Optimal Adaptive vs. Optimal Group Sequential Designs

Keaven Anderson[1] and Qing Liu[2]

[1]Merck Research Laboratories

[2]Johnson & Johnson Pharmaceutical Research and Development

BASS XI

November 2, 2004

Outline

- Introductory example
- Adaptive design
  - Background
  - Conditional error functions
  - Optimization
- Optimal group sequential design
- Examples comparing optimized adaptive and optimized group sequential designs
- Summary and unresolved issues

What is adaptive design?

- For this paper:
  - Sample size adjustment at a single interim analysis
  - This is a very narrow definition!

Designs with Interim Analysis

Design	Analysis 1 n=	Analysis 2 n=	Analysis 3 n=
Fixed (no interim)	1051	NA	NA
2-stage group-sequential	440	1225	NA
3-stage group-sequential	306	679	1347
2-stage adaptive (LC)	430	679 to 1402	NA
2-stage adaptive (PH)	700	903 to 1866	NA

Tradeoffs

- Time and expense for an interim analysis
- Do final analysis at the ‘right sample size’ versus at a predictable sample size
- Partial knowledge of results from adjusted sample size (small sample size means ‘looks good!’)
- Statistical efficiency?
  - PH (whatever it is) looks bad!

Background

- Adaptive designs allow ‘redesign’ of trial based on interim data
  - have been criticized for not using sufficient statistics
- Tsiatis and Mehta, Biometrika, 2003 prove group-sequential can be used to improve on a given adaptive design
  - May require additional interim analyses compared to adaptive
- Jennison and Turnbull, Statistics in Medicine, 2003
  - Suggest that if group sequential design is planned for all contingencies, it will have better power and sample size characteristics across a broad range of treatment differences than an adaptive design
  - Use adaptive design basing adjustment on interim estimated treatment difference

Background

- Posch, Bauer and Brannath, Statistics in Medicine, 2003
  - Start with a given 2-stage group sequential design
  - Find optimal adaptive design with
    - same timing of interim analysis
    - same critical value at interim analysis
    - restricts maximum sample size
    - sets second stage sample size based on conditional power for a minimum treatment effect of interest
    - minimize expected sample size averaged over a fixed set of alternatives
    - These designs can improve average sample size over given group sequential designs
- Design strategy presented here is a generalization
  - optimize over a broad class of conditional error functions
    - replaces need to set maximum sample size
  - does not restrict timing or critical value of interim analysis
  - minimize expected value of loss function over a prior distribution for treatment effect size

Background

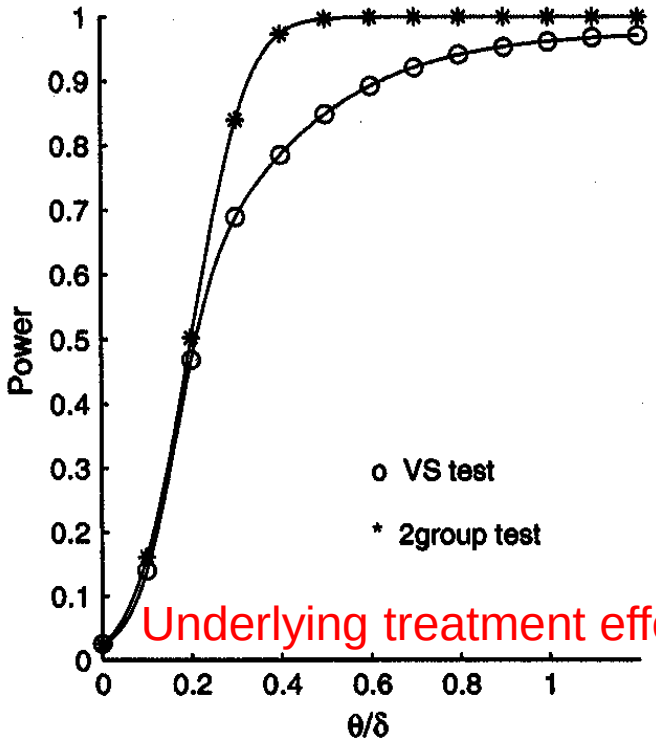
- Lokhnygina and Tsiatis, 2004
  - Fully optimized 2-stage adaptive designs
    - Not confined to a limited class as here
    - Otherwise, the optimization objective is the same
    - Dynamic programming algorithm for optimization

STATISTICS IN MEDICINE
Statist. Med. 2003; **22**:971–993 (DOI: 10.1002/sim.1457)

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

Christopher Jennison[1,*,†] and Bruce W. Turnbull[2]

[1]*Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, U.K.*
[2]*Department of Statistical Science, 227 Rhodes Hall, Cornell University, Ithaca, New York 14853-3801, U.S.A.*



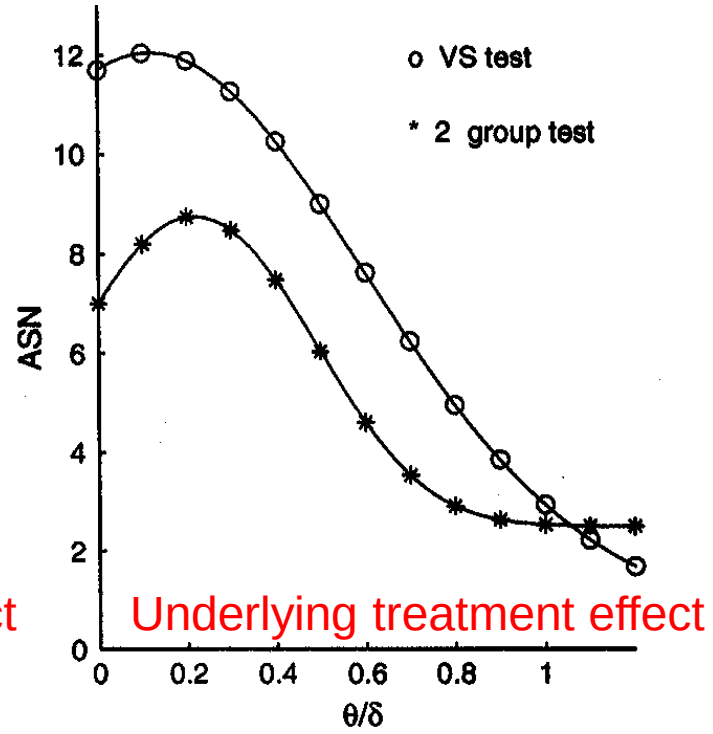


Figure 7. Power and ASN curves of the variance spending (VS) test and two-stage group sequential test. ASN scale is in multiples of the original fixed sample size, *n*.

Question

- Can we compare ‘best’ adaptive and group sequential designs, with each using a fixed number of interim analyses?
  - Optimal group sequential design problem solved by Barber and Jennison, Biometrika, 2002
  - Similarly optimized of 2-stage adaptive designs are presented here with either:
    - Interim estimated treatment effect (Proschan & Hunsberger, Bcs 1995)
    - Minimal treatment effect of interest (Liu & Chi, Bcs, 2001)

2-stage Adaptive design

2-Stage Adaptive Design

- Hypotheses
  - H_0: $\theta=0$
  - H_a: $\theta>0$
- Interim analysis with test statistic Z_1
  - Pre-specify $a_1<b_1$
  - If $Z_1>b_1$ then stop and reject H_0
  - If $Z_1<a_1$ then stop and 'accept' H_0
  - If $a_1 \leq Z_1 \leq b_1$, continue to stage 2
    - Estimate required stage 2 sample size
    - Compute critical value for stage 2 data

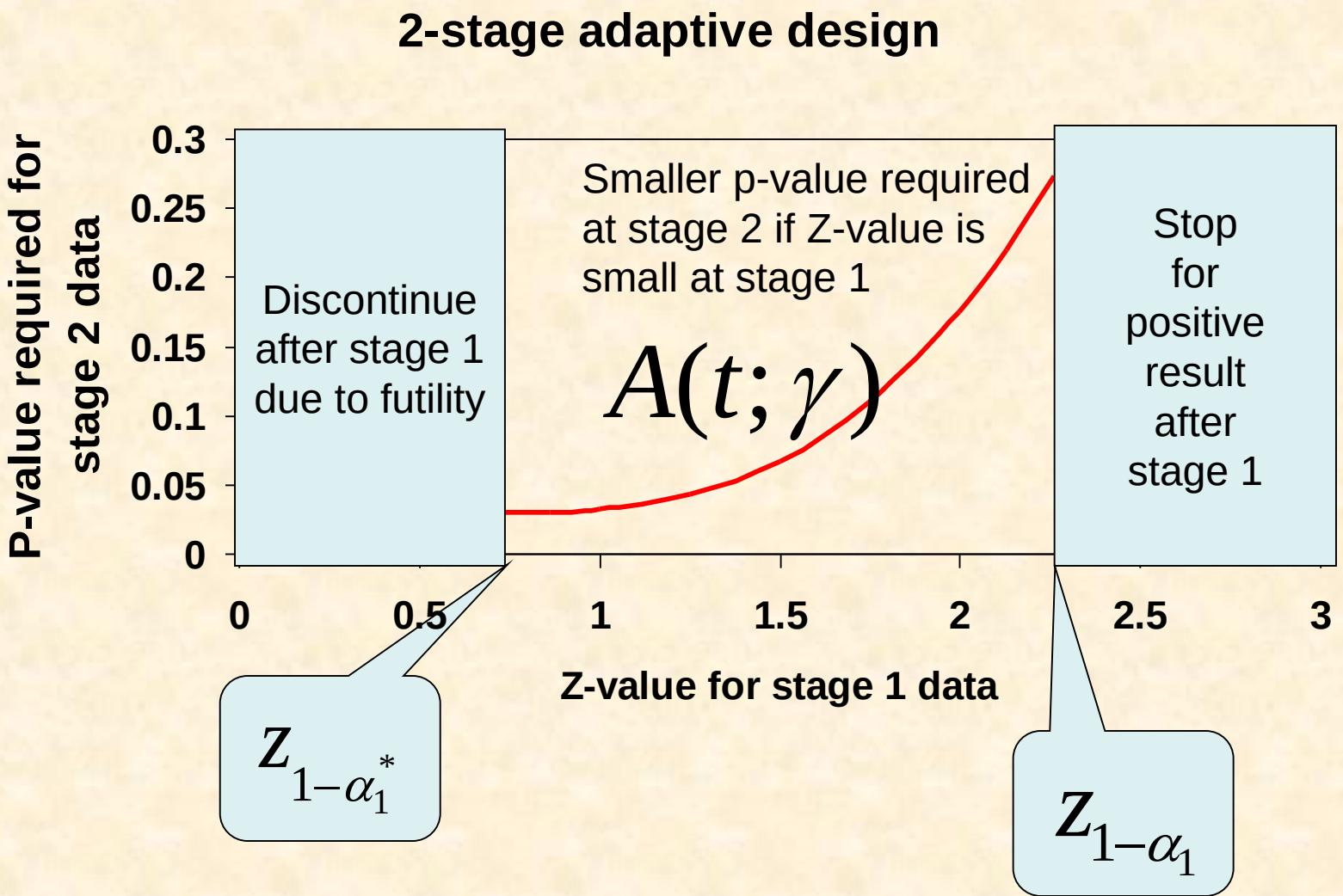
2-Stage Adaptive Design

- Sample n_1 subjects in stage 1
- Compute p-value for stage 1 data
  - May stop if positive or futile
- If trial continues, map observed stage 1 p-value to required stage 2 critical value
- Second stage sample size n_2 determined at end of stage 1
- Compute observed p-value for stage 2 data (excluding stage 1 data) and compare to stage 2 critical value
  - Trial positive if 1st or 2nd stage is positive

Conditional Power for Computing n_2

- Proschan and Hunsberger, Biometrics,1995
  - Estimate $\hat{\theta}$ from stage 1 data
  - Given this value and a critical value for stage 2, compute n_2 to achieve desired power
- Liu and Chi, Biometrics, 2001
  - Substitute θ_0, a minimum value of interest, for $\hat{\theta}$
  - This has the effect of reducing maximum sample size

Conditional Error Function



Properties of $A(t;\gamma)$

- Values in [0,1/2)
  - Require additional evidence in stage 2
- As a function of t, $A(t,\gamma)$ is
  - Defined on (-1,1)
  - Non-decreasing
- 'Nuisance' parameter γ
  - $A(t,\gamma)$ increasing in γ
  - Used to obtain the desired overall Type I error given the stage 1 critical values α_1 and α_1^*

Type I Error

$$\alpha = \alpha_1 + \alpha_2$$

where

$$\alpha_2 = \int_{z_{1-\alpha_1^*}}^{z_{1-\alpha_1}} A(t;\gamma)\varphi(t)dt$$

γ is typically used to adjust α_2 appropriately

Generalizing $A(t;\gamma)$

- Want to choose from a broad class of $A()$ functions to get a ‘good’ one
- Add 2 more parameters (η, ν) to allow flexibility in the shape and range of values
- Optimize over α_1, α_1^*, η and ν
- Still use γ to get desired overall Type I error given values of α_1, α_1^*, η and ν

Power Function Family

$$A(t;\gamma,\eta,\nu) = \eta + \nu\left(\frac{t - z_{1-\alpha_1^*}}{z_{1-\alpha_1} - z_{1-\alpha_1^*}}\right)^{\gamma}$$

- $A()$ increasing in t
- $A(z_{1-\alpha_1^*};\gamma,\eta,\nu)=\eta$
- $A(z_{1-\alpha_1};\gamma,\eta,\nu)=\eta+\nu$
- γ determines shape

What to optimize?

Optimization Problem Set-up

- Assume a prior distribution for θ
- Choose a 'loss' function (e.g., expected sample size)
- Fixed parameters:
  - α : Type I error
  - 1- β: Power at minimum parameter value of interest θ_0
- Variable
  - n_1, sample size at stage 1
  - α_1, stage 1 Type I error
  - α_1^*, probability of futility at stage 1 for θ=0
  - η, ν : determine shape of A()

Minimize wrt n_1, α_1, α_1^*, η, ν

$$\int_{z_{1-\alpha_1^*}}^{z_{1-\alpha_1}} \int L(n_1, n_2(t; n_1, \alpha_1, \alpha_1^*, \eta, \nu)) dF(t; \theta, n_1) d\pi(\theta)$$

- $L()$ is the loss function
- $n_2(t;)$ is the sample size for stage 2 given a z-value of t was observed at stage 1 (formula not shown, but it is simple and is made of 'standard' components)
- t is the z-value at stage 1
- $F()$ is a normal distribution with variance 1
- $\pi()$ is the prior distribution for θ

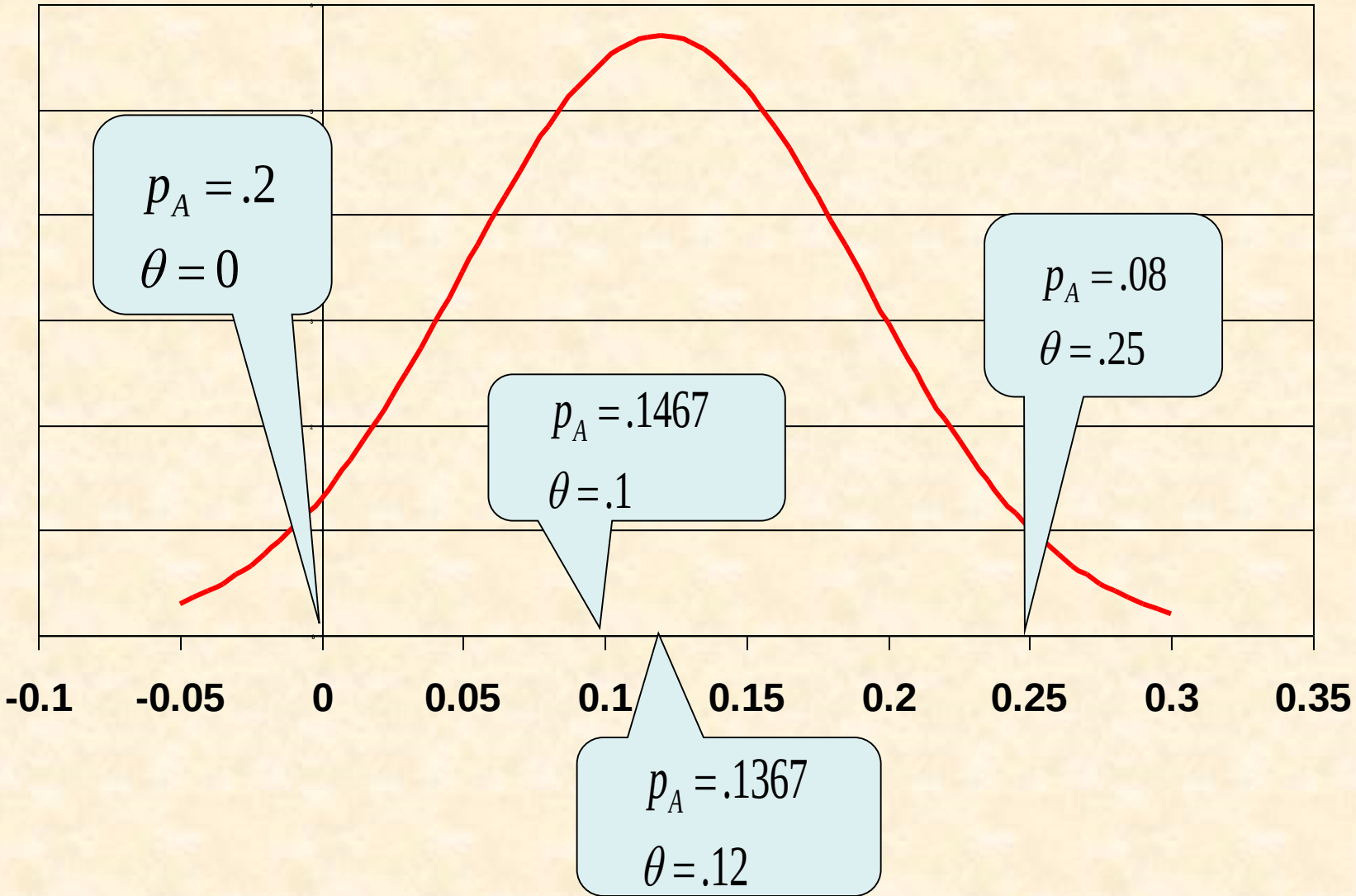
Optimization
(in a nutshell)

- All functions are continuous in the given parameters
- Transform problem to an unconstrained optimization
- Use numerical integration to compute function
- Use off-the-shelf optimization for function without known derivatives (Powell's method)

Example

- Binary outcome
  - Control event rate estimate: p_C=20%
  - Reduction by > 25% (say, p_A=14.67%) considered clinically meaningful
    - θ=arcsin(.20$^{1/2}$)-arcsin(.1467)$^{1/2}$)=0.10
  - Reduction by 30% considered likely
    - θ=arcsin(.20$^{1/2}$)-arcsin(.1367$^{1/2}$)=0.12
  - Moderately weak prior distribution:
    - θ ~ Normal(μ=0.12,σ=.07)
    - Implies ~5% chance of no effect or worse

Prior density for θ (assume p_C=0.2)



Example (cont.)

- Suppose n_1 observations ($n_1/2$ per arm) collected in stage 1
- At that time

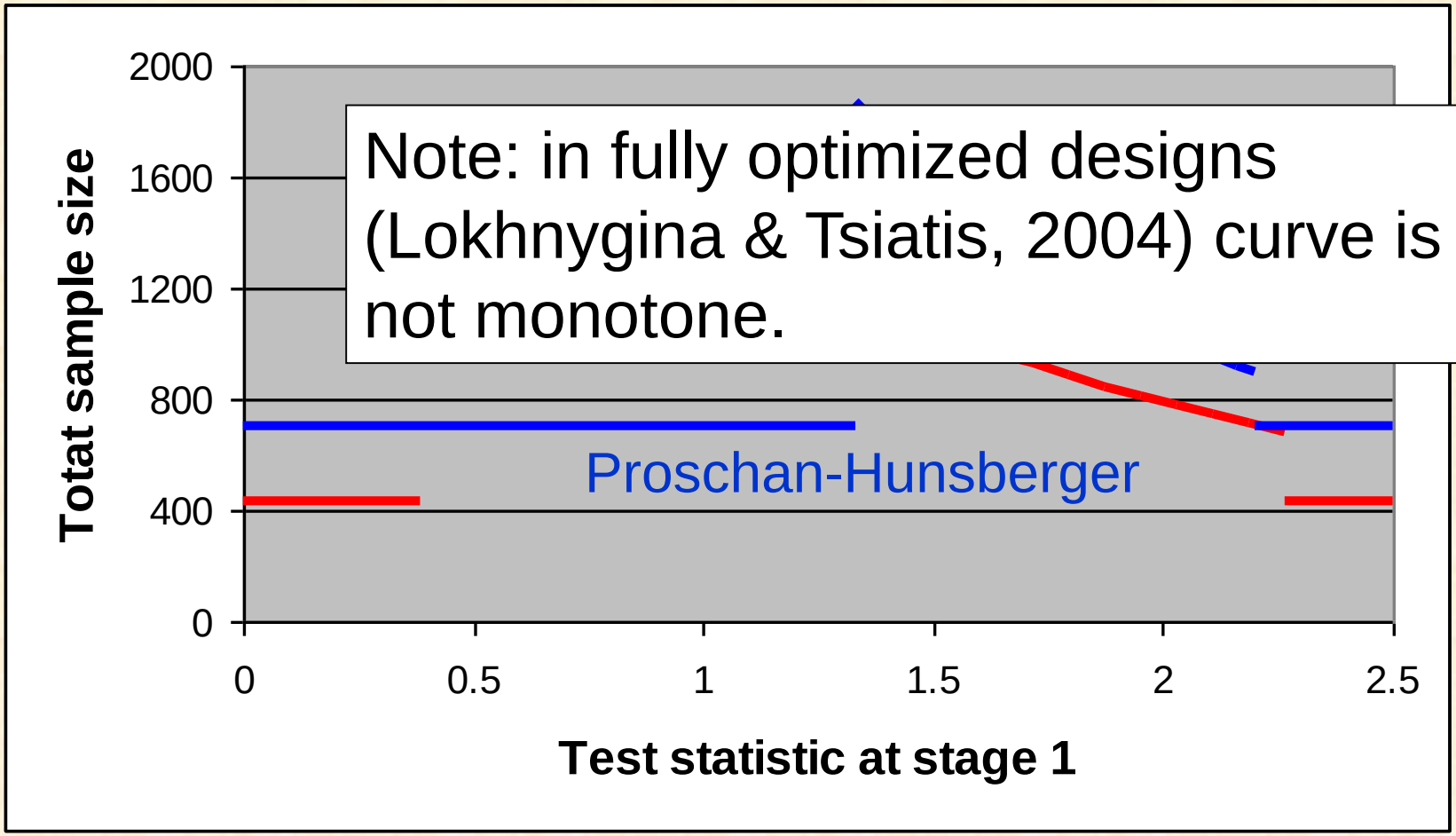
$$\hat{\theta} = \sqrt{n_1}\left\{\arcsin(\sqrt{\hat{p}_C})\text{-}\arcsin(\sqrt{\hat{p}_A})\right\}$$

is distributed approximately Normal($\theta n_1^{1/2}$,1)

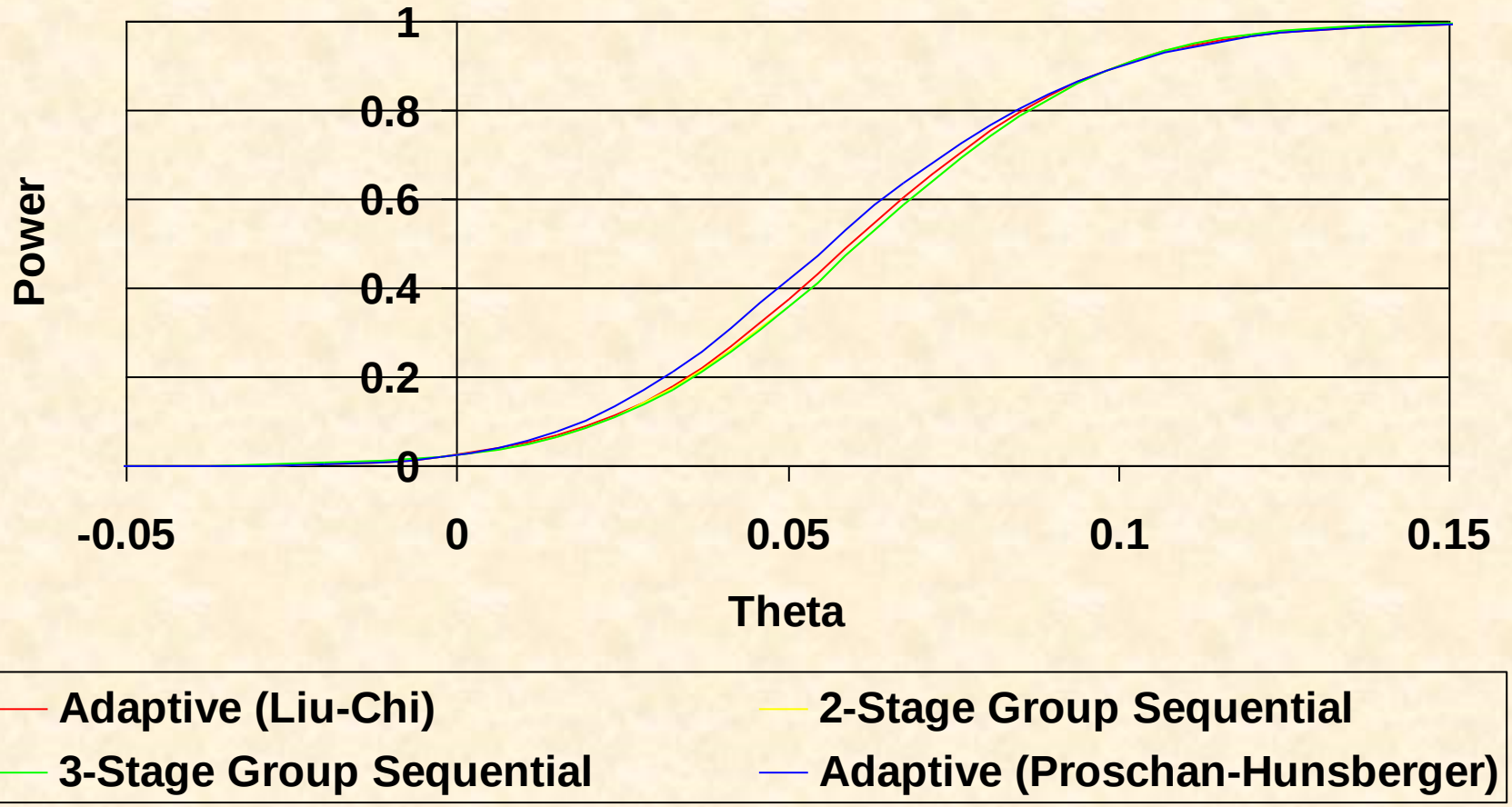
Designs compared

- All have
  - 90% power when θ=0.1
  - Type I error (one-sided) = 0.025
- Designs
  - Optimal adaptive (among Liu & Chi designs)
  - Optimal group sequential
    - 2-stage
    - 3-stage
  - Optimal adaptive (among Proschan-Hunsberger designs)

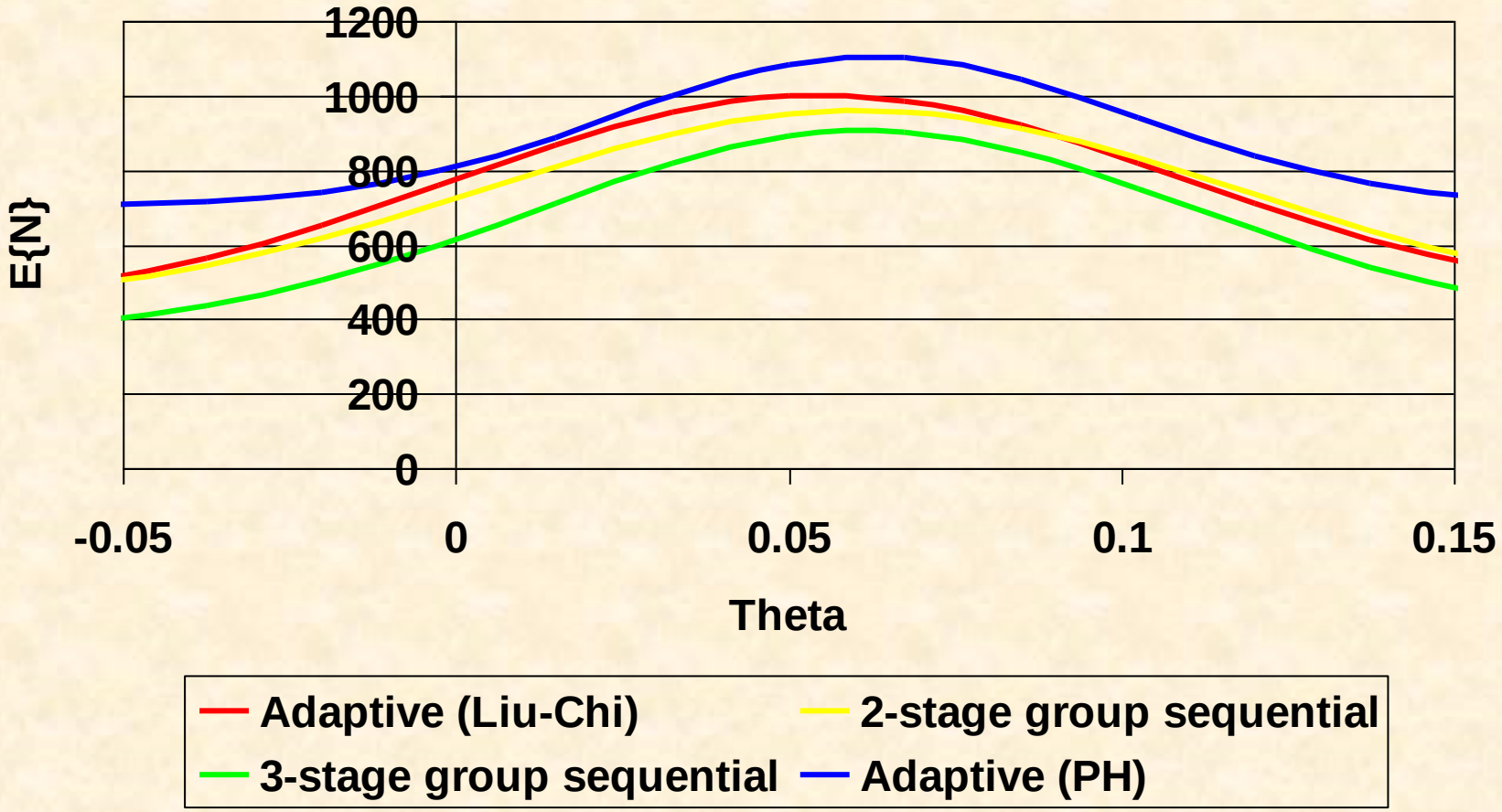
Final Sample Size Based on Interim Analysis: Optimal Designs



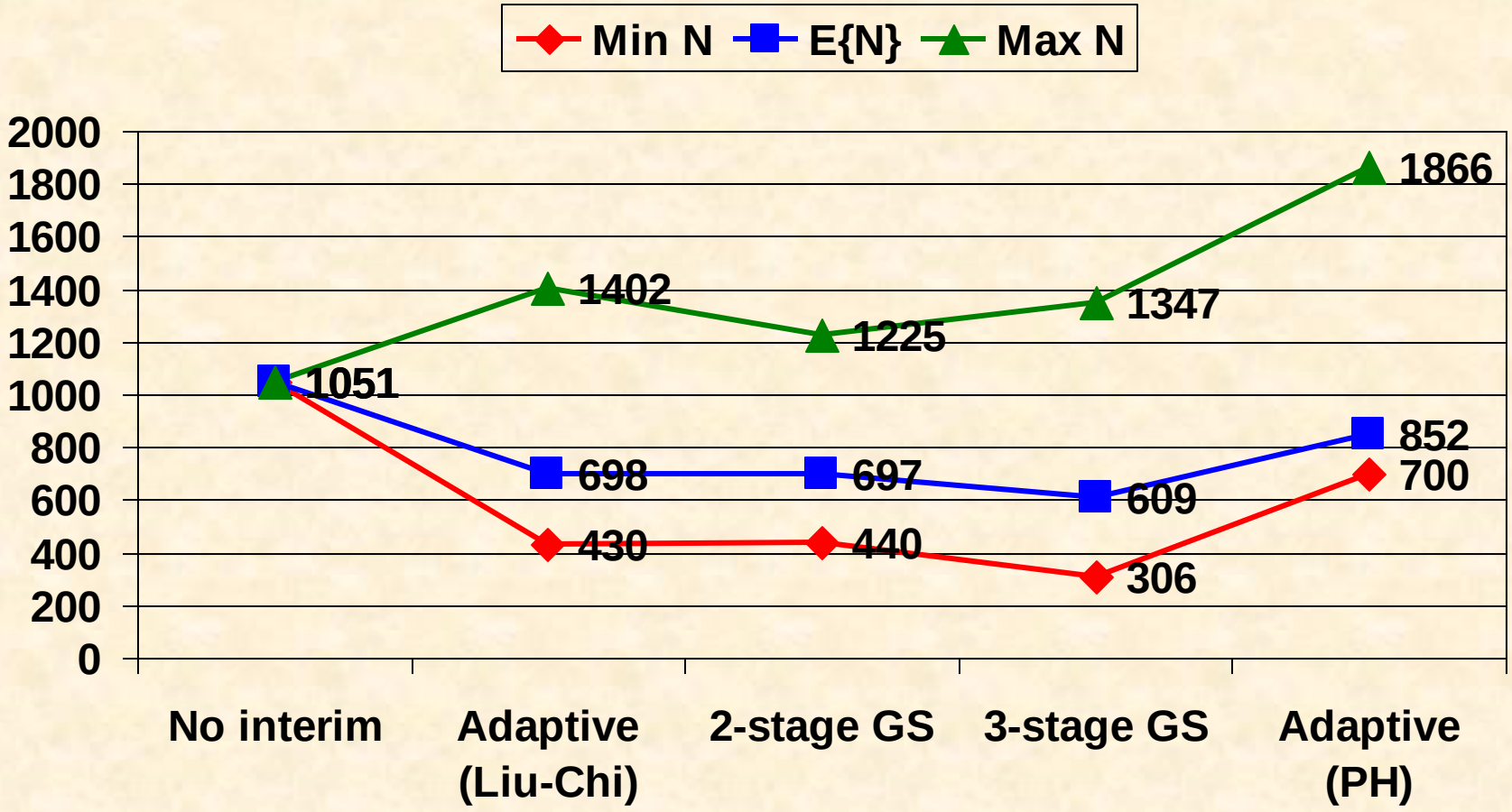
Power of Optimal Tests



E{N} for Optimal Designs



Sample Size for Optimal Designs



Summary

- Adapting using a fixed, minimum treatment effect of interest (Liu-Chi method) appears to be better than adapting to the estimated effect at the time of interim analysis (Proschan-Hunsberger method)

Summary

Assuming a single interim analysis we have shown an example where best adaptive and group sequential designs have essentially identical:

- Power over a range of parameter values
- Expected sample size when averaged over possible parameter values using a prior distribution

Issues…

- Can we improve optimized adaptive designs by not insisting on constant conditional power?
  - Set maximum sample size (Posch, et al, 2001)
  - Lokhnygina & Tsiatis (2004)
- Other methods of comparing adaptive and group sequential designs
  - Qing Liu: ‘effectiveness’

REFERENCES

1. Barber S,.Jennison C. Optimal asymmetric one-sided group sequential tests. *Biometrika* 2002;**89**:49-60.
2. Jennison C,.Turnbull BW. Mid-course sample size modification in clinical trials based on the observed treatment effect. *Stat.Med* 2003;**22**:971- 93.
3. Jennison C, Turnbull BW. Group Sequential Methods with Applications to Clinical Trials. 2002.
4. Liu Q,.Chi GY. On sample size and inference for two-stage adaptive designs. *Biometrics* 2001;**57**:172-7.
6. Liu, Q., Anderson, K. M., and Pledger, G. W. Benefit-risk evaluation of multi-stage adaptive designs. *Sequential Analysis*, in press, 2004.
7. Lokhnygina, Y. and Tsiatis, A. Optimal Two-stage Adaptive Designs. Submitted for publication.
8. Posch, M, Bauer, P and Brannath, W. Issues in designing flexible trials. Stat Med 2003;22:953-969.
9. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical Recipes in C: The Art of Scientific Computing. 1992.
10. Proschan MA,.Hunsberger SA. Designed extension of studies based on conditional power. *Biometrics* 1995;**51**:1315-24.
11. Tsiatis AA,.Mehta C. On the inefficiency of the adaptive design for monitoring clinical trials. *Biometrika* 2003;**90**:367-78.

Backup slides

- Conditional error function families

Generalized Proportional Error Function Family

$$A(t;\gamma,\eta,\nu)=\eta\left\{1-\Phi\left(\left[\gamma(z_{1-\alpha_1}-t)\right]^{\nu}\right)\right\}$$

- $A()$ increasing in t
- $A(z_{1-\alpha_1};\gamma,\eta,\nu)=\eta$
- γ and ν together determine shape