** conditional power when $n_2(z_1) = n_{\text{max}}$
Algorithm for reverse implementation modified LSW design

1. Identify a fixed sample size design with type I error $\alpha$ and power $1-\beta$ at $\delta = \delta_{H_1}$. Additionally fix the maximum value of $(n_1 + n_2(z_1))$, $n_{T_{max}}$ say, and set $C$ equal to $Z_\alpha$.

2. Given $n_{max} = n_{T_{max}} - n_1$, find the joint values of $(h,k,Z_{\beta_1},n_1,n_{max})$ such that:

   1. $(h,k,Z_{\beta_1},n_1,n_{max})$ are consistent with $\alpha$ and $C = Z_\alpha$
   2. $n_1$ is minimised given the joint values of $(h,k,Z_{\beta_1},n_{max})$
   3. The unconditional power at $\delta = \delta_{H_1}$ equals $1-\beta$