



Shaping the Future of
Drug Development

Group Sequential and Adaptive designs in East[®]

Cyrus Mehta

Cytel Corporation and Harvard TH Chan
April 25, 2016

This workshop will teach you how to best use East® to design, simulate, and monitor adequately and well controlled trials, while incorporating group sequential and adaptive elements into the design.

East® is the industry standard software package for the design of clinical trials. It directly addresses many of the requirements given in the FDA "Draft Guidance on Adaptive Design for Clinical Trials for Drugs and Biologics."

- To provide the general background of adaptive designs in clinical trials
- To work through some practical examples using East
- To inform about regulatory positions regarding adaptive designs

1 Introduction

2 Statistical Methods for Group Sequential Designs

- Distribution theory
- Efficacy stopping rules
- Power and sample size calculations
- Example: The CAPTURE Trial
- Futility stopping rules

3 Survival Designs in East

- Example: The JUPITER Study
- Calculating the required number of events
- Handling drop-outs, variable accrual and non-constant hazards
- Non-proportional hazards

4 Workshops 1 - 2 (Group Sequential)

Warm-up Exercise; Recall Fixed Sample Size Session

Cytel



Sample Size Calculation 1



Design a Phase 3 trial for a normal endpoint (difference of means for two independent populations). What is the total sample size, n , required?

- ① Given two-sided $\alpha = 0.05$,
- ② power $1 - \beta = 0.85$, so $Z_{\frac{\alpha}{2}} + Z_\beta \approx 3$,
- ③ treatment effect $\delta = 0.3$,
- ④ and nuisance parameter $\sigma = 1$,
- ⑤ compute the sample size $n = 4 \left(\frac{Z_{\frac{\alpha}{2}} + Z_\beta}{\delta/\sigma} \right)^2 = ?$

Sample Size Calculation 2

Cytel

Confirm your answer in East:



Design: Continuous Endpoint: Two-Sample Test - Parallel Design - Difference of Means

Design Type: Superiority Number of Looks: 1

Design Parameters

Test Type:	2-Sided	Input Method:	Difference of Means	Test Statistic:	Z		
Type I Error (α):	0.05	<input type="radio"/>	Diff. in Means ($\delta = \mu_t - \mu_c$):	0.3	<input type="radio"/>	Std. Deviation (σ):	1
Power:	0.85	<input type="radio"/>					
Sample Size (n):	Computed	<input checked="" type="radio"/>					
Allocation Ratio:	1	(n_t / n_c)					

Assurance (Probability of Success)

Sample Size Calculation 3

Cytel

Select Des1 in Output Preview, and click 'Output Summary' icon:

The screenshot shows two windows from the Cytel software interface. The top window is titled "Output Summary" and displays "Des1" parameters. The bottom window is titled "Output Preview" and also displays "Des1" parameters. A red box highlights the "Output Summary" icon in the Output Preview window toolbar.

Output Summary

Des1	
Mnemonic	MN-2S-DI
Test Parameters	
Design Type	Superiority
No. of Looks	1
Test Type	2-Sided
Specified α	0.05
Power	0.851
Model Parameters	
Input Method	Difference of Means
Diff. in Means ($\delta = \mu_t - \mu_c$)	0.3
Std. Deviation (σ)	1
Test Statistic	Z
Allocation Ratio (nt/nc)	1
Sample Size	
Maximum	400

output Summary icon

Output Preview

ID	Design Type	No. of Looks	Test Type	Specified α	Power	nt/nc	Sample Size	Input Method	δ	σ	Test Statistic
Des1	Superiority	1	2-Sided	0.05	0.851	1	400	Difference of Means	0.3	1	Z

Select Des1 in Output Preview, and click 'Save in Workbook' icon:

The screenshot shows the "Output Preview" window from the Cytel software. A red box highlights the "Save in workbook" icon in the toolbar. The table below shows the same data as the previous screenshot.

Output Preview

ID	Design Type	No. of Looks	Test Type	Specified α	Power	nt/nc	Sample Size	Input Method	δ	σ	Test Statistic
Des1	Superiority	1	2-Sided	0.05	0.851	1	400	Difference of Means	0.3	1	Z

Sample Size Calculation 4

Cytel

Select Des1 in the Library, and click 'Details' icon:

The screenshot shows the Cytel software interface. On the left, there is a 'Library' pane with a tree view. A red box highlights the 'Details' icon (a magnifying glass) on the toolbar above the tree view. The tree view shows 'Root' with a folder 'Wbk1' containing a file 'Des1'. The main window displays the 'Design: Continuous Endpoint: Two-Sample Test - Parallel Design - Difference of Means' dialog. This dialog is divided into sections: 'Test Parameters' and 'Model Parameters' under 'Test Parameters', and 'Output' under 'Output'. The 'Test Parameters' section contains the following data:

Design ID:	Des1
Design Type:	Superiority
Number of Looks:	1
Test Type:	2-Sided
Specified α :	0.05
Power:	0.851

The 'Model Parameters' section contains the following data:

Test Statistic:	Z
Input Method:	Difference of Means
$\delta = \mu_t - \mu_c$	
Under H0:	0
Under H1:	0.3
Std. Deviation (σ):	1
Allocation Ratio (n_t/n_c):	1

The 'Output' section displays the following results:

Upper Critical Point	1.96
Lower Critical Point	-1.96
Sample Size (n)	400
Sample Size Treatment (n_t)	200
Sample Size Control (n_c)	200
Max. Information (I_{max})	100

Introduction to Adaptive Designs

What are Adaptive Design Clinical Trials?



- “...defined as a study that includes a prospectively planned opportunity for modification of one or more specified aspects of the study . . .”
- (FDA, 2010)

Guidance for Industry

Adaptive Design Clinical Trials for Drugs and Biologics

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Comments may be submitted via email to CBERAdvisory@fda.hhs.gov, or by mail to U.S. Food and Drug Administration, 5630 Fishers Lane, rm. 1041, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability published in the *Federal Register*.

For questions regarding this draft document contact Robert O'Neill or Sue-Jane Wong at 301-796-7100; Marc Wallen at 301-796-2060 (CBER), or the Office of Communication, Outreach and Development (OCORD) at 888-535-4709 or 301-435-1380.

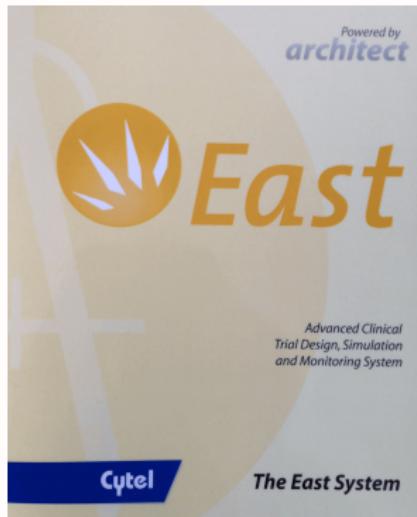
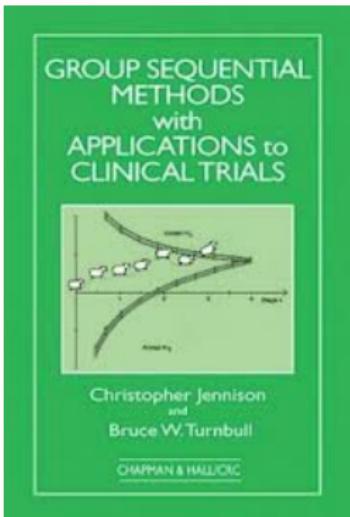
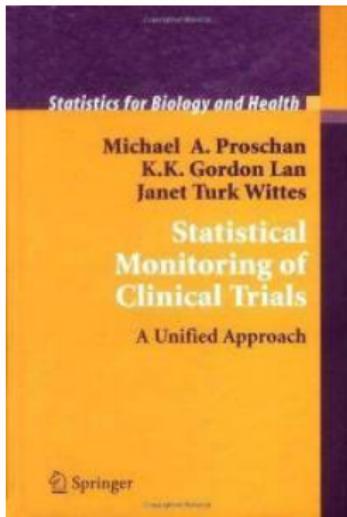
U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

February 2010
Clinical/Medical

- **Traditional Design:** fix the sample size in advance, and perform one analysis after all subjects have been enrolled and evaluated
- **Adaptive Design:** monitor the accruing data, and make interim decisions concerning the future course of the study (e.g., stop early, select dose, increase sample size)
- **Group Sequential Design:** a type of adaptive design in which the only interim decision is whether to stop early (for harm, efficacy, futility)

Background Readings

Cytel



Statistical Methods for Group Sequential Designs

Example: The CAPTURE Trial

Cytel

THE LANCET

Randomised placebo-controlled trial of abciximab before and during coronary intervention in refractory unstable angina: the CAPTURE study

*The CAPTURE investigators**

Summary

Background Platelet aggregation is a dominant feature in the pathophysiology of unstable angina. Percutaneous transluminal coronary angioplasty (PTCA) in patients with this disorder carries an increased risk of thrombotic complications. Abciximab (c7E3) blocks the platelet glycoprotein IIb/IIIa receptor, thus preventing platelet adhesion and aggregation. The CAPTURE study was a randomised placebo-controlled multicentre trial to assess whether abciximab can improve outcome in patients with refractory unstable angina who are undergoing PTCA.

Introduction

In coronary artery disease, unstable angina may develop in association with plaque fissuring or rupture, with subsequent aggregation, platelet adhesion, and intracoronary thrombosis.^{1,2} An episode of unstable angina may progress to myocardial infarction or death, or may stabilise after one or more ischaemic episodes. Different classes of unstable angina have been recognised, with different event rates.^{3,4} In patients who have recurrent episodes of myocardial ischaemia despite intensive medical therapy (refractory unstable angina), percutaneous transluminal coronary angioplasty (PTCA)

- CAPTURE (1997) was a parallel arm placebo-controlled trial, with binary primary endpoint
- Use EasT to compute the sample size required to detect a reduction in event rates (difference of proportions) from 15% to 10% with 80% power using a two-sided level-0.05 test

Single-Look Design

Cytel

	Des 1
Mnemonics	PN-2S-DI
Test Parameters	
Trial Type	Superiority
No. of Looks	1
Test Type	2-Sided
Variance	Unpooled Estimate
Specified α	0.05
Specified Power	0.8
Attained Power	0.8
Model Parameters	
Proportion under Control (π_c)	0.15
Proportion under Treatment (π_t)	0.1
Diff. in Prop. ($\pi_t - \pi_c$)	-0.05
Allocation Ratio (n_t/n_c)	1
Sample Size	
Maximum	1366

Very large sample size commitment with no possibility of early termination for benefit, harm or futility

- Suppose Abciximab is actually beneficial: Do we really have to randomize 683 patients to placebo before we know for sure?
- What if Abciximab is no different than placebo or even harmful? Can we avoid randomizing 683 patients to a treatment that is no better or worse than placebo?

Interim monitoring can help in both these cases

- ① Understand maximum information, and information fractions
- ② Contrast Pocock and O'Brien-Fleming boundaries
- ③ Choose appropriate error spending functions
- ④ Compare designs (e.g., fixed vs group sequential) in East
- ⑤ Display boundaries on Z-scale and other scales
- ⑥ Calculate expected sample sizes for group sequential designs

1 Introduction

2 Statistical Methods for Group Sequential Designs

- Distribution theory
- Efficacy stopping rules
- Power and sample size calculations
- Example: The CAPTURE Trial
- Futility stopping rules

3 Survival Designs in East

- Example: The JUPITER Study
- Calculating the required number of events
- Handling drop-outs, variable accrual and non-constant hazards
- Non-proportional hazards

4 Workshops 1 - 2 (Group Sequential)

Let δ denote the true, unknown treatment effect. We always define δ to be a difference between the treatment and control groups. For instance:

- a difference of two means
- a difference of two binomial probabilities
- a log hazard ratio
- a log odds ratio
- any general coefficient in a regression model

Monitor the data K times at calendar times $\tau_1, \tau_2, \dots, \tau_K$

- For normal and binomial endpoints let

$$n_j = \text{sample size at calendar time } \tau_j$$

- For time-to-event endpoints let

$$d_j = \text{number of events at calendar time } \tau_j$$

- More generally, in terms of Fisher information let

$$I_j = \text{information at calendar time } \tau_j \approx \left[se\left(\hat{\delta}_j\right) \right]^{-2}$$

where $\hat{\delta}_j$ is an efficient estimate of δ at calendar time τ_j .

The Maximum Information



The maximum information is the precision of $\hat{\delta}$ needed to achieve the desired power

- For normal and binomial endpoints let

n_{\max} = the maximum sample size required for the trial

- For time-to-event endpoints let

d_{\max} = maximum number of events required for the trial

- More generally, in terms of Fisher information, let

$$I_{\max} \approx \left[se\left(\hat{\delta}_{\max}\right) \right]^{-2} = \text{maximum information to be collected}$$

In a group sequential design we will keep the trial open until either the maximum information is obtained or a stopping boundary is crossed.

The Information Fraction



Define the information fraction t_j at calendar time τ_j

$$t_j = \begin{cases} \frac{n_j}{n_{\max}} & \text{for normal and binomial} \\ \frac{d_j}{d_{\max}} & \text{for time-to-event} \\ \frac{I_j}{I_{\max}} & \text{in general} \end{cases}$$

- If K is intended to be the last look, we will often denote I_{\max} by I_K , n_{\max} by n_K , and d_{\max} by d_K
- We may regard the information fraction t , $0 \leq t \leq 1$, as the internal time axis of the clinical trial

1 Introduction

2 Statistical Methods for Group Sequential Designs

- Distribution theory
- **Efficacy stopping rules**
- Power and sample size calculations
- Example: The CAPTURE Trial
- Futility stopping rules

3 Survival Designs in East

- Example: The JUPITER Study
- Calculating the required number of events
- Handling drop-outs, variable accrual and non-constant hazards
- Non-proportional hazards

4 Workshops 1 - 2 (Group Sequential)

- Let $\{c_1, c_2, \dots, c_K\}$ be the corresponding two-sided stopping boundaries
- Stop the trial and reject H_0 at the first t_j such that

$$|Z(t_j)| \geq c_j$$

- We must select the c_j 's so as to preserve the type-1 error

$$P_0\left\{\bigcap_{j=1}^K |Z(t_j)| < c_j\right\} = 1 - \alpha$$

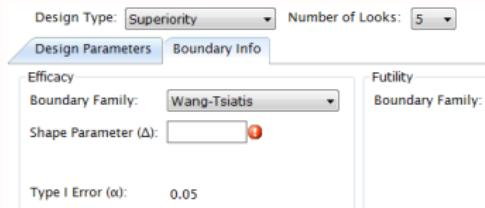
- Many c_j 's satisfy this condition.

- Pocock (1977) was the first to propose a group sequential design: a constant boundary $c_j = C$ on the Wald scale.
 - ▶ An equally spaced 5-look level-0.05 two-sided Pocock test utilizes constant boundaries $c_j = 2.413$ for $j = 1, 2, \dots, 5$
- The next proposal came from O'Brien and Fleming (1979) who introduced boundaries of the form $c_j = C/\sqrt{t_j}$.
 - ▶ An equally spaced 5-look level-0.05 two-sided O'Brien-Fleming design utilizes the boundaries $c_j = 2.04/\sqrt{t_j}$ for $j = 1, 2, \dots, 5$

- Wang and Tsiatis (1987) proposed a family of boundary shapes:

$$c_j = Ct_j^{\Delta-1/2}$$

where Δ is known as a **shape parameter**

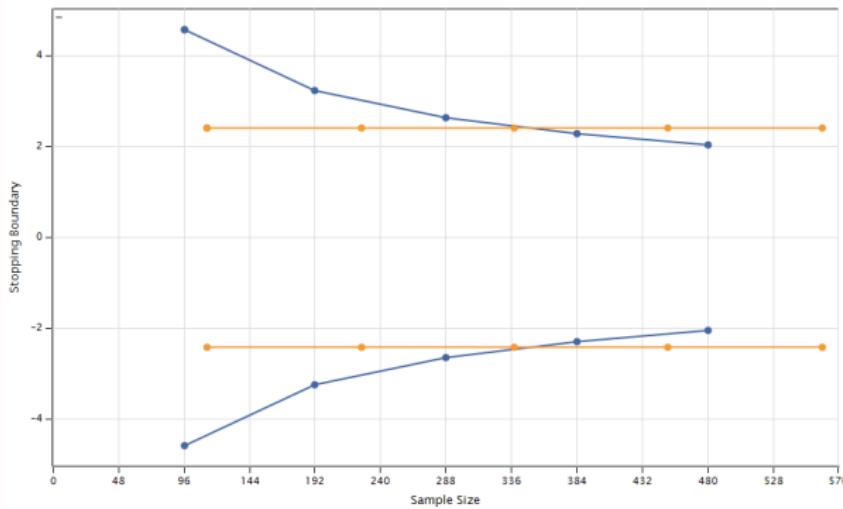


- Which values of Δ yield the Pocock (1977) and O'Brien-Fleming (1979) boundaries, respectively?

Exercise: Pocock vs O'Brien-Fleming

Cytel

- In East, plot the Pocock (1977) and O'Brien-Fleming (1979) boundaries on the Wald scale.



- What key differences do you notice between these boundaries? Are they advantages or disadvantages?

- Wang-Tsiatis boundaries depend on pre-specified values of the information fraction t_1, t_2, \dots, t_K
- In practice, it might be necessary to alter the spacing or number of looks after the trial has started, which requires changing remaining boundary values
- Lan and DeMets (1983) proposed the [error spending function](#) approach, where α is treated as a budgeted quantity to be spent

The α -Spending Function Approach



- Specify a monotone increasing function of t for $t \in [0, 1]$, with $\alpha(0) = 0$ and $\alpha(1) = \alpha$
- Solve recursively for c_1, c_2, \dots, c_K

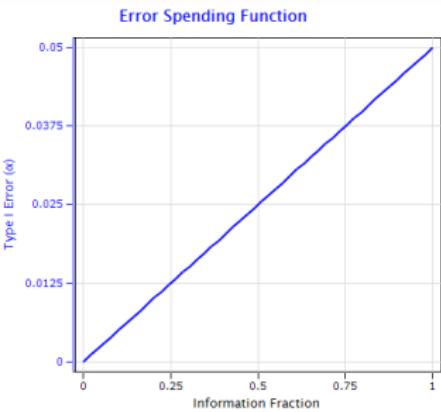
$$P_0\{ |Z(t_1)| \geq c_1 \} = \alpha(t_1)$$

and for $j = 2, \dots, K$

$$\alpha(t_{j-1}) + P_0\{ |Z(t_1)| < c_1, \dots, |Z(t_{j-1})| < c_{j-1}, |Z(t_j)| \geq c_j \} = \alpha(t_j)$$

Exercise: Cumulative α spent and Type I error

Consider a linear spending function ($\alpha = 0.05$) with 4 equally-spaced looks:



- What are the values of the information fractions: t_1, t_2, t_3, t_4 ?
- What are the values of cumulative alpha spent: $\alpha(t_1), \alpha(t_2), \alpha(t_3), \alpha(t_4)$?
- What is the probability of committing a Type I error by Look 2?
- What is the probability of committing a Type I error at Look 3?

Spending Functions and Stopping Boundaries



- Lan and DeMets proposed two spending functions
 - ▶ The $LD(OF)$ spending function

$$\alpha(t) = \begin{cases} 4 - 4\Phi(z_{\alpha/4})/\sqrt{t} & \text{for two-sided tests} \\ 2 - 2\Phi(z_{\alpha/2})/\sqrt{t} & \text{for one-sided tests} \end{cases}$$

yields boundaries that approximate the O'Brien-Fleming boundaries

- ▶ The $LD(PK)$ spending function

$$\alpha(t) = \alpha \log\{1 + (e - 1)t\}$$

yields boundaries that approximate the Pocock boundaries

LD(OF) vs LD(PK)

Cytel

- Which of these spending functions corresponds to LD(OF) and LD(PK)?



Gamma Family

- Hwang, Shih, & DeCanis (1990) proposed

$$\alpha(t) = \alpha \frac{(1 - e^{-\gamma t})}{(1 - e^{-\gamma})}, \text{ where } \gamma \neq 0$$

- $\gamma = -4$ or -5 (similar to O'Brien-Fleming); $\gamma = 1$ (similar to Pocock)

1 Introduction

2 Statistical Methods for Group Sequential Designs

- Distribution theory
- Efficacy stopping rules
- **Power and sample size calculations**
- Example: The CAPTURE Trial
- Futility stopping rules

3 Survival Designs in East

- Example: The JUPITER Study
- Calculating the required number of events
- Handling drop-outs, variable accrual and non-constant hazards
- Non-proportional hazards

4 Workshops 1 - 2 (Group Sequential)

Three-Step Procedure for Sample Size / Events



- ① Compute **boundaries**, (c_1, c_2, \dots, c_K) , given α , spacing, number of looks and spending function
- ② Compute **maximum information** I_{max} needed to achieve the desired power, given boundaries, treatment effect δ and desired power $1 - \beta$
- ③ Convert I_{max} into **sample size** n_{max} or **events** D_{max} , for given endpoint, and nuisance parameters

Maximum Information I_{\max} required to reject $H_1 : \delta = \delta_1$ with $1 - \beta$ power



- Jennison and Turnbull (1997) showed that $Z(t_j) \sim N(\eta\sqrt{t_j}, 1)$, where $\eta = \delta_1\sqrt{I_{\max}}$, and, for any $t_{j_1} < t_{j_2}$, $\text{cov}\{Z(t_{j_1}), Z(t_{j_2})\} = \sqrt{\frac{t_{j_1}}{t_{j_2}}}$
- Find the value of the drift parameter η that satisfies the equation

$$P_\eta\left\{\bigcup_{j=1}^K |Z(t_j)| \geq c_j\right\} = 1 - \beta$$

- Solve for

$$I_{\max} = \left[\frac{\eta}{\delta_1}\right]^2$$

- The effect size δ is estimated by the difference of the two group sample means

$$\begin{aligned} I_{\max}^{-1} \equiv \text{var}[\hat{\delta}(t_K)] &= \text{var}[\bar{X}_t(t_K) - \bar{X}_c(t_K)] \\ &= \frac{\sigma^2}{(n_{\max}/2)} + \frac{\sigma^2}{(n_{\max}/2)} \\ &= \frac{4\sigma^2}{n_{\max}} \end{aligned}$$

where σ is the common standard deviation and assuming balanced allocation

- Therefore the relationship between n_{\max} and I_{\max} is

$$n_{\max} = 4\sigma^2 I_{\max}$$

- To compute n_{\max} , we need σ , a “nuisance parameter”

- The effect size δ is estimated by the difference of the two group binomial response rates

$$\begin{aligned} I_{\max}^{-1} \equiv \text{var}[\hat{\delta}(t_K)] &= \text{var}[\hat{\pi}_t(t_K) - \hat{\pi}_c(t_K)] \\ &= \frac{\pi_t(1-\pi_t)}{(n_{\max}/2)} + \frac{\pi_c(1-\pi_c)}{(n_{\max}/2)} \end{aligned}$$

where we have assumed balanced allocation

- Solve for n_{\max} under $H_1 : \pi_t - \pi_c = \delta_1$ to obtain

$$n_{\max} = 2[\pi_c(1 - \pi_c) + (\pi_c + \delta_1)(1 - \pi_c - \delta_1)]I_{\max}$$

- To compute n_{\max} , we need π_c , a “nuisance parameter”

- For time-to-event endpoints, we usually assume the **proportional hazards** alternative. In this case, the effect size δ is the logarithm of the hazard ratio and the variance of $\hat{\delta}_j$ at any interim look j is proportional to the number of events D_j .
- For balanced randomization, it can be shown that, approximately

$$I_{\max}^{-1} \equiv \text{var}[\hat{\delta}(t_K)] \approx [D_{\max}/4]^{-1}$$

- Thus

$$D_{\max} = 4I_{\max}$$

1 Introduction

2 Statistical Methods for Group Sequential Designs

- Distribution theory
- Efficacy stopping rules
- Power and sample size calculations
- Example: The CAPTURE Trial
- Futility stopping rules

3 Survival Designs in East

- Example: The JUPITER Study
- Calculating the required number of events
- Handling drop-outs, variable accrual and non-constant hazards
- Non-proportional hazards

4 Workshops 1 - 2 (Group Sequential)

Example: The CAPTURE Trial Revisited



THE LANCET

Randomised placebo-controlled trial of abciximab before and during coronary intervention in refractory unstable angina: the CAPTURE study

*The CAPTURE investigators**

Summary

Background Platelet aggregation is a dominant feature in the pathophysiology of unstable angina. Percutaneous transluminal coronary angioplasty (PTCA) in patients with this disorder carries an increased risk of thrombotic complications. Abciximab (c7E3) blocks the platelet glycoprotein IIb/IIIa receptor, thus preventing platelet adhesion and aggregation. The CAPTURE study was a randomised placebo-controlled multicentre trial to assess whether abciximab can improve outcome in patients with refractory unstable angina who are undergoing PTCA.

Introduction

In coronary artery disease, unstable angina may develop in association with plaque fissuring or rupture, with subsequent aggregation, platelet adhesion, and intracoronary thrombosis.^{1,2} An episode of unstable angina may progress to myocardial infarction or death, or may stabilise after one or more ischaemic episodes. Different classes of unstable angina have been recognised, with different event rates.^{3,4} In patients who have recurrent episodes of myocardial ischaemia despite intensive medical therapy (refractory unstable angina), percutaneous transluminal coronary angioplasty (PTCA)

- CAPTURE (1997) was a parallel arm placebo-controlled trial, with binary primary endpoint
- Use East to compute the sample size required to detect a reduction in event rates (difference of proportions) from 15% to 10% with 80% power using a two-sided level-0.05 test

Single-Look Design

Cytel

	Des 1
Mnemonics	PN-2S-DI
Test Parameters	
Trial Type	Superiority
No. of Looks	1
Test Type	2-Sided
Variance	Unpooled Estimate
Specified α	0.05
Specified Power	0.8
Attained Power	0.8
Model Parameters	
Proportion under Control (π_c)	0.15
Proportion under Treatment (π_t)	0.1
Diff. in Prop. ($\pi_t - \pi_c$)	-0.05
Allocation Ratio (n_t/n_c)	1
Sample Size	
Maximum	1366

Very large sample size commitment with no possibility of early termination for benefit, harm or futility

- Suppose Abciximab is actually beneficial: Do we really have to randomize 683 patients to placebo before we know for sure?
- What if Abciximab is no different than placebo or even harmful? Can we avoid randomizing 683 patients to a treatment that is no better or worse than placebo?

Interim monitoring can help in both these cases

- The investigators planned to take up to two interim looks using the $LD(OF)$ spending function. The three looks will be equally spaced
- Use East to compute the maximum sample size for this group sequential design

Summary of Three-Look Design

Cytel

	Des8
Mnemonics	PN-2S-DI
Test Parameters	
Trial Type	Superiority
No. of Looks	3
Test Type	2-Sided
Variance	Unpooled Estimate
Specified α	0.05
Specified Power	0.8
Attained Power	0.8
Model Parameters	
Proportion under Control (π_c)	0.15
Proportion under Treatment (π_t)	0.1
Diff. in Prop. ($\pi_t - \pi_c$)	-0.05
Allocation Ratio (n_t/n_c)	1
Boundary Parameters	
Efficacy Boundary	LD (OF)
Spacing of Looks	Equal
Sample Size	
Maximum	1384
Expected Under H0	1378.32
Expected Under H1	1182.678
Completers	
Expected Under H0	1378.32
Expected Under H1	1182.678

Detailed Properties of Three-Look Design



Look #	Info. Fraction (n/n_max)	Sample Size (n)	Cumulative α Spent	Boundaries		Incr. Boundary Crossing Prob.			
				Efficacy Z		Under H0		Under H1	
				Upper	Lower	Upper	Lower	Upper	Lower
1	0.333	461	0.0002	3.7118	-3.7118	0.0001	0.0001	0	0.0186
2	0.667	923	0.0121	2.5109	-2.5109	0.006	0.006	0	0.3995
3		1	0.05	1.9931	-1.9931	0.0189	0.0189	0	0.3821

- The sum of incremental boundary crossing probabilities under $[H_0 : \delta = 0]$ is 0.05, the type-1 error
- The sum of incremental boundary crossing probabilities under $[H_1 : \delta = -0.05]$ is 0.8, the power

Exercise: Expected Sample Size



Look #	Info. Fraction (n/n_{\max})	Sample Size (n)	Cumulative α Spent	Boundaries		Incr. Boundary Crossing Prob.			
				Efficacy Z		Under H_0		Under H_1	
				Upper	Lower	Upper	Lower	Upper	Lower
				3.7118	-3.7118	0.0001	0.0001	0	0.0186
1	0.333	461	0.0002	3.7118	-3.7118	0.0001	0.0001	0	0.0186
2	0.667	923	0.0121	2.5109	-2.5109	0.006	0.006	0	0.3995
3	1	1384	0.05	1.9931	-1.9931	0.0189	0.0189	0	0.3821

- ① Let n_1 , n_2 , and n_3 be the sample sizes at Looks 1, 2, and 3, respectively.
- ② Let p_1 , p_2 , and p_3 be the incremental boundary crossing probabilities under H_1 at Looks 1, 2, and 3, respectively.
- ③ Write down an expression to calculate the expected sample size, from the variables: n_1 , n_2 , n_3 , p_1 , p_2 , p_3 . Confirm by calculation.

1 Introduction

2 Statistical Methods for Group Sequential Designs

- Distribution theory
- Efficacy stopping rules
- Power and sample size calculations
- Example: The CAPTURE Trial
- **Futility stopping rules**

3 Survival Designs in East

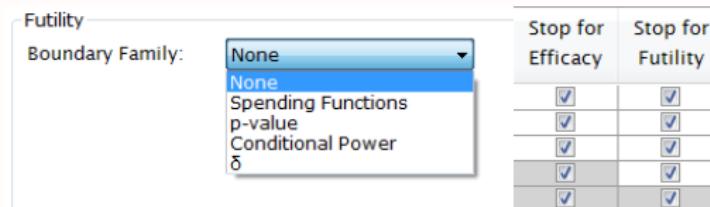
- Example: The JUPITER Study
- Calculating the required number of events
- Handling drop-outs, variable accrual and non-constant hazards
- Non-proportional hazards

4 Workshops 1 - 2 (Group Sequential)

Futility Boundaries in East

Cytel

- An efficacy boundary offers possibility of early stopping and savings in sample size if H_1 is true, but what if H_0 is true?
- Need to add a futility stopping rule.
- Many options for futility boundaries in East (for one-sided test type):



The β -Spending Function Approach

Cytel

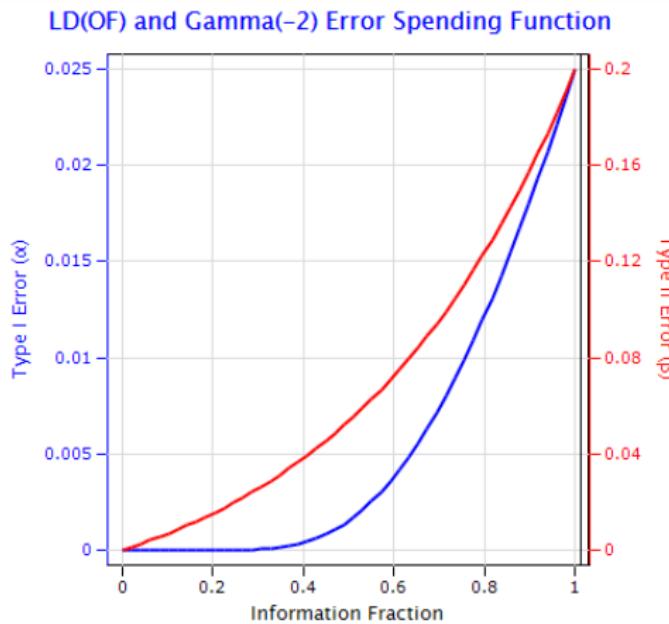
(Pampallona, Tsiatis, and Kim, 2001)

- Just as we use an α -spending function to generate efficacy boundaries that preserve type-1 error, we can also use a β -spending function to generate futility boundaries that control type-2 error
- The probability of crossing the efficacy boundary under H_0 is α
- The probability of crossing the futility boundary under H_1 is β
- We force the two boundaries to meet at the last look to ensure that either the null or the alternative hypothesis is rejected

α -Spending and β -Spending

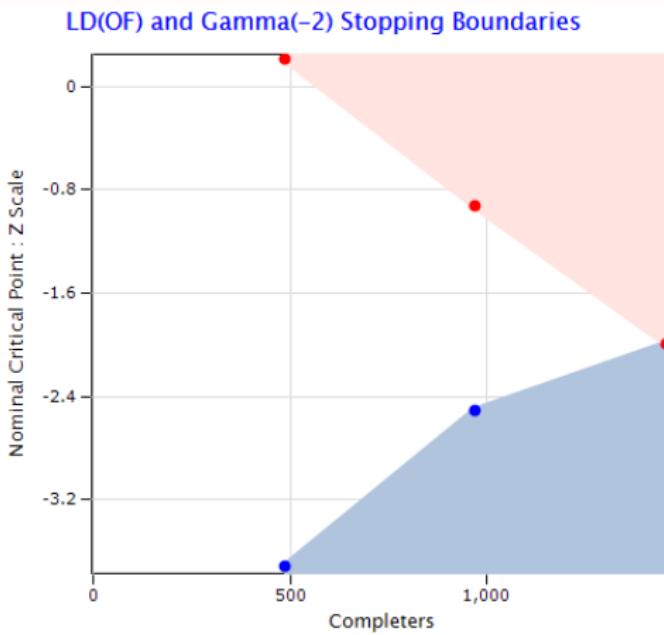
Cytel

For the CAPTURE trial, we continue to use the $LD(OF)$ spending function for α , but spend β using the $\gamma(-2)$ spending function



Both Efficacy and Futility Boundaries

- The two boundaries meet at $l_3 = u_3 = -1.993$, exactly as in the efficacy boundary only design
- Thus the futility boundary can be safely ignored (**Non-Binding**) if crossed at an early look without inflating the study's type-1 error



Savings under both H0 and H1

Cytel

- This design protects the sample size **both** if $\delta = 0$ and if $\delta = -0.05$

Des11	
Mnemonics	PN-2S-DI
Test Parameters	
Trial Type	Superiority
No. of Looks	3
Test Type	1-Sided
Variance	Unpooled Estimate
Specified α	0.025
Attained α	0.023
Specified Power	0.8
Attained Power	0.8
Model Parameters	
Proportion under Control (nc)	0.15
Proportion under Treatment (nt)	0.1
Diff. in Prop. (nt - nc)	-0.05
Allocation Ratio (nt/nc)	1
Boundary Parameters	
Efficacy Boundary	LD (OF)
Futility Boundary	Gm (-2) (NB)
Spacing of Looks	Equal
Sample Size	
Maximum	1455
Expected Under H0	848
Expected Under H1	1174.624
Completers	
Expected Under H0	848
Expected Under H1	1174.624

- Why is the attained α lower than the specified α ?

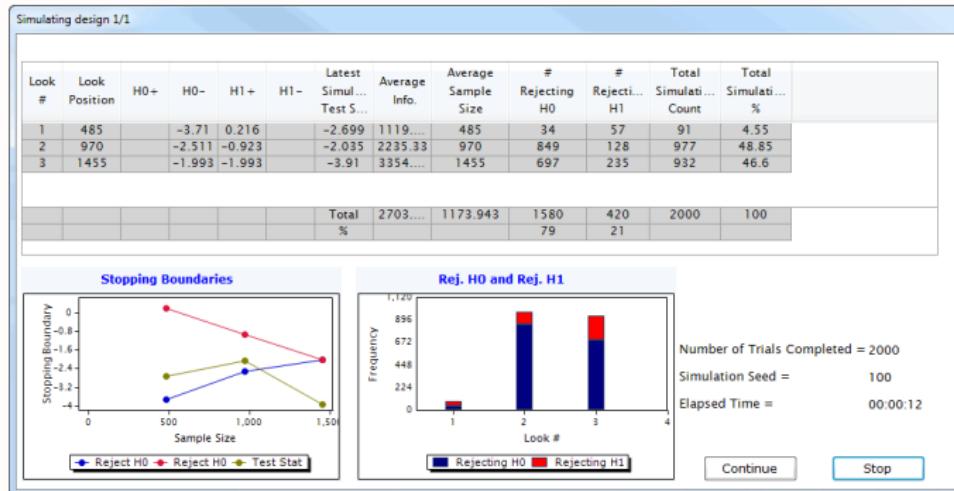
The Role of Simulations

Cytel

- ① Sensitivity analysis for changes in assumptions (eg., treatment effect δ or nuisance parameters)
- ② Breakdown of asymptotics when samples are small
- ③ Operating characteristics in complex survival designs, adaptive sample size re-estimation, or dose selection trials, are possible only via simulation
- ④ Visual communication of properties and variability of study design

Demo: Simulate the CAPTURE Trial

Cytel



Interim Monitoring of CAPTURE



The Capture trial was designed for three equally spaced looks with $n_{\max} = 1383$ and stopping boundaries derived from the LD(OF) spending function. But the trial was actually monitored with unequal spacing as shown below:

Look j	N_j	$(N_j/1383)$ t_j	Resp. Plcbo	Resp. Abcix	Wald Z_j	Old c_j	New c_j
1	350	0.253	30/175	14/175	-2.6046	-3.71	-4.3048
2	700	0.506	55/353	37/347	-1.9332	-2.5114	-2.9429
3	1050	0.759	84/532	55/518	-2.485	-1.993	-2.3426

Look 3 was unplanned and occurred before the planned end of the trial

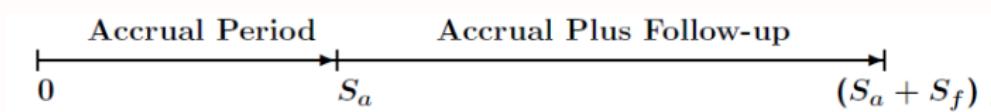
Demo: Monitor the CAPTURE Trial

Cytel

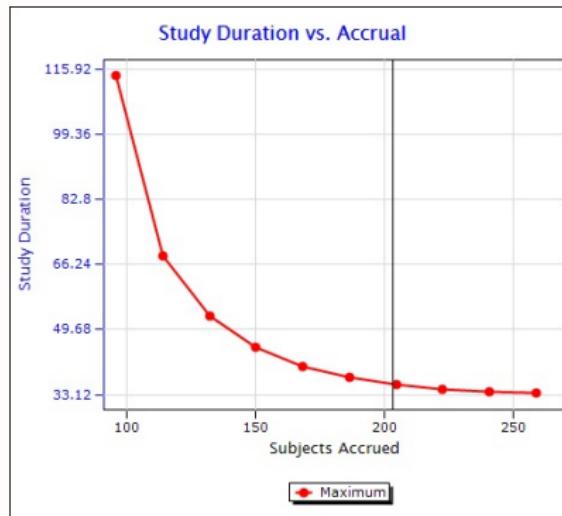
Edit Interim Data								Capture: 3-look-fut:Interim Monitoring			
Look #	Information Fraction	Cumulative Sample Size	Test Statistic	δ	Standard Error	Efficacy	97.5% RCI for δ		Repeated p-value	CP	Predictive Power
							Upper	Lower			
1	0.2529	350	-2.6046	-0.0914	0.0351	-4.3061	0.0597	-1	0.1541	0.9999	0.9706
2	0.5058	700	-1.9332	-0.0492	0.0254	-2.9439	0.0257	-1	0.1162	0.8538	0.7743
3	0.7587	1050	-2.485	-0.0517	0.0208	-2.3434	-0.0029	-1	0.0183	NA	NA

Designing and Simulating for Survival Endpoints

- For studies with survival or time-to-event endpoints, the asymptotic distribution theory and the derivation of stopping boundaries remains the same
- There are however some special considerations
 - The information is directly proportional to the number of events
 - Thus the number of events, not the number of patients, determines the power of the study
 - If study duration is not fixed, there is a trade-off between sample size and study duration.



- If we recruit more patients to the study, we obtain the required number of events sooner and the total study duration is reduced



1 Introduction

2 Statistical Methods for Group Sequential Designs

- Distribution theory
- Efficacy stopping rules
- Power and sample size calculations
- Example: The CAPTURE Trial
- Futility stopping rules

3 Survival Designs in East

- Example: The JUPITER Study
- Calculating the required number of events
- Handling drop-outs, variable accrual and non-constant hazards
- Non-proportional hazards

4 Workshops 1 - 2 (Group Sequential)

Example: The JUPITER study

Cytel

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812

NOVEMBER 20, 2008

VOL. 359 NO. 21

Rosuvastatin to Prevent Vascular Events in Men and Women with Elevated C-Reactive Protein

Paul M Ridker, M.D., Eleanor Danielson, M.I.A., Francisco A.H. Fonseca, M.D., Jacques Genest, M.D., Antonio M. Gotto, Jr., M.D., John J.P. Kastelein, M.D., Wolfgang Koenig, M.D., Peter Libby, M.D., Alberto J. Lorenzatti, M.D., Jean G. MacFadyen, B.A., Børge G. Nordestgaard, M.D., James Shepherd, M.D., James T. Willerson, M.D., and Robert J. Glynn, Sc.D., for the JUPITER Study Group*

- Justification for the Use of Statins in Prevention: an Intervention Trial Evaluating Rosuvastatin (JUPITER) examined the question of whether treatment with 20 mg of rosuvastatin daily, as compared with placebo, would reduce the rate of first major cardiovascular events (Ridker et al., 2008)

Example: The JUPITER study (cont.)



- JUPITER was a randomized, double-blind, placebo-controlled, multicenter trial conducted between 2003 and 2008 by AstraZeneca at 1315 sites in 26 countries
- Composite primary endpoint: occurrence of a first major cardiovascular event, defined as nonfatal myocardial infarction, nonfatal stroke, hospitalization for unstable angina, an arterial revascularization procedure, or confirmed death from cardiovascular causes

Example: The JUPITER study (cont.)



- Designed for statistical power of 90% to detect a 25% reduction in the rate of the primary end point, with a two-sided significance level of 0.05
- We are interested in a 25% reduction in the rate of the primary endpoint, i.e. a hazard ratio $\lambda_t/\lambda_c = 0.75$; the effect size is thus $\delta = -\ln(\lambda_t/\lambda_c) = 0.2877$
- Baseline hazard rate in placebo arm of 0.0077
- Study to complete in 7.5 years, with 4 years accrual and 3.5 years of follow-up
- Two interim analyses are planned with LD(OF) spending function defined boundaries at 37.5% and 75% of the information
- How do we design such a trial?

1 Introduction

2 Statistical Methods for Group Sequential Designs

- Distribution theory
- Efficacy stopping rules
- Power and sample size calculations
- Example: The CAPTURE Trial
- Futility stopping rules

3 Survival Designs in East

- Example: The JUPITER Study
- **Calculating the required number of events**
- Handling drop-outs, variable accrual and non-constant hazards
- Non-proportional hazards

4 Workshops 1 - 2 (Group Sequential)

Maximum Number of Events



- To convert information into number of events, we use Schoenfeld's approximation (Biometrika, 1981)

$$D_{\max} = \frac{I_{\max}}{p(1-p)}$$

where D_{\max} is the maximum number of events required, p is the proportion of these events that occur on the treatment arm, and $(1 - p)$ is the proportion of these events that occur on the control arm

- If the null hypothesis holds, we can set $p = r$, where r is the fraction of patients randomized to the treatment arm. For balanced randomization where $p = 1/2$ we have

$$D_{\max} = 4I_{\max}$$

Equating Required and Expected Number of Events to Estimate Sample Size and Study Duration



- For a K look design, if D_{\max} events are desired, then the study must remain open until a boundary is crossed or calendar time τ_K where τ_K satisfies the relationship

$$D_{\delta_1}(\tau_K) = D_{\max}$$

where $D_{\delta_1}(\tau_K)$ is the expected number of events at time τ_K

- We can derive an expression for $D_{\delta_1}(\tau_K)$ from exponential assumptions and thereby estimate sample size and study duration

JUPITER Study Design in East



- We can use the Logrank Test Given Accrual Duration and Study Duration design in East to obtain a 3-look GSD for the JUPITER trial

Wbk1:Des1	
Mnemonics	SU-2S-LRSD
Test Parameters	
Trial Type	Superiority
No. of Looks	3
Test Type	2-Sided
Specified α	0.05
Specified Power	0.9
Attained Power	0.9
Model Parameters	
Hazard Ratio (Alt.)	0.75
Var (Log HR)	Null
Allocation Ratio (Int/nc)	1
Boundary Parameters	
Efficacy Boundary	LD (OF)
Spacing of Looks	Unequal
Sample Size	
Maximum	14229
Expected Under H0	14229
Expected Under H1	14229
Events	
Maximum	517
Expected Under H0	514,406
Expected Under H1	414,514
Accrual Duration	
Maximum	4
Expected Under H0	4
Expected Under H1	4
Study Duration	
Expected Under H0	6.787
Expected Under H1	6.397

- East tells us that the required number of events is $D_{\max} = 517$ as previously determined, and that we will require 14,229 subjects accrued over 4 years

Early Stopping and Expected Study Duration



Events, Sample Size, Dropouts, Pipeline and Analysis Times: Look by Look (Under H1)

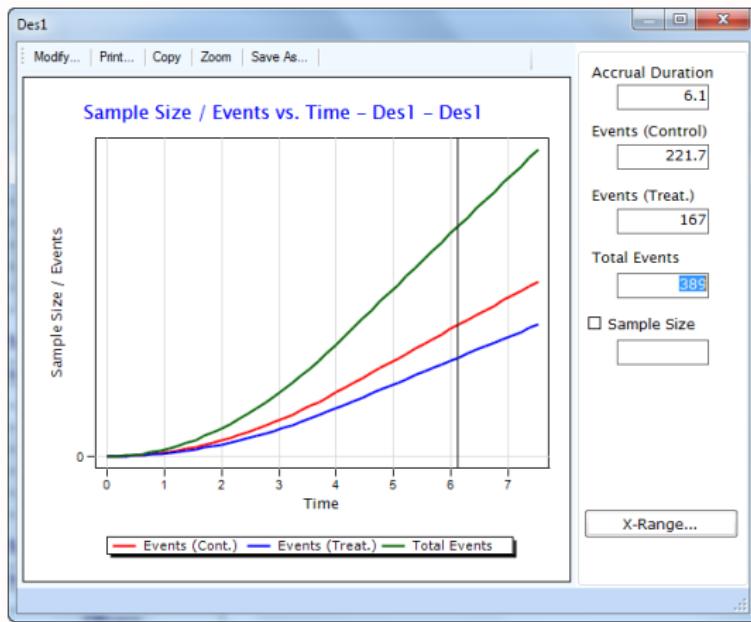
Look #	Info. Fraction (s/s_max)	Sample Size (n)	Events (s)	Pipeline (n-s)	Analysis Time	Incr. Boundary Crossing Prob.	
						Under H1	
						Efficacy	
						Upper	Lower
1	0.375	14229	194	14035	4.042	0	0.0703
2	0.75	14229	388	13841	6.109	0	0.6185
3	1	14229	517	13712	7.5	0	0.2114

$$\begin{aligned} E(\text{Study Duration} \mid H_1) &= 0.0703 \times 4.042 \\ &\quad + 0.6186 \times 6.109 \\ &\quad + 0.2114 \times 7.5 \\ &\quad + (1 - 0.0703 - 0.6186 - 0.2114) \times 7.5 \\ &= 6.397 \end{aligned}$$

Planning for DMC Meetings

Cytel

- The Events vs. Time Chart is useful in planning for interim analyses & DMC meetings; eg., after the second look, **389 events** should occur roughly **6.1 years** into the study assuming the treatment effect is δ_1



1 Introduction

2 Statistical Methods for Group Sequential Designs

- Distribution theory
- Efficacy stopping rules
- Power and sample size calculations
- Example: The CAPTURE Trial
- Futility stopping rules

3 Survival Designs in East

- Example: The JUPITER Study
- Calculating the required number of events
- **Handling drop-outs, variable accrual and non-constant hazards**
- Non-proportional hazards

4 Workshops 1 - 2 (Group Sequential)

- **Dropouts:** 5% per year in both the treatment and placebo arms

Piecewise Dropout Information

of Pieces: Input Method:

Period #	By Time	Prob. of Dropout (Control)	Prob. of Dropout (Treatment)
1	1.000	0.05	0.05

Note: Period 1 hazard rates apply after time 1.

- Like events: Dropouts assumed exponential, with corresponding hazard rates
- Like Cumulative % survival: Cumulative % dropout calculated for patient time (from study entry), not calendar time (from study start)

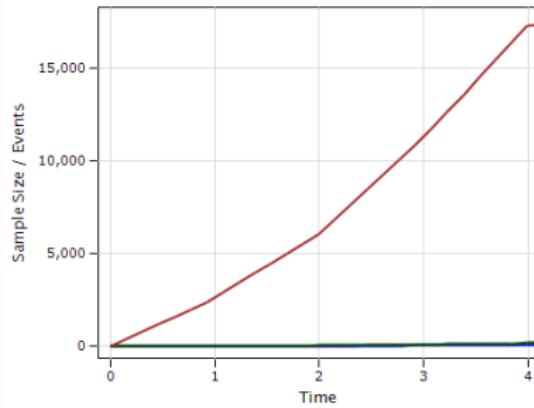
JUPITER Design Accruals

Cytel

- Non constant accrual (piece-wise linear): 15% by end of year 1; 35% by end of year 2; 65% by end of year 3

Accrual Info

Accrual Duration:	4	Study Duration:	7.5
# of Accrual Periods:	4		
Period #	By Time	Cum. % Accrued	
1	1.000	15.000	▲
2	2.000	35.000	▼
3	3.000	65.000	



Revisited JUPITER Design in East



Events, Sample Size, Dropouts, Pipeline and Analysis Times: Look by Look (Under H1)

Look #	Info. Fraction (s/s_max)	Sample Size (n)	Events (s)	Dropouts (d)	Pipeline (n-s-d)	Analysis Time	Incr. Boundary Crossing Prob.	
							Under H1	
							Efficacy	
							Upper	Lower
1	0.375	17344	194	1478	15672	4.131	0	0.0703
2	0.75	17344	388	2955	14001	6.074	0	0.6185
3	1	17344	517	3938	12889	7.5	0	0.2114

- Note that since we have not changed the effect size δ_1 that we are powering the trial for, $D_{\max} = 517$ has not changed. However, the dropouts and slow starting accrual means we now need **17,344 patients** if we want to finish the study within 7.5 years

1 Introduction

2 Statistical Methods for Group Sequential Designs

- Distribution theory
- Efficacy stopping rules
- Power and sample size calculations
- Example: The CAPTURE Trial
- Futility stopping rules

3 Survival Designs in East

- Example: The JUPITER Study
- Calculating the required number of events
- Handling drop-outs, variable accrual and non-constant hazards
- Non-proportional hazards

4 Workshops 1 - 2 (Group Sequential)

- All survival designs in East assume the proportional hazards assumption holds
- Non-proportional hazards may arise in clinical trials for many reasons (e.g., delayed treatment effects)
- In East, simulated data may be generated according to non-proportional hazards
- Thus, it is possible to evaluate the impact of non-proportional hazards on the power of a study

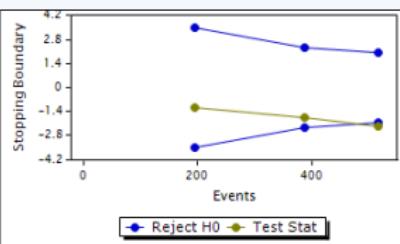
Demo: Simulate the JUPITER Trial

Cytel

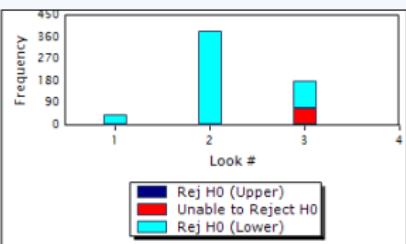
Simulating design 1/1

Look #	Look Position	H0+	H0-	H1+	H1-	Latest Simul... Test S...	Average Events	# Rejecting Up(H0+)	# Rejecting Low(H0-)	# Unable to Reject...	Total Simulation Count	Total Simulation %
1	194	3.477	-3.477			-1.201	194	0	41		41	6.833
2	388	2.342	-2.342			-1.745	388	0	381		381	63.5
3	517	2.012	-2.012			-2.24	517	0	110	68	178	29.667
							Total	413.013			532	68
							%		0	88.667	11.333	600
												100

Stopping Boundaries



Rej. H0 and Unable to Rej. H0



Number of Trials Completed = 600

Simulation Seed = 100

Elapsed Time = 00:01:01

Continue

Stop

AstraZeneca Stops Cholesterol Drug Trial Because Of Promising Results

01 Apr 2008

[Click to Print](#)

Pharmaceutical giant AstraZeneca announced yesterday, Monday 31st March, it was stopping the JUPITER clinical trial on its cholesterol-busting drug Crestor (rosuvastatin calcium) because early findings showed that the drug reduced deaths and risk of heart problems in patients compared to placebo.



A press statement on the company's website said that the trial's Steering Committee and also the Independent Data Monitoring Board recommended that JUPITER be stopped early because there was "unequivocal evidence of a reduction in cardiovascular morbidity and mortality amongst patients" who took Crestor compared to patients who took placebo.

Monitoring the JUPITER Study



- At the first interim analysis in September 2007, the efficacy boundary was crossed. However, the DMC voted to continue the trial for an additional 6 months. Thus, the next interim analysis in March 2008 was not originally planned.
- Suppose that in March 2008, 142 events had been observed on the Rosuvastatin arm against 251 events on the placebo arm
- The estimated hazard ratio was 0.56 so that $\hat{\delta}_1 = \ln(0.56)$. The standard error can be estimated as $se(\hat{\delta}) = I_1^{-1/2} = \sqrt{4/D_1}$ where $D_1 = 142 + 251 = 393$
- The JUPITER study was terminated early for overwhelming efficacy in the primary endpoint: HR 0.56 (95% CI 0.46-0.69) $P=<0.00001$ (Ridker, 2008)

Demo: Monitor the JUPITER Trial



Test Statistic Calculator

Editing Look #1

Set Current Look as Last

Cumulative Events:

Input for Survival end point

Estimate of δ :
 $\delta = \ln(\lambda_t / \lambda_c)$

Standard Error of Estimate of δ :

Output

Test Statistic:

Demo: Monitor the JUPITER Trial

Cytel

Enter Interim Data										Interim Monitoring: Des2
Look #	Information Fraction	Cumulative Events	Test Statistic	δ	Standard Error	Efficacy		Repeated 95% CI for δ		Repeat ... p-value
	Upper	Lower	Upper	Lower						
1	0.76	393	5.747	0.58	0.101	2.321	-2.321	0.814	0.346	0



Final Inference	
Final Outputs at Look #	1
Adj. p-value	0
Adj. Pt. Est. for δ	0.58
Adj. 95% CI for δ	
Upper Confidence Bound	0.778
Lower Confidence Bound	0.382

East provides great flexibility, but in practice, most group sequential designs have a few common features:

- ① **Conservative α -spending:** Discourages termination during early adjustments (patient treatment, trial management). Also, wait until enough safety data have been collected

- ② **Non-binding futility:** DMCs may wish to continue a trial even after the futility boundary has been crossed (eg., to follow secondary endpoint)

General Recommendations for Design (cont.)



- ③ **Delay time of first interim:** The first interim analysis is usually performed after enough data to provide stable estimates
- ④ **Anticipate and avoid over-runs:** The final interim analysis should be performed while it is still possible to reduce the sample size.
- ⑤ **Fewer interim analyses:** Usually between 2 and 3. (Large trials with mortality endpoints might have as many as 6)

Questions?

Cytel



Design of a Phase 3 Clinical Trial in Hypertension

- You work for Lifebeat Inc., a multinational pharmaceutical company which has in-licensed a product called ANTIHYPE®
- Your team has been tasked with developing a phase 3 clinical study design for this product in the essential hypertension indication to serve in a New Drug Application submission for the US Food and Drug Administration regulatory agency.
- Your team is told that the phase 3 trial must be completed within **two years**

- Primary endpoint: Change from baseline to Week 6 (1.5 months) in the 24-hour mean systolic blood pressure by ambulatory blood pressure monitor
- Based on preliminary phase 2 study results and the commercial team's feedback, you should power your trial for a treatment effect of $\delta = -1.5$ mm Hg when compared to placebo in a 1 : 1 allocation
- A common standard deviation of $\sigma = 6$ mm Hg can be assumed
- Enrollment rate can be assumed to be 40 patients per month (Click "Include Options", then "Accrual/Dropout Info")
- Question 1: Creating a fixed-sample design (Design 1) with 90% power and two-sided $\alpha = 0.05$ type-1 error. What is the sample size in each group?
(Note that the sample size provided by East is the **total** sample size)

$$N_{Treatment} : \underline{\hspace{2cm}} \quad N_{Placebo} : \underline{\hspace{2cm}}$$

- Question 2 : How long will the study last?

Three-Look Group Sequential Design



- Convert Design 1 to a 3-look group sequential study design with the default $LD(OF)$ spending function (Design 2)
- Question 3:** What is the maximum total sample size of this 3-look group sequential design? (And what is the penalty in maximum sample size for Design 2, compared to Design 1)

N_{\max} : _____

- Examine the boundaries chart, the spending function chart, and the power chart by clicking on their respective icons.
- Question 4:** From the spending function chart, what amount of alpha would have been spent by information fraction = 0.5?

alpha spent : _____

Exploring the Group Sequential Design Properties



- Observe that the boundaries chart can be displayed on various scales. The delta scale is particularly useful for communicating with clinicians (NB: contrast with design delta)
- Question 5: What is final critical value on the Z-scale and delta-scale for the 3-look design (cf. for the 1-look design)?

critical values : _____

- Question 6: What is the expected savings in sample size under H_1 for Design 2 compared to Design 1?

Savings: _____

- Do you know how to calculate the expected sample size by hand?
- Question 7: If the treatment effect is indeed $\delta_1 = -1.5$, with what sample size are you most likely to stop the trial and reject H_0 ? With what probability does this occur?

Sample Size : _____ probability : _____

Sensitivity of the Design to Number and Spacing of Looks



- Suppose you want to take only 2 looks, not 3. And suppose you want to space them unequally, after 75% and 100% of the information has accrued (Design 3).
- Question 8: What is the maximum sample size of this 2-look group sequential design (Design 3)?

Sample Size : _____

Adding in a Futility Boundary



- Edit the original 3-look equally-spaced design (Design 2). Change to a one-sided with $\alpha = 0.025$. Add a non-binding futility boundary, with an aggressive β -spending function: $\gamma(-2)$. This will be Design 4
- Question 9: What is the maximum total sample size of this design?

$N_{\max} :$ _____

- Question 10: At the first look, what is the futility boundary on the CP delta1 scale?

CP boundary : _____

- Let's simulate Design 4 under different conditions.
- Question 11:** Verify the operating characteristics of this design by simulating the study under $H_0 : \delta = 0$ and under $H_1 : \delta = -1.5$. Use the same fixed seed for each simulation.

$\hat{\alpha} : \underline{\hspace{2cm}}$ $1 - \hat{\beta} : \underline{\hspace{2cm}}$

- Question 12:** Investigate each of the following scenarios (1) $\delta = -0.5$; (2) $\delta = -1.8$; (3) $\delta = -1.5; \sigma = 9.5$. Provide below the estimated power, the probability of stopping early for efficacy, and the probability of stopping for futility for each of these settings. You may need to complete some calculations by hand.

Scenario	Power	Pr(early efficacy)	Pr(early futility)
(1) $\delta = -0.5$			
(2) $\delta = -1.8$			
(3) $\sigma = 9.5$			

Design of a Phase 3 Survival Study in Small Cell Lung Cancer

- You now work for Cancergene Inc., a multinational pharmaceutical company which has developed an in-house product called ONCOBLAST®
- Your team has been tasked with developing a phase 3 clinical study design for this product in the small cell lung cancer indication to serve in a Market Authorisation Application submission for the European Medicines Agency regulatory agency.
- Your final safety database must include at least 500 patients.
- Your team is told that the phase 3 trial must be completed within five years.

- Primary endpoint: overall survival (OS)
- Based on early phase results, the medical literature, and the commercial team's feedback, you should power your trial for a **30% improvement** in median survival time for ONCOBLAST when administered with Standard of Care (SOC) relative to the SOC alone
- Median survival time on the SOC is assumed to be **11 months**
- The study must enroll within **3 years** and be completed within **5 years**. Note that you are using **Logrank Given Accrual Duration and Study Duration**.
- **Question 1:** Create a fixed-sample design (Design 1) with **90% power** and **one-sided $\alpha = 0.025$** type-1 error. How many events do you need to observe?

D : _____

- **Question 2:** How many subjects do you need to enroll in the 3 year period to complete the study within 5 years?

Three-Look Group Sequential Design



- Convert this design to a 3-look group sequential study design. Choose the Lan-Demets (OF) boundary family with equal spacing of the looks (Design 2)
- Question 3: What is the maximum number of events needed to power this 3-look group sequential design?

D_{\max} : _____

- Question 4: How many patients will you need to enroll to achieve study completion within 5 years?

N_{\max} : _____

Additional Assumptions

Cytel

- Add in a Lan-Demets(OF) non-binding **futility boundary** (Design 3)
- Plan for **3% dropout per year in each arm** [NB: Use “Dropout Rates”]
- Also, accrual is slower than expected: **20% by end year 1; 50% by end year 2; and the rest by end year 3**
- **Question 5:** With these added assumptions, what is the **maximum number of events** needed to power this trial?

D_{\max} : _____

- **Question 6:** How many patients will you need to enroll to achieve study completion within 5 years?

N_{\max} : _____

Exploration of the Design Characteristics



- Question 7: In Design 3, select Charts:Power vs. Sample size. Enter 500 in the Sample Size field. What power is achieved at 500 patients, the minimum needed for your safety database?

Power : _____

- Question 8: In the table from Design Details, identify the timing from study start of your two interim analyses if the treatment effect is indeed δ_1 . Alternatively, find these times from the events/accruals vs. time Chart.

Look #1 : _____ Look #2: _____

- In the 'Simulation Control Info' tab, set a fixed random number seed.
- Question 9: Simulate Design 3 under the following three hypotheses:
(1) $HR = 0.9$; (2) $HR = .5$, and (3) $HR = 1$. What is the power and expected sample size under each scenario?

Scenario	Power	E(Sample Size)
(1) $HR = .9$		
(2) $HR = 0.5$		
(3) $HR = 1$		

- Now suppose the baseline hazard rate in the control arm remains unchanged, but that $HR = 1.0$ in the first 6 months, yet $HR = 0.72$ thereafter
- Question 10:** Simulate this study under the Design 3. What is the simulated power of the trial?

Power : _____

- Question 11:** How often do we stop early for futility (1st or 2nd look)?

$\text{Pr}(\text{futility})$: _____