An adaptive-flavored group sequential design with connection to conditional error

Dong Xi, Paul Gallo
Novartis
Group sequential design

• During a trial, look at cumulative data multiple times after groups of patients complete assessment

• Decisions at interim analyses
  – Stop for futility, Stop for efficacy, Continue to the next analysis

• Ethics
  – Stop unnecessary exposure to toxic or ineffective treatments
  – Timely access to new effective treatments for all patients

• Economics
  – Reduce loss (futility)
  – Market new product early (efficacy)
Interim analyses for efficacy

• Interim analyses are usually conducted by Data Monitoring Committee (DMC)

• If interim data suggest that likelihood to be significant at the final analysis is very high, DMC may recommend stopping for efficacy
  – Sponsor may unblind data and claim success

• Challenge: Multiple chances to claim success
  – Inflating Type I error

• Opportunity: Cumulative data used for analyses
  – Test statistics are correlated among interim and final analyses
Group sequential design with efficacy interim analyses

• Test $H_0: \theta \leq 0$ against $H_1: \theta > 0$ with $K$ analyses
  - $K - 1$ interim and 1 final analysis

• Interim analyses are often planned according to sample size or number of events

• E.g., A trial with four analyses when 125, 250, 375, 500 patients complete assessment
  - $t_1 = \frac{125}{500} = 0.25$, $t_2 = 0.5$, $t_3 = 0.75$, $t_4 = 1$

• More formally, the timing of interim analyses is according to statistical information time (or fraction)
  $$t_k = \frac{\text{information up to analysis } k}{\text{information up to the final analysis}}$$

• Information: inverse of the variance of the estimate
  - Normal endpoint: proportional to the sample size
  - Survival endpoint: proportional to the number of events
Canonical distribution

- $Z_k, k = 1, \ldots, K$ are test statistics at the interim and final analyses
  - E.g., z test at information time $t_k$
- Under $H_0$, $Z_k$’s follow an asymptotic multivariate normal distribution (Jennison and Turnbull, 2000)
  - $E(Z_k) = 0$
  - $\text{Var}(Z_k) = 1$
  - $\text{Corr}(Z_j, Z_k) = \sqrt{t_j/t_k}$, $1 \leq j < k \leq K$
- E.g., A trial with four analyses at $t_k = k/4$, $k = 1, \ldots, 4$
  - $\text{Corr}(Z_1, Z_2, Z_3, Z_4) = \begin{bmatrix} 1 & \sqrt{1/2} & \sqrt{1/3} & \sqrt{1/4} \\ \sqrt{1/2} & 1 & \sqrt{2/3} & \sqrt{2/4} \\ \sqrt{1/3} & \sqrt{2/3} & 1 & \sqrt{3/4} \\ \sqrt{1/4} & \sqrt{2/4} & \sqrt{3/4} & 1 \end{bmatrix}$
Boundary for test statistics

- Let $c_k$ be the boundary value for analysis $k$
- $H_0$ is rejected if $Z_k \geq c_k$ for any $k = 1, \ldots, K$
- Type I error—crossing the boundary at any analysis
  \[
P \left( \bigcup_{k=1}^{K} Z_k \geq c_k \right) = \alpha
\]
- In addition, constraints on the relationship among $c_k$’s
  - Pocock (1977): $c_k = c_K$ for $k = 1, \ldots, K$
  - O’Brien-Fleming (1978): $c_k = c_K \times t_k^{-0.5}$
  - Wang and Tsiatis (1987): $c_k = c_K \times t_k^\Delta^{-0.5}$
  - These were originally proposed assuming equal-spaced analyses
Properties of the boundary approach for four equally-spaced analyses

- Easy to communicate for decision making
  - Constraints on the scale of test statistics
- Less flexible that the number and timing of analyses are pre-specified
  - Boundary value for the current analysis depends on the timing of future analyses
Error spending function

• An extension is to “spend” $\alpha$ across the interim and final analysis

• Error spending function $\alpha(t)$ specifies how much $\alpha$ is spent up to time $0 \leq t \leq 1$
  - $\alpha(0) = 0$
  - $\alpha(1) = \alpha$
  - $\alpha(t)$ is increasing in $t$

• Boundary values ($c_k$) can be obtained by equating
  - Error spent to analysis $k$: $\alpha(t_k) - \alpha(t_{k-1})$ and
  - Probability of crossing the boundary only at analysis $k$: $P(Z_1 < c_1, ..., Z_{k-1} < c_{k-1}, Z_k \geq c_k)$
Properties of error spending function approach

• Flexible in the number and timing of analyses
  – Boundary value for the current analysis does not depend on the timing of future analyses

• Less transparent to derive boundary values

• Lan-DeMets (1983): \( \alpha(t) = 2 - 2\Phi \left( \frac{z\alpha/2}{\sqrt{t}} \right) \)
  – Approximates O'Brien-Fleming
  – Derived using the Brownian motion

• Lan-DeMets (1983): \( \alpha(t) = \alpha \log[1 + (e - 1)t] \)
  – Approximates Pocock
  – Derived to approximate Pocock boundary numerically
More flexible error spending functions

• Gamma family (Hwang, Shih, DeCani, 1990)
  \[ \alpha_\gamma(t) = \alpha \frac{1 - e^{-\gamma t}}{1 - e^{-\gamma}} \]
  – \( \gamma = -4 \) approximates O’Brien-Fleming and \( \gamma = 1 \) approximates Pocock

• Power family (Jennison and Turnbull, 2000)
  \[ \alpha_\rho(t) = \alpha t^\rho \]
  – \( \rho = 3 \) approximates O’Brien-Fleming and \( \rho = 0.75 \) approximates Pocock

• Less interpretable in their mathematical forms

• No statistical meaning of the parameter except the above approximation
ACCOMPLISH trial

• ACCOMPLISH trial compared the effect in hypertensive patients at high risk for CV events between
  – benazepril / amlodipine (B/A) -- treatment
  – benazepril / hydrochlorothiazide (B/H) -- control

• Large outcome trial (11506 patients and 1642 events) included a group sequential scheme using the Lan-DeMets O’Brien-Fleming spending function

• At the first two interim analyses, stopping boundaries were not approached

• Third analysis (720 events, information time 0.44), a strong observed effect favoring the B/A arm ($z_3 = 3.18$); however, slightly below the O’Brien-Fleming criterion ($c_3 = 3.2$)

• DMC suspected that the boundary might well be crossed if the outcome of cases awaiting adjudication were known

• DMC suggested the next DMC meeting much sooner than had previously been planned (without conveying unblinded information)
Conditional error in O’Brien-Fleming boundary

• How to determine the stopping boundary for the next analysis?
  – Spending function allows changing the timing but it does not allow choosing a sooner time because of a positive treatment effect

• The conditional error approach in adaptive designs (Müller and Shäfer, 2001) is valid
  – Conditional on the observed effect at the current analysis, the probability of rejecting $H_0$ in the future
  – It should be the same between
    – Proceeding at analyses as originally planned
    – Proceeding with the expedited analysis as an additional analysis

• The conditional error rate is 0.5 and $c = 2.92$ for the expedited analysis
• If we plugged in the timing to the spending function, $c = 2.94$
Conditional error in O’Brien-Fleming boundary

• The closeness between these two boundary values is not a coincidence
• O’Brien-Fleming boundary has a conditional error of 0.5 (Jennison and Turnbull, 2000)
• At an interim analysis, if
  – the test statistic is equal to the boundary value and
  – the future data are generated under $H_0$,

then the expectation of test statistics for the next analysis will be equal to the boundary again
  – 50% chance to cross the boundary $\rightarrow$ conditional error = 0.5
Questions of interest

• How does the boundary behave if it has a conditional error of say, 0.6 or 0.4?

• Is there a family of group sequential designs that is connected to adaptive designs via conditional error?

• To derive the group sequential design, is it possible to connect the boundary and spending function in a closed form?
  – E.g., Brownian motion was used for Lan-DeMets O’Brien-Fleming spending function
Proposed design

• Consider an interim analysis at time $0 < t < 1$ and the final analysis at time $t_F = 1$
• Let $c_t$ and $c_{t_F}$ be the respective boundary values
• Suppose that at the interim analysis, the observed test statistic is right on the boundary, i.e., $Z_t = c_t$
• Under $H_0$, the conditional distribution of the final statistic is

$$Z_{t_F} | Z_t = c_t \sim N(c_t \sqrt{t}, 1 - t)$$
Conditional error

• The conditional error is

\[ P_{H_0}(Z_{t_F} \geq c_{t_F} | Z_t = c_t) = 1 - \Phi \left( \frac{c_{t_F} - c_t \sqrt{t}}{\sqrt{1 - t}} \right) \]

• Set \( P_{H_0}(Z_{t_F} \geq c_{t_F} | Z_t = c_t) = \gamma \) and we have

\[ c_t \sqrt{t} = c_{t_F} - z_\gamma \sqrt{1 - t}, \text{ where } z_\gamma = \Phi^{-1}(1 - \gamma) \]

• This provides the relationship between \( c_t \) and \( c_{t_F} \) such that the conditional error is \( \gamma \) at the interim analysis

• Stopping rule
  – In a group sequential design with the conditional error being \( \gamma \), stop at the interim analysis if the actual conditional error \( \geq \gamma \)
Brownian motion

• The standard Brownian motion process

\[ B(t) = Z_t \sqrt{t} \sim N(0, t) \] under \( H_0 \)

• Let the boundary value be \( b_t = c_t \sqrt{t} \) for \( B(t) \)
Spending function

• From $c_t \sqrt{t} = c_{t_F} - z_\gamma \sqrt{1 - t}$, we have $b_t = b_{t_F} - z_\gamma \sqrt{1 - t}$

• Let $t$ be the first exit time across the boundary $b_t$ from below

• Standard Brownian motion calculation gives spending function as the probability of exiting before time $t$

  $$P(\tau \leq t) = 2 - 2\Phi\left(\frac{z_{\alpha/2} - z_\gamma \sqrt{1 - t}}{\sqrt{t}}\right) = \alpha_\gamma(t)$$

• Lan-DeMets O’Brien-Fleming boundary is a special case
  - $\gamma = 0.5, z_\gamma = 0$
  - $c_t \sqrt{t} = c_{t_F}$
  - $\alpha(t) = 2 - 2\Phi\left(\frac{z_{\alpha/2}}{\sqrt{t}}\right)$
Properties for four-equally spaced analyses

- The larger $\gamma$ (conditional error) is, the more conservative early boundaries are (the less conservative the final boundary is)
- Large $\gamma$ means higher power to reject the null hypothesis, under the same sample size (e.g., O'Brien-Fleming is more powerful than Pocock)
Approximation of spending function approach

• The boundary approach provides the exact conditional error of $\gamma$

• The spending function approximates the $\gamma$ well
  – Within $\pm 0.1$ of $\gamma$ for all values investigated

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>Method</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
<th>$c_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8</td>
<td>Boundary approach</td>
<td>5.389</td>
<td>3.621</td>
<td>2.756</td>
<td>1.966</td>
</tr>
<tr>
<td></td>
<td>Actual conditional error</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>Spending function</td>
<td>5.826</td>
<td>3.845</td>
<td>2.863</td>
<td>1.963</td>
</tr>
<tr>
<td></td>
<td>Actual conditional error</td>
<td>0.864</td>
<td>0.857</td>
<td>0.849</td>
<td></td>
</tr>
</tbody>
</table>

• The spending function approach provides slightly more conservative boundaries for early analyses
Sometimes a perfect ending is not always what you expect it to be.

-- from the movie *A Perfect Ending*

What about $\gamma < 0.5$?
Valid range of $\gamma$

- A requirement on spending function to be increasing in $t$
- This means $0.5 \leq \gamma < \infty$
- This also means that the previous $\alpha_{\gamma}(t)$ does not allow less conservative interim boundaries than O’Brien-Fleming

- Could modify $\sqrt{1 - t}$ to allow a wide range and good operating characteristics

\[
\alpha_{\gamma}(t) = \begin{cases} 
2 - 2\Phi\left(\frac{z_{\alpha/2} - \gamma(1-t)}{\sqrt{t}}\right), & \text{if } 1 - \Phi\left(\frac{z_{\alpha/2}}{2}\right) \leq \gamma < 0.5 \\
2 - 2\Phi\left(\frac{z_{\alpha/2} - \gamma\sqrt{1-t}}{\sqrt{t}}\right), & \text{if } 0.5 \leq \gamma < 1
\end{cases}
\]

$-0.131 < \gamma < 1$ when $\alpha = 0.025$
Properties for four-equally spaced analyses

- The larger $\gamma$ (conditional error) is, the more conservative early boundaries are (the less conservative the final boundary is)
Approximation of spending function approach

• The spending function approximates the $\gamma$ well
  – Within ± 0.1 of $\gamma$ for all values investigated

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>Method</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
<th>$c_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>Spending function</td>
<td>3.016</td>
<td>2.350</td>
<td>2.208</td>
<td>2.224</td>
</tr>
<tr>
<td></td>
<td>Actual conditional error</td>
<td>0.204</td>
<td>0.213</td>
<td>0.267</td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>Spending function</td>
<td>3.516</td>
<td>2.574</td>
<td>2.239</td>
<td>2.097</td>
</tr>
<tr>
<td></td>
<td>Actual conditional error</td>
<td>0.348</td>
<td>0.348</td>
<td>0.376</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>Spending function</td>
<td>3.940</td>
<td>2.774</td>
<td>2.295</td>
<td>2.044</td>
</tr>
<tr>
<td></td>
<td>Actual conditional error</td>
<td>0.466</td>
<td>0.454</td>
<td>0.455</td>
<td></td>
</tr>
</tbody>
</table>

• The combined family has a wide range and good operating characteristics
Connection with adaptive designs

• In the conventional adaptive design framework using the conditional error
  – Start with a group sequential boundary (e.g., gamma family)
  – Then change the design at an interim analysis
  – Calculate the updated boundary so that the conditional error rate is maintained

• But the conditional error has almost no link with the gamma parameter

• Why don’t use the conditional error rate as a parameter in the group sequential boundary
  – If we set a conditional error rate (e.g., specify $\gamma$)
  – We will know the boundary follows $\alpha_{\gamma}(t)$ approximately to give a conditional error of $\gamma$
Conclusion

• A group sequential design that
  – Has a statistically interpretable parameter (conditional error)
  – Closed-form relationship between boundary values and the spending function

• The use of the Brownian motion introduces a general framework for further extensions

• The proposed group sequential design connects with the adaptive design via conditional error


• R program https://github.com/xidongdxi/gsd

26
Reference


• Jennison C, Turnbull BW. Group Sequential Methods with Applications to Clinical Trials. Chapman & Hall/CRC.


• Hwang IK, Shih WJ, De Cani JS. Group sequential designs using a family of type I error probability spending functions. Statist Med. 1990.

Thank you