A Comparison of Approaches for Unplanned Sample Size Changes in Phase II Clinical Trials

Molly Olson
molly.a.olson@vanderbilt.edu
Advisor: Tatsuki Koyama

Vanderbilt University
Department Of Biostatistics

June 13, 2017

Outline

Phase II Trials

- Phase I: Evaluate safety and dose
- Phase II: Evaluate initial effect to determine phase III trial
- Phase III: Evaluate efficacy

- Phase II - Two-stage
  - Mitigate the risk of exposure
  - Don't want to "waste" resources

Two-stage Phase II Trial

$H_0 : p \leq p_0$, $H_1 : p > p_1$

1. stage 1: n_1 patients are enrolled
   - $X_1 \sim \text{Binomial}(n_1, p) = \#$ of successes in first stage
2. If number of responses is r_1 or fewer, trial stopped for futility
3. Otherwise, stage 2: n_2 patients are enrolled ($n_t = n_1 + n_2$ total patients now)
   - $X_2 \sim \text{Binomial}(n_2, p) = \#$ of successes in second stage
   - $X_t = X_1 + X_2$
4. If number of responses is r_t or fewer, lack of efficacy concluded
5. Otherwise efficacy concluded

- $p_0, p_1, n_1, n_t, r_1, r_t, \alpha, \beta$ are design parameters
- n_1, n_t, r_1, r_t are chosen so that type I error rate is less than α and the type II error rate is less than β.

Types of Two-Stage Designs

- Simon introduced Optimal and Minimax criteria for good designs
  - Optimal minimizes the expected sample size under H_0
  - Minimax minimizes the maximum sample size
- Jung *et al.* introduced Admissible designs
  - Compromise between Optimal and Minimax
  - Similar maximum sample sizes as Minimax
  - Similar expected sample size under H_0 as Optimal
- Suppose $H_0 : p_0 \leq 0.25$, $H_1 : p_1 > 0.4$, $\alpha = 0.05$, $\beta = 0.2$

Design	n_t	n_1	r_1	r_t	EN_0	PET_0
Optimal	71	20	5	23	39.5	0.617
Minimax	60	51	16	20	52	0.886
Admissible	63	25	6	21	41.7	0.561

Deviation from the design

- Attain different enrollment than planned in first and/or second stage
- Why would we deviate?
  - Unanticipated recruitment speed
  - Unanticipated drop out rates
  - Delay in communication for multi-center trials
  - Ethical considerations
  - Shopping for sponsors
- Nice properties go out the window
- Currently, common practice is to treat attained sample size as planned
- Leads to invalid inference
- Hypothesis testing is not straightforward

Setting the Scene

- Goal is to make a decision
- How do we do this if our attained sample size is different than planned?
- P-value calculations are complicated - we don't consider these solutions
- Consider prespecified "redesigns" - recalculating critical values
- Primary focus on deviation in first stage

Deviation from Planned Sample Sizes in Second Stage

- Over-enrollment in first stage: perform interim analysis on the planned number of first stage, adjust testing procedure for attained second stage
- Under-enrollment: just wait
- Literature exists for point estimation, calculation of p-values when stage II differs (Review: Porcher *et al.*)
- P-values in two-stage trials depend on planned design and attained data, complicated when attained SS differ than planned [Koyama and Chen]

Koyama and Chen

- Koyama and Chen, StatMed, 2008
- Notation:
  - Planned design parameters: $n_1, n_t, n_2 = n_t - n_1 r_1, r_t, \alpha, \beta$.
  - Attained design parameters: $n_1, n_t^*, n_2^* = n_t^* - n_1, r_1, r_t^*, \alpha^*, \beta^*$
- Let first stage remained as planned and change testing procedure in stage II
- Calculate new critical value, r_t^*, by finding maximum integer s.t.

$$P[X_2^* \geq r_t^* | X_1 = x_1] \leq P]X_2 \geq r_t | X_1 = x_1$$

$$X_2^* \sim \text{Binomial}(n_2^*, p_0)$$

Koyama and Chen

- Results in controlled unconditional type I error rate - new CV gives more conservative conditional type I error rate
- New critical value depends on number of positive responses

Zeng *et al.*

- Zeng *et al.*, StatMed, 2015
- Attempts to maximize unconditional power while controlling type I error
- r_2^* new stage II critical value and $r_t^* \equiv r_2^* + x_1$
- Second stage CV is integer that maximizes unconditional power while subject to type I error $\leq \alpha$
- Theoretically possible, computationally difficult. No closed form solution.
- Propose normal approximation to ease computation of power
- Math (Lagrange multipliers, derivatives, substitution, searching over λs)
- Solve an ugly equation for r_2^*

Zeng *et al.*

$$\left(\frac{1}{p_0(1-p_0)}-\frac{1}{p_1(1-p_1)}\right)r_2^{*2}-\frac{2n_2^*(p_0-p_1)}{(1-p_0)(1-p_1)}r_2^*+\frac{n_2^{*2}(p_0-p_1)}{(1-p_0)(1-p_1)}-2n_2^*log\left(\frac{\lambda a(x_1)}{b(x_1)}\right)=0$$

$$a(x_1)=\binom{n_1}{x_1}p_0^{x_1}(1-p_0)^{n_1-x_1}$$

$$b(x_1)=\binom{n_1}{x_1}p_1^{x_1}(1-p_1)^{n_1-x_1} \tag{1}$$

λ is the Lagrange multiplier.

Deviation from planned sample sizes in first stage

- SWOG:
  - $\alpha = 0.05$, $\beta = 0.1$
  - Interim: $H_0 : p = p_1, H_1 : p < p_1$, stop if p-value is significant at 0.02-level
  - Stage II: $H_0 : p = p_0, H_1 : p > p_0$
- Green and Dahlberg, StatMed, 1992
  - Use SWOG, but use attained sample size
  - Test stage II at 0.055 level
- Unclear how to generalize
- Arbitrary and lacks theoretical justification [Li *et al.*]

Deviation from planned sample sizes in first stage

- Chen and Ng, StatMed, 1998
- Consider range of sample sizes
- Search these ranges for the Minimax or Optimal design that satisfy error constraints using the average PET and EN
- Limitation: attained sample sizes may fall outside of ranges
- Limitation: average probabilities rather then actual for attained SS

Chang *et al.*

- Chang *et al.* Biometrics & Biostatistics, 2015
- Recall: $n_1, n_t, r_1, r_t, p_0, p_1, \alpha, \beta$
- Notation: n_1^{**}, n_2^{**} - attained sample sizes
- Notation: s_1, s_t - new critical values
- Choose s_1 by first using β-spending function

$$\beta(m) = \begin{cases} \beta_1 m/n_1 & \text{if } m \leq n_1 \\ \beta_1 + (\beta - \beta_1)(m - n_1)/n_2 & \text{if } m > n_1 \end{cases}$$

- Integer s.t. type II error probability given s_1, n_1^{**} is closest to $\beta(n_1^{**})$
- Choose s_t s.t. type I error $\leq \alpha$

Olson and Koyama

- Select s_1 s.t. $\text{PET}_0^{**} \approx \text{PET}_0$
- Conservative
- Could have done $\text{PET}_0^{**} \leq \text{PET}_0$

Background: Likelihood

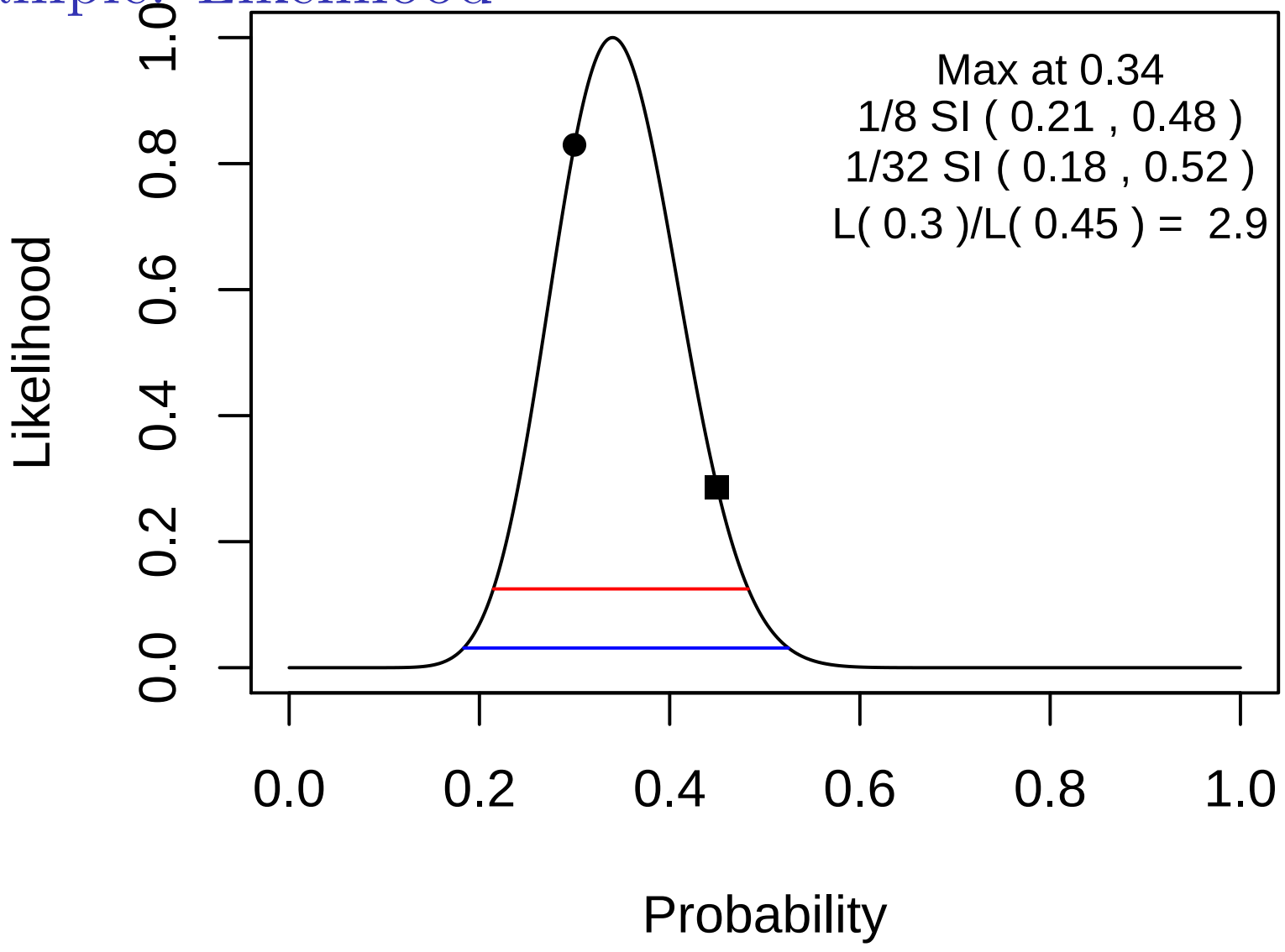
- Law of likelihood: "If $H_1 \Rightarrow P(X = x) = P_1(X)$, $H_2 \Rightarrow P(X = x) = P_2(X)$, then the observation X=x is evidence supporting H_1 over H_2 iff $P_1(X) > P_2(X)$. Likelihood ratio measures strength of evidence.
- Likelihood function:

$$\begin{aligned} \mathrm{L}_n(p) &= P(X|p, n) \\ &= \binom{n}{x} p^x (1-p)^{n-x} \\ &\propto p^x (1-p)^{n-x} \end{aligned} \tag{2}$$

- Likelihood ratio:

$$\mathrm{LR}_n = \frac{\mathrm{L}_n(p_1)}{\mathrm{L}_n(p_2)} \tag{3}$$

Example: Likelihood



Likelihood

- Can calculate probability of observing weak evidence, strong evidence, misleading evidence
- Universal bound of misleading evidence is $\leq 1/k$

Ayers and Blume

- Likelihood two-stage design
- Enroll n_1,
  - $1/k < LR_{n_1}$ ¡ k $\rightarrow$ second stage
  - $LR_{n_1} < 1/k \rightarrow$ stop for futility
  - $LR_{n_1} > k \rightarrow$ stop for efficacy
- Enroll n_2, $LR_{n_t} = LR_{n_1} LR_{n_2}$
  - $1/k < LR_{n_1}$ ¡ k $\rightarrow$ conclude weak
  - $LR_{n_1} < 1/k \rightarrow$ conclude futility
  - $LR_{n_1} > k \rightarrow$ conclude efficacy

Ayers and Blume

- Emulate conventional two-stage designs
- Notation: $k_{a_1}, k_{a_t}, k_{b_1}, k_{b_t}$
- Start with conventional two-stage design, set $k_{b_1}, k_{b_t} = \infty$, redefine k_{a_1}, k_{a_t}

$$s_1 = \frac{log(k_{a_1}) - n_1^{**} log\left(\frac{1-p_1}{1-p_0}\right)}{log\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right)}$$
$$s_t = \frac{log(k_{a_t}) - n_t^{**} log\left(\frac{1-p_1}{1-p_0}\right)}{log\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right)} \quad (4)$$

Ayers and Blume

- Can recalculate probability of weak, strong, and misleading evidence, $\mathrm{PET}_0, \mathrm{EN}_0$ under attained
- Minimizes average of error rates
- Type I error rate often below nominal rates

Comparison of methods

- Don't consider added cohorts
- Original total sample size ($n_t^{**} = n_t, n_2^{**} = n_t - n_1^{**}$)
- Original second stage sample size ($n_t^{**} = n_1^{**} + n_2$)

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item

Title

- Item