Adaptive Sample Size Design

BASS 2008

Jialu Zhang
John Lawrence
Steve Bai
Center for Drug Evaluation and Research,
Food and Drug Administration

* The views expressed here are those of the authors and not necessarily those of the U.S. Food and Drug Administration

Outline

- Overview
- Motivating example
- Estimates, confidence intervals and more
- Interpretations
- Discussions and Summary

Introduction

- Sample size adjustment
  - Conditional power
    - Weighted test
    - Change of critical value
    - More

Example

- Phase III trial for a cardiovascular indication
- Initial planned total sample size: 4000 subjects
- Expected number of events is 1600
- Two interim analyses were conducted
- By the time for third interim analysis, 830 events were collected from 3900 subjects

Example (2)

- The sponsor proposed to adjust sample size at the third interim analysis based on paper by Chen, DeMets and Lan (CDL)
- Advantage of CDL method per Chen et al:
  - No inflation of type I error if the conditional power at interim analysis is larger than 50%
  - Ordinary test statistic can still be used

Example (3)

- Some concerns:
  - Is the traditional point estimate and confidence interval still valid under this design?
  - The sample size adjustment is near the end of the trial enrollment
  - Potential operational bias

CDL Method

- Chen, DeMets and Lan (2004) showed that if increasing sample size when the conditional power is greater than 50% at the interim look, the regular unweighted test statistic can still be used without inflating the type I error rate

- In this presentation
  - Sample size increase only
    - No need for sample size reduction since there is efficacy boundary for early stopping
  - Sample size increase only at the last interim look prior to the final analysis
    - More data and better estimate of treatment effect
  - No more interim analysis after the sample size increase
- First, consider one interim analysis during a trial

Notation

- N_0: originally planned total sample size
- N: new total sample size ($N \geq N_0$)
- t_1: information time where the interim analysis is conducted
- t^*: new maximum information time after sample size adjustment, $t^*=N/N_0$
- $Z(t_1)$: statistic computed at the interim analysis based on the first n subjects

Notation (2)

- $Z(t_2)$: final statistic based on the total N_0 subjects
- $Z(t^*)$: final statistic based on the total N subjects
- c_k: critical value for kth interim analysis, k=1,2

CDL Method

- Under the null hypothesis $H_0 : \theta = 0$, the change of the conditional type I error is

$$\Delta = P(Z(t^*) > c_2 \mid Z(t_1) = z, \theta = 0) - P(Z(t_2) > c_2 \mid Z(t_1) = z_1, \theta = 0)$$

$$= \Phi\left(\frac{\sqrt{t_1} z_1 - \sqrt{t^*} c_2}{\sqrt{t^* - t_1}}\right) - \Phi\left(\frac{\sqrt{t_1} z_1 - c_2}{\sqrt{1 - t_1}}\right)$$

- Chen, DeMets and Lan showed that when the conditional power is greater than 50%, this change in conditional type I error decreases when N increase.

CDL Method

- the conditional power is greater than 50% if and only if the observed test statistic at the interim analysis is at least $\sqrt{t_1}c_2$
- Under the condition $z_1 \geq \sqrt{t_1}c_2$, the change in type I error rate is

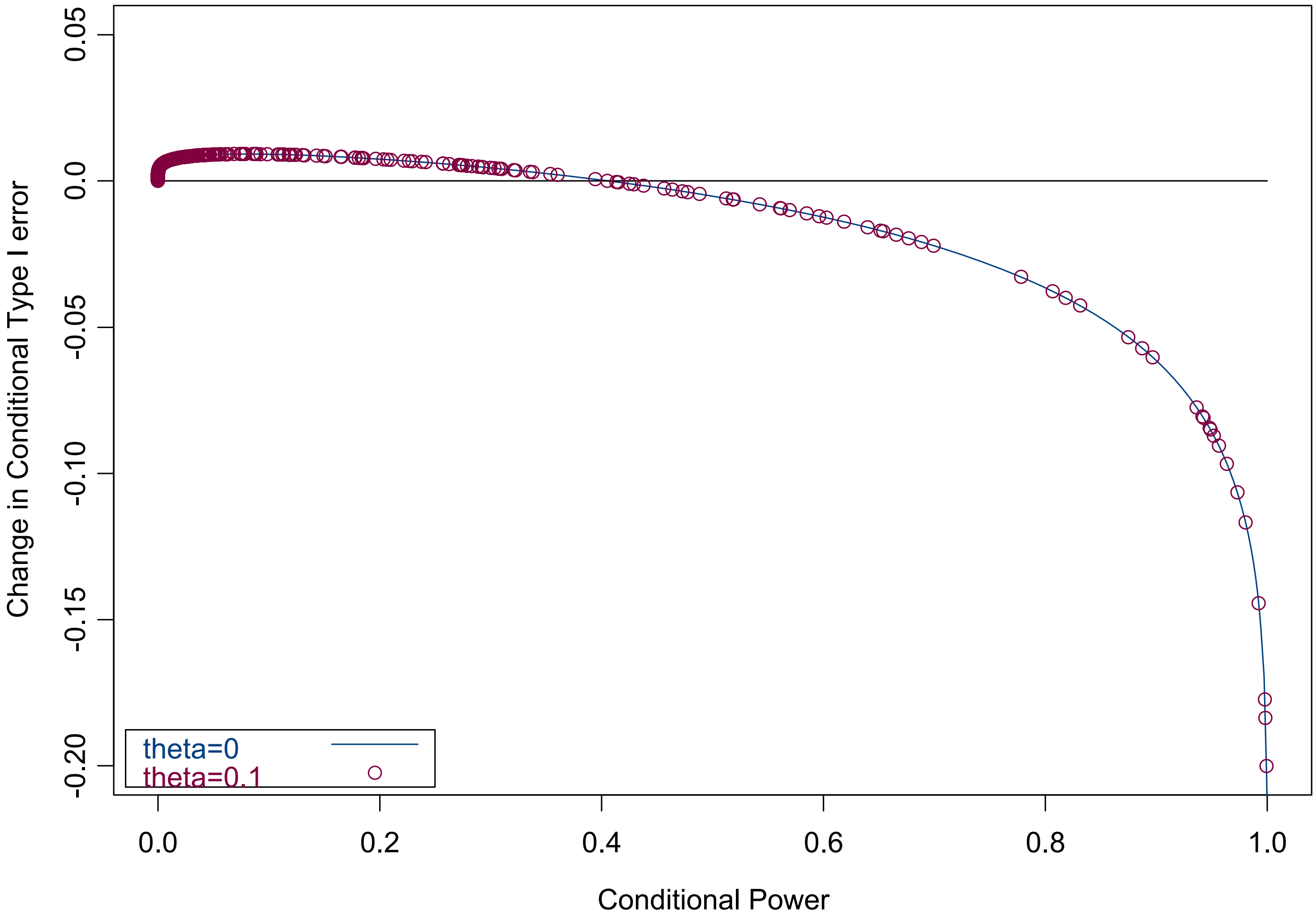
$$E\{P(Z(t^*) > c_2 \mid Z(t_1) = z_1, \theta = 0)\} - E\{P(Z(t_2) > c_2 \mid Z(t_1) = z_1, \theta = 0)\}$$
$$\leq 0$$

CDL Method

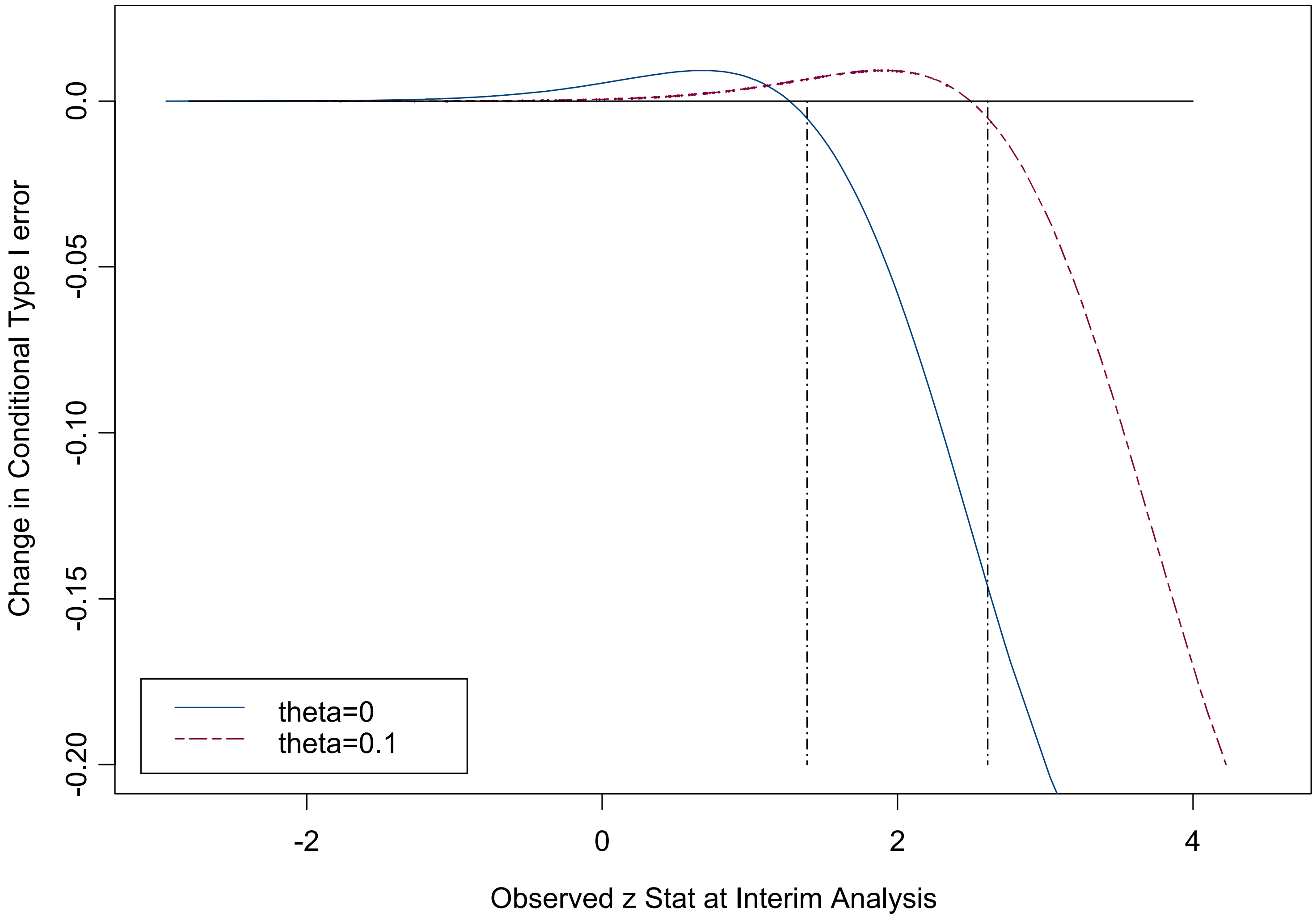
- Proposal: increase the sample size only when the first interim analysis test statistic is at least $\sqrt{t_1 c_2}$. The unweighted test statistics can still be used and the critical values remain the same.
- The simulation in the paper also showed that the type I error is controlled if the CDL method applies to a trial with two or more interim analyses

Extension

- Under null hypothesis $H_0 : \theta = \theta_0$
- Increasing sample size when conditional power is no less than 50% will not inflate type I error.
- CP≥50% is equivalent to $z_1 \geq \sqrt{t_1} c_2 + \sqrt{t_1} k \theta_0$ at the interim analysis
- k depends on the total sample size N_0 and θ is the treatment effect



0.05
0.0
-0.05
-0.10
-0.15
-0.20
Change in Conditional Type I error
0.0
0.2
0.4
0.6
0.8
1.0
Conditional Power
theta=0
theta=0.1



Change in Conditional Type I error
0.0
-0.05
-0.10
-0.15
-0.20
-2
0
2
4
Observed z Stat at Interim Analysis
theta=0
theta=0.1

- In fact, increasing sample size at the interim look will not inflate type I error as long as the following condition is satisfied

$$z_1 > z^* = \sqrt{n}\theta_0 + \frac{\sqrt{t^* - t_1} - \sqrt{t^*(1-t_1)}}{(\sqrt{t^* - t_1} - \sqrt{1-t_1})\sqrt{t_1}} c_2$$

- z* is always smaller than $\sqrt{t_1}c_2 + \sqrt{t_1}k\theta_0$ as long as N>N0

- Increasing the sample size may be conducted when the conditional power is less than 50% as long as the observed test statistic at interim look is larger than z*
- z* depends on the new information fraction t* and cannot be decided until the interim look.

Hypothesis Testing

- The total sample size will be increased if the conditional power exceeds ***L*** but below ***U***
  - In CDL method, ***L***=50%
- The sample size remains the same
  - if the conditional power exceeds ***U***, for example, ***U***=80%. The trial runs as expected
  - Or if the conditional power is below ***L***. The trial is not promising enough to increase the sample size
- We did not take futility into consideration

Hypothesis Testing (2)

- Reject null hypothesis $H_0 : \theta = \theta_0$ when
  - The total sample size is increased (***50%***≤CP<***U***) and the final test statistic $Z(t^*)$ based on the new sample size N exceeds critical value c2

$$C_1 : \frac{z - \sqrt{t_1} z_\alpha - \sqrt{t_1(1-t_1)} z_{1-\gamma}}{\sqrt{t_1} k} < \theta_0 \le \frac{z - \sqrt{t_1} z_\alpha}{\sqrt{t_1} k} \quad \text{and} \quad Z^{(N)} - \sqrt{t^*} k \theta_0 > z_\alpha$$

  - The total sample size remains the same (CP<***50%***, or CP≥***U***) and the final test statistic $Z(t_2)$ based on N_0 exceeds critical value c2

$$C_2 : \theta_0 > \frac{z - \sqrt{t_1} z_\alpha}{\sqrt{t_1} k} \quad \text{or} \quad \theta_0 \le \frac{z - \sqrt{t_1} z_\alpha - \sqrt{t_1(1-t_1)} z_{1-\gamma}}{\sqrt{t_1} k} \quad \text{and} \quad Z^{(N_0)} - k \theta_0 > z_\alpha$$

Hypothesis Testing (3)

- Reject hypothesis if one of the two conditions is satisfied:
  - Condition 1: $b_2 < \theta_0 \le b_1 \text{ and } \theta_0 < b_3$
  - Condition 2: $(\theta_0 > b_1 \text{ or } \theta_0 \le b_2) \text{ and } \theta_0 < b_4$
- b1, b2, b3, and b4 are values depending on the interim test statistic, final test statistic, and information time t1

$$b_1 = \frac{\mathrm{z} - \sqrt{\mathrm{t}_1} z_\alpha}{\sqrt{t_1} k}, b_2 = \frac{\mathrm{z} - \sqrt{\mathrm{t}_1} z_\alpha - \sqrt{t_1(1-t_1)} z_{1-\gamma}}{\sqrt{t_1} k}, b_3 = \frac{\mathrm{z}^{(\mathrm{N})} - z_\alpha}{\sqrt{t^* k}}, b_4 = \frac{\mathrm{z}^{(\mathrm{N}_0)} - z_\alpha}{k}$$

Confidence Interval

- Depending on the relative positions of b1, b2, b3 and b4, the one-sided confidence interval can be written as

$$
\begin{cases}
[b_4, \infty) & \text{if} \quad b_3 = b_4 \quad or \quad b_1 < \min(b_3, b_4) \quad or \quad b_2 \geq \max(b_3, b_4) \\
[b_3, \infty) & \text{if} \quad b_2 < \min(b_3, b_4) \quad and \quad b_1 \geq \max(b_3, b_4) \\
(b_2, \infty) & \text{if} \quad b_3 \leq b_2 < b_4 \leq b_1 \\
(b_1, \infty) & \text{if} \quad b_2 < b_4 \leq b_1 < b_3 \\
[b_3, b_1] \cup [b_4, \infty) & \text{if} \quad b_2 < b_3 \leq b_1 < b_4 \\
(b_2, b_1] \cup [b_4, \infty) & \text{if} \quad b_3 \leq b_2 < b_1 < b_4 \\
[b_4, b_2] \cup (b_1, \infty) & \text{if} \quad b_4 \leq b_2 < b_1 < b_3 \\
[b_4, b_2] \cup [b_3, \infty) & \text{if} \quad b_4 \leq b_2 < b_3 \leq b_1
\end{cases}
$$

True Treatment difference	theta	COND 1	COND 2	COND 3	COND 4	COND 5	COND 6	COND 7	COND 8
0	0	77.0	21.9	0.4	0.1	0.1	0.1	0.0	0.4
0.02	0	76.4	21.5	0.6	0.2	0.3	0.1	0.1	0.8
0.05	0	75.1	20.8	0.8	0.8	0.6	0.1	0.4	1.5
0.1	0	73.9	19.2	0.6	2.6	0.8	0.1	0.9	1.9
0.15	0	75.3	19.3	0.2	3.1	0.4	0.0	0.9	0.9
0.2	0	77.0	20.9	0.0	1.5	0.1	0.0	0.3	0.2
0.02	0.02	77.0	21.9	0.4	0.1	0.1	0.1	0.0	0.4
0.05	0.02	76.0	21.3	0.7	0.4	0.3	0.1	0.2	1.0
0.1	0.02	74.0	19.9	0.7	1.8	0.8	0.1	0.7	1.9
0.15	0.02	74.6	18.9	0.3	3.1	0.6	0.0	1.0	1.3
0.2	0.02	76.5	20.1	0.1	2.2	0.2	0.0	0.5	0.4
0.15	0.15	77.0	21.9	0.4	0.1	0.1	0.1	0.0	0.4
0.2	0.15	75.2	20.7	0.8	0.8	0.5	0.1	0.4	1.5

Based on 1,000,000 replications

Confidence Interval (2)

- Simplified confidence interval

$$\begin{cases} [b_4, \infty) & if \quad b_1 < \min(b_3, b_4) \quad or \quad b_2 \geq \max(b_3, b_4) \quad or \quad b_4 \leq b_3 \\ [b_3, \infty) & if \quad \left(b_2 < \min(b_3, b_4) \quad and \quad b_1 \geq \max(b_3, b_4)\right) \quad or \quad b_4 > b_3 \end{cases}$$

- The simplified confidence interval always has slightly larger probability coverage than the confidence interval shown previously

Point Estimate

- The point estimate can be derived by changing the significance level in the two one-sided confidence intervals so that the upper and the lower confidence intervals intersect on a single point
- The intersection of the two confidence intervals is either $\dfrac{z(t_2)}{k}$ or $\dfrac{z(t^*) + z(t_2)}{(1+\sqrt{t^*})k}$

Practicality Issue

- Disjoint confidence interval
- Two-sided confidence interval
- Statistical property of the estimator

Confidence Interval (3)

- The naive confidence interval may cause some problem.
- Naïve confidence interval:

$$\left[\frac{\sqrt{t_1}z_1+\sqrt{t^*-t_1}z_2-c_2}{t^*\sqrt{N_0}},\frac{\sqrt{t_1}z_1+\sqrt{t^*-t_1}z_2+c_2}{t^*\sqrt{N_0}}\right)$$

- z_1, z_2 are the before-interim and after-interim stage statistics

Example

- Initial planned sample size: 600 subjects (300 per group)
- This can detect an effect size of 0.25 with 86% power
- True treatment effect is zero
- A single interim analysis at 50% information fraction (300 patients) is planned
- If the conditional power is greater than 50% but less than 80% in the interim analysis, the sample size can be increased.

Example (2)

- The test statistic in the interim analysis is 1.47 and the corresponding conditional power is 57%
- So the sample size is increased to 1090 subjects
- The confidence intervals and treatment effect estimates are computed based on the 1090 subjects

Example (3)

- The one-sided naive confidence interval is [0.003, ∞)
- The new confidence interval is [-0.08, ∞)
- The naive confidence interval does not include zero

CHW method

- CHW method: weighted statistics
- The usual point estimate and naïve confidence interval are not valid
- Lawrence and Hung (2003) proposed a point estimate and a confidence interval based on CHW method

Comparison

- Compare CDL method with CHW method
  - Type I error
  - Power
  - Coverage probability

θ_0	CHW	CDL
0	**0.0251**	**0.0243**
0.02	**0.0251**	**0.0243**
0.05	**0.0250**	**0.0242**
0.1	**0.0249**	**0.0240**
0.15	**0.0249**	**0.0240**
0.25	**0.0251**	**0.0244**

Table 2: Estimated Type I error based on 1,000,000 simulation

True Treatment Difference	θ_0	CHW	CDL
0	0	0.025	0.024
0.02	0	0.045	0.043
0.05	0	0.095	0.092
0.1	0	0.251	0.246
0.15	0	0.485	0.480
0.25	0	0.884	0.883
0.02	0.02	0.025	0.024
0.05	0.05	0.025	0.024
0.1	0.05	0.094	0.091
0.15	0.05	0.252	0.247
0.25	0.05	0.720	0.718
0.1	0.1	0.025	0.024
0.15	0.15	0.025	0.024
0.25	0.15	0.252	0.247

Table 3: Power based on 1,000,000 simulation

True Treatment Difference	θ_0	CHW	CDL	CDL simplified	Naïve
0	0	0.975	0.976	0.977	0.970
0.02	0	0.975	0.975	0.976	0.970
0.05	0	0.975	0.975	0.975	0.968
0.1	0	0.975	0.975	0.975	0.967
0.15	0	0.975	0.975	0.975	0.969
0.25	0	0.975	0.975	0.975	0.974
0.05	0.05	0.975	0.976	0.977	0.971
0.1	0.05	0.975	0.975	0.975	0.968
0.15	0.05	0.975	0.975	0.975	0.967
0.25	0.05	0.975	0.975	0.975	0.972
0.15	0.15	0.975	0.976	0.977	0.971
0.25	0.15	0.975	0.975	0.975	0.967
0.25	0.25	0.975	0.976	0.977	0.970

Table 4: Coverage Probability of Confidence Interval

Summary

- CDL method appears to provide a way to increase sample size during the interim analysis and still use the regular test statistic and critical values
- Without a valid confidence interval and point estimate, it is difficult to put the adaptive design into practice
- Statistical validity is not good enough. It has to be sensible in practice

Reference

- Cui L., Hung H. M. J. and Wang S. J. (1999). Modification of Sample Size in Group Sequential Clinical Trials. *Biometrics*
- Lawrence J. and Hung H. M. J. (2003). Estimation and Confidence Intervals after Adjusting the Maximum Information. *Biometrical Journal*
- Chen Y. H. J., DeMets D. L. and Lan K. K. G. (2004). Increasing the Sample Size When the Unblinded Interim Result is Promising. *Statistics in Medicine*