Flexible Designs in Clinical Trial

Tatsuki Koyama, PhD

4/12/2013
A friendly investigator comes to your office with data on 12 mice each in the control and treatment groups.

```r
sort(cont1)
## [1]  10.4  16.7  18.5  19.7  19.8  20.2  20.2  21.4  21.9  24.3  26.7  29.1

sort(trea1)
## [1]  10.9  13.7  13.9  14.5  14.9  15.3  16.5  19.9  20.8  21.0  22.0  27.5

t.test(cont1, trea1)$p.value
## [1] 0.1153
```

“Not statistically significant.”
A friendly investigator comes back to your office with new data on 12 more mice each in the control and treatment groups.

\[
c(\text{sort}(\text{cont1}), \text{sort}(\text{cont2}))
\]

```r
## [1] 10.4 16.7 18.5 19.7 19.8 20.2 20.2 21.4 21.9 24.3 26.7 29.1 15.9 19.5 20.3
## [16] 21.4 21.7 22.1 22.9 23.5 23.6 24.9 25.9 27.3
```

\[
c(\text{sort}(\text{trea1}), \text{sort}(\text{trea2}))
\]

```r
## [1] 10.9 13.7 13.9 14.5 14.9 15.3 16.5 19.9 20.8 21.0 22.0 27.5 8.8 13.2 14.8
## [16] 15.4 15.5 16.1 16.1 21.3 21.4 21.7 25.5 28.7
```

\[
t.\text{test}(c(\text{cont1}, \text{cont2}), c(\text{trea1}, \text{trea2}))$p\text{.value}
\]

```r
## [1] 0.007601
```

“Statistically significant!”
Two-stage clinical trial

A clinical trial is being planned with $N_t = 200$ and $N_c = 200$.

- Let’s look at the data (break the blind) after we have $n_t = 100$ and $n_c = 100$.
  - If $p$-value < 0.025, reject $H_0$ and stop, otherwise continue to the end.
  - At the end $N_t = 200$ and $N_c = 200$, if $p$-value < 0.025, reject $H_0$.

- Stop for futility?
- Change sample size?
- Change hypothesis? treatment groups? primary endpoint? inclusion criteria?
Design of the second stage depends on the results of the first stage.

- Group sequential designs
- Sample size modification
- Internal pilot
- Phase II/III designs
- Superiority/Non-inferiority
Aside: Two-stage designs in academia?

Given a conventional single stage design (sample size $N$), I can probably come up with a two stage design with the same type I error rate and power with a smaller sample size\textsuperscript{1}.

\textsuperscript{1}expected (average) sample size. Maximum sample size is bigger than $N$. 
What do they say about adaptive designs?

**PhRMA (2006)** “A clinical study design that uses accumulating data to decide how to modify aspects of the study as it continues, without undermining the validity and integrity of the trial.” “... changes are made by design, and not on an ad hoc basis; therefore, adaptation is a design feature aimed to enhance the trial, not a remedy for inadequate planning.”

**EMA (2006)** “A study design is called ‘adaptive’ if statistical methodology allows the modification of a design element (e.g., sample-size, randomisation ratio, number of treatment arms) at an interim analysis with full control of type I error rate.” “... adaptive designs should not be seen as a means to alleviate the burden of rigorous planning of clinical trials.”

**FDA (2010)** “… adaptive design clinical study is defined as a study that includes a prospectively planned opportunity for modification of one or more aspects of the study design and hypotheses based on analysis of data (usually interim data) from subjects in the study.”

- “The objectives ... were to foster and facilitate wider usage and regulatory acceptance of adaptive designs to enhance clinical development, ...”

- “We endorse routine consideration of adaptive designs for clinical trials and encourage critical evaluation of when and where they can be appropriately deployed.”

- Their target: adaptive dose finding, seamless phase II/III designs, sample size reestimation
Issues

Statistical

- Preservation of the type I error rate (“have been solved broadly”).
- Treatment effect estimation and confidence interval coverage (“have not been fully resolved”).
- Encourage the exploration of both frequentist and Bayesian approaches.

Logistical

- Adaptive designs require rapid data collection.
- Require the integration of data capture, drug supply management, interactive communication system between patients/investigators/study sites/randomization center.
- Statistical analysis is complex (“customized software may be required”).

Procedural

- The interim data only monitored by a Data Monitoring Committee (DMC).
- DMC may be augmented with additional experts of adaptation.
- Sponsor representation may be required for critical business decisions.
- Unintentional data leak (“not unique to adaptive designs”).
Seamless phase IIb/III design

Some characteristics

- Placebo and 5 daily dose levels (10mg, 50mg, 150mg, 300mg, 450mg) of the test drug.
- Endpoint is a continuous variable.
- Pairwise comparisons with placebo.
- Type I error rate is 2.5%/5 each
  \[ N = 2\sigma^2(z_{\alpha/5} + z_\beta)/\delta^2 \approx 150. \]

Why adaptive design?

- Relatively quick outcome
- Relatively slow accrual
Seamless phase IIb/III design

- Stage 1 (phase IIb): \( n_1 = 50 \) per arm. Reduce the number of dose levels to 2 to 4 and determine sample size based on the prespecified rules.

- Sample size for stage 2: at most \( n_2 = 200 \), at least \( n_1 = 100 \). Based on conditional power.

\[ CP(n_2) = \text{Conditional power} = P[\text{Reject } H_0 \text{ in stage 2 | Stage 1 results is the truth}] \]

- If \( CP(100) > 90\% \) then \( n_2 = 100 \).
- If \( CP(100) < 90\% \) but \( CP(200) > 90\% \) then pick \( n_2 \) so that \( CP(n_2) = 90\% \).
- If \( 50\% < CP(200) < 90\% \) then \( n_2 = 200 \).
- If \( CP(200) < 50\% \) then no stage 2.

- If a dose level is to continue, all the dose levels above it will continue.

- May change based on the safety data.

- Conclusions at end of stage 2: take into account adaptive changes and multiple comparisons.

- Avoiding unintentional unblinding is a challenge.
Guidance for industry: Adaptive design clinical trials for drug and biologics.

- Two types of adaptive trials.
  - A & WC (adequate and well-controlled): intended to support marketing of a drug (pivotal studies - phase III). Type I error rate control is critical.
  - Exploratory: complement of A & WC.

- Encouraging with caution
  Recognizes its potential to improve efficacy of drug development, but raises concerns on their use in pivotal studies.

- Main concerns = type I error rate inflation and operational bias (validity, interpretability of results).
  All involved parties need to remain blinded
  Positive results may be difficult to interpret.

- **Well understood and accepted AD**
  - Adaptation of study eligibility criteria based on baseline data.
  - Blinded sample size re-assessment.
  - Group sequential designs.
  - Limited adaptations in statistical analysis plan without breaking the blind.
Group sequential designs

```r
(gd <- gsDesign(k = 4, test.type = 2, alpha = 0.025, beta = 0.1, delta0 = 0, delta1 = 0.25, n.fix = 1, sfu = "OF"))
```

## Symmetric two-sided group sequential design with
## 90% power and 2.5% Type I Error.
## Spending computations assume trial stops
## if a bound is crossed.

```r
## Sample
## Size
## Analysis Ratio* Z Nominal p Spend
## 1 0.256 4.05 0.0000 0.0000
## 2 0.511 2.86 0.0021 0.0021
## 3 0.767 2.34 0.0097 0.0083
## 4 1.022 2.02 0.0215 0.0145
## Total 0.0250
## ++ alpha spending:
## O'Brien-Fleming boundary
## * Sample size ratio compared to fixed design with no interim
## Boundary crossing probabilities and expected sample size
## assume any cross stops the trial
##
## Upper boundary (power or Type I Error)
## Analysis
## Theta 1 2 3 4 Total E{N}
```

Tatsuki Koyama, PhD
Flexible Designs in Clinical Trial
O’Brien Fleming boundary

Information relative to fixed sample design

Normal critical value

\[
\begin{array}{c|c|c|c|c}
\text{alpha spending} & 0.0000 & 0.0021 & 0.0083 & 0.0145 \\
\hline
\text{Normal critical value} & 4.0 & 2.0 & 0.0 & -2.0 \\
\end{array}
\]

Information relative to fixed sample design

Tatsuki Koyama, PhD
Flexible Designs in Clinical Trial
Less well understood AD

- Anything based on interim analysis of unblinded treatment effects.
  - Sample size modifications.
  - Group sequential designs?
- Adaptive dose selection: Complex (eg Bayesian) adaptation is recommended only for exploratory studies.
- Response adaptive randomization.
- Adaptation in non-inferiority trials.
- Type I error rate by simulation: “Simulation-based type I error rate control is controversial and not fully understood.”
- ...
- ...
- Seamless phase IIb/III design

FDA will probably require an “Adaptive design protocol” separately from the regular clinical trial protocol.
Study rational, justification for adaptation, anticipated impact on operating characteristics, plans to maintain study integrity, DMC’s roles, analytical derivation for type I error rate control.
Flexible Designs in Clinical Trial

Tatsuki Koyama, PhD

4/12/2013