Cytel
Shaping the Future of Drug Development

# Group Sequential Designs and Sample Size Re-estimation – Modern Uses

Professor Chris Jennison
University of Bath, UK

*Introduction to Complex Innovative Trial Design Webinar Series*

# Group Sequential Designs and Sample Size Re-estimation – Modern Uses

## *Agenda:*

**Group Sequential Methods:**

- Challenges of Traditional Designs and Benefits of Group Sequential Methods
- Method Theory
- Case Studies
- Design Options and Discussion

**Sample Size Re-estimation**

- Challenges of Traditional Designs and Benefits of Sample-Size Re-estimation Methods
- Method Theory
- Case Studies
- Discussion

**Factors to Consider and Summary**

**Q&A**

**Conclusion**

Cytel


# *Group Sequential Design*

## *Challenges of Traditional Design and Benefits of Advanced Methods*

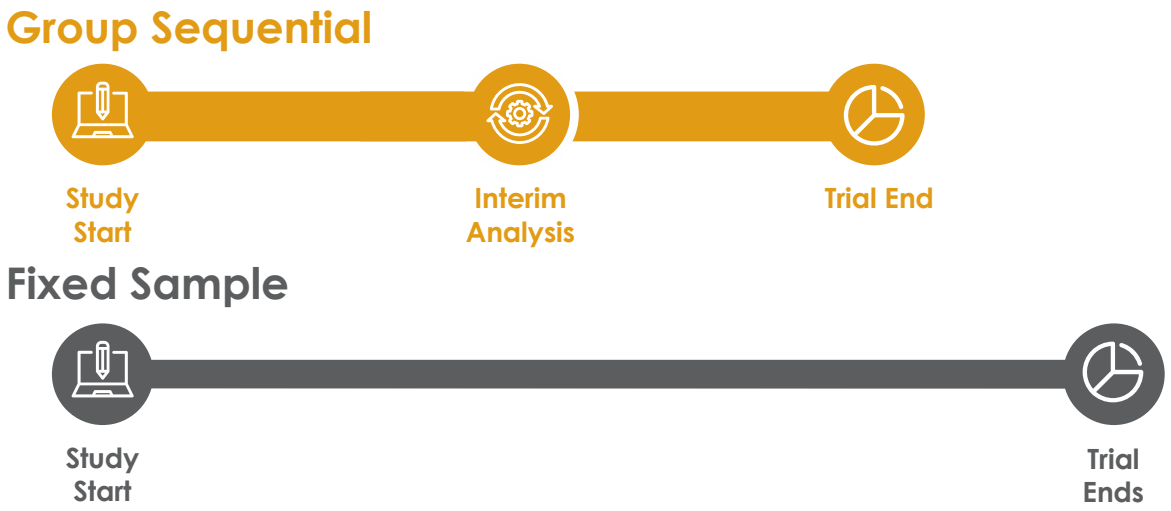
# Why Use Group Sequential Methods?

## I shall focus on the design of Phase 3 trials. These studies are

- A critical step on the way to drug approval
- Large and expensive

## A group sequential design can reduce the length of a Phase 3 trial

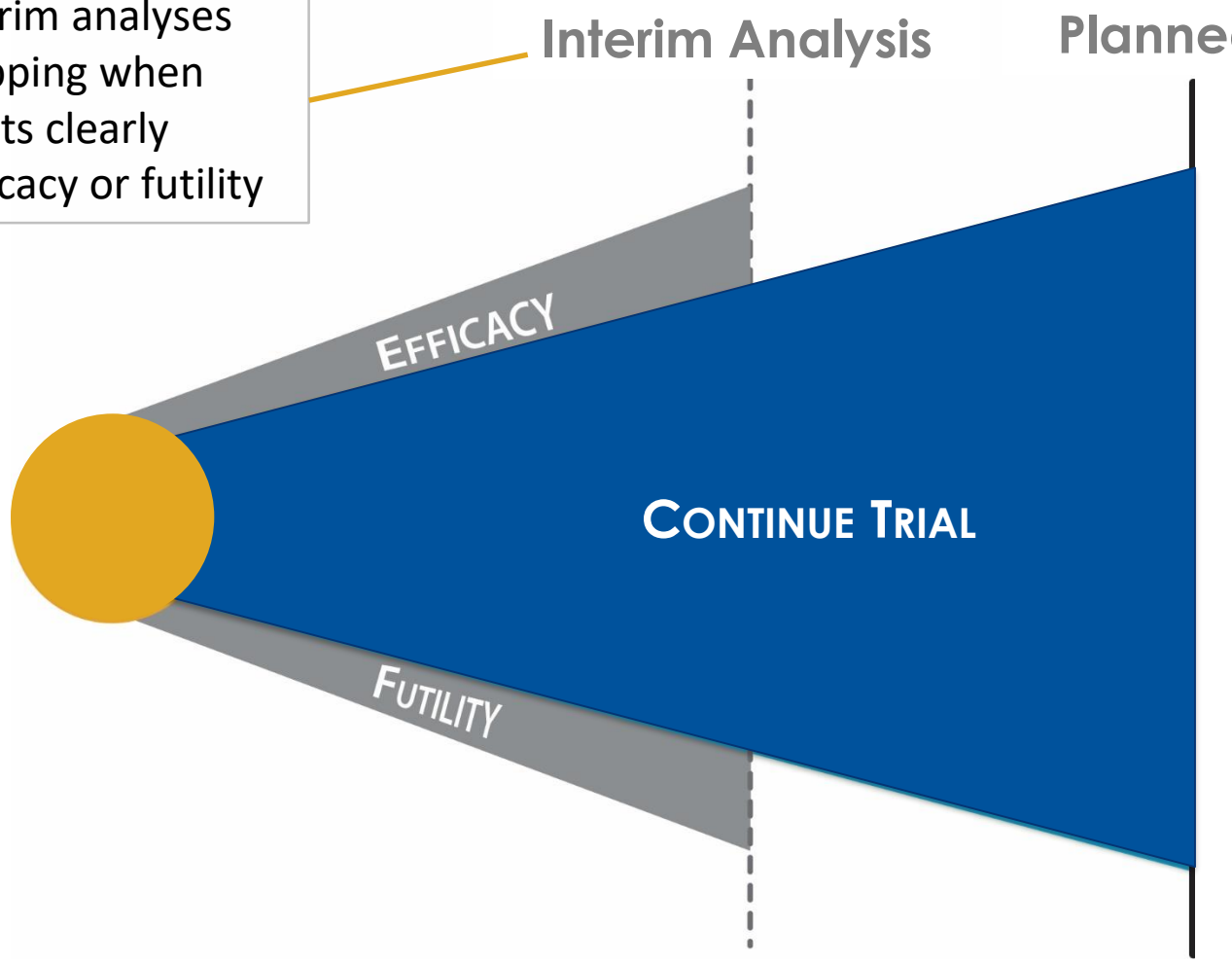
- Saving resources
- Reaching a conclusion sooner




Cytel

# Group Sequential Design Can Reduce Sample Size And Duration, Conserving Resources & Accelerating Development

Pre-planned interim analyses enable early stopping when preliminary results clearly demonstrate efficacy or futility




Designs with 2 or 3 analyses and a maximum sample size 5% or 10% greater than the fixed sample test can **reduce expected sample size by around 30%** versus fixed designs.



Cytel

# What Do Group Sequential Methods Offer?

## Consider A Phase 3 Trial with

- Type I error rate $\alpha = 0.025$,
- Power 0.9 at treatment effect $\theta = 1$

**Fixed Sample Size Design:**
100 patients per treatment

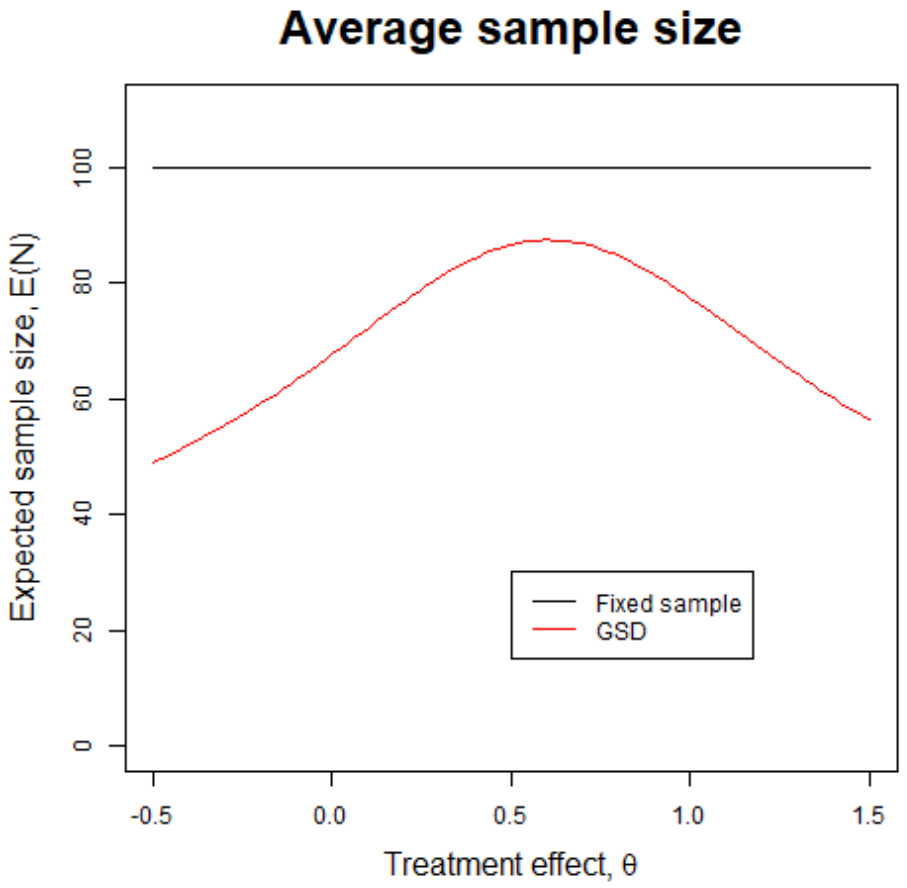
**Group Sequential Trial, 3 Analyses:**
Up to 105 patients per treatment

**Group Sequential Trial has expected sample size E(N):**

66 at treatment effect $\theta = 0$

87 at treatment effect $\theta = 0.5$

77 at treatment effect $\theta = 1$




Potential to reduce patients needed by 30%

Cytel

# Poll Question

**Do you use Group Sequential designs?**

A: No

B: Yes, but only rarely

C: Yes

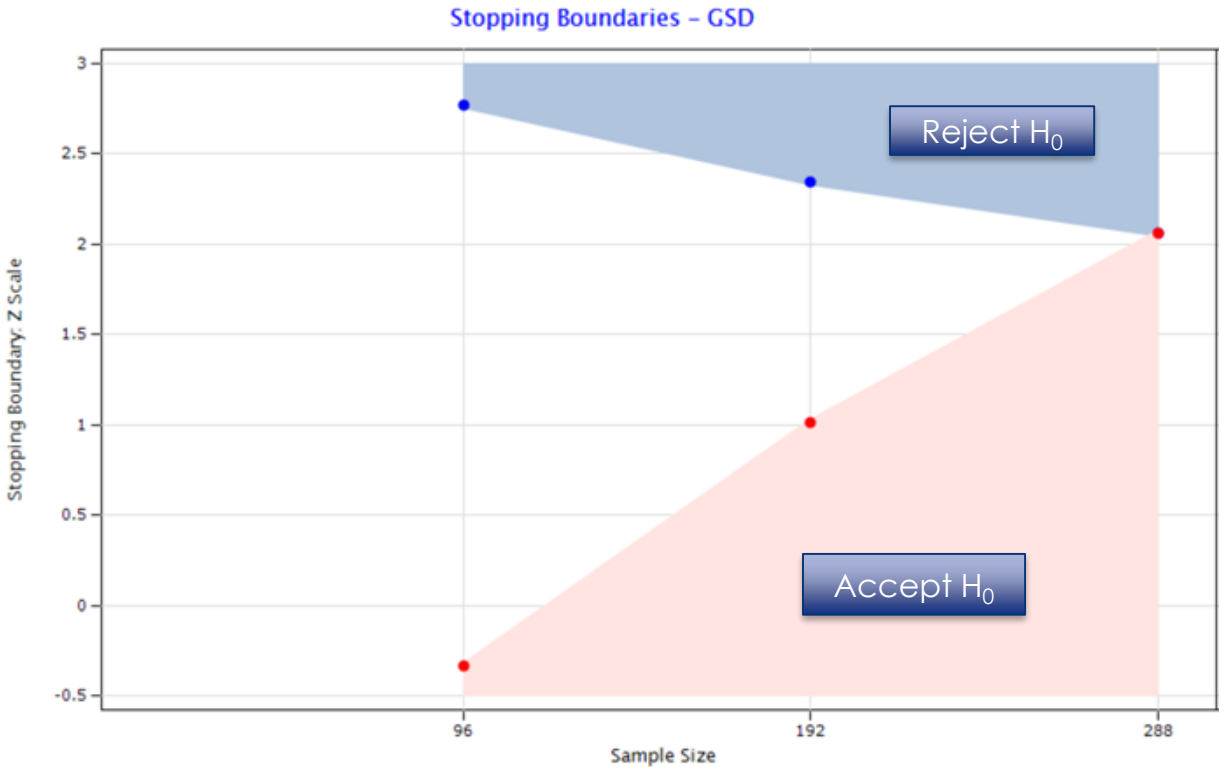


Cytel


# *Group Sequential Design*

## *Method Theory*

# How Do Group Sequential Tests Work?




At each analysis, we calculate the $Z$-statistic for testing $H_0$:

$\theta \leq 0$ vs $\theta > 0$

We stop the trial when this statistic crosses a boundary

Note:

$Z_k$ is a measure at analysis k of the evidence against the null hypothesis $H_0$, which states that the new treatment is no better than the control.

The boundary is chosen so that

- Type I error rate is $\alpha = 0.025$

  Note: Multiple testing can inflate Type 1 error, so the boundary is set to avoid this
- Power is $0.9$ at treatment effect $\theta = \delta$

Cytel

# How Do Group Sequential Tests Work?

**Underpinning Theory**

The joint distribution of the sequence of $Z$-statistics has a standard form for many different response distributions, including survival data.

**Computation**

Type I error probability and power are computed by numerical integration – which is fast and accurate.

**Software**

is available to implement these methods

Cytel

# Error Spending Group Sequential Tests

Sample size or, more generally, observed information, is unpredictable.

In an ***error spending*** design, we define boundaries at each analysis so that the cumulative type I and type II error probabilities are equal to a certain function of the observed information.

Here, the information for the treatment effect, $\theta$, at analysis $k$ is defined as

$$I_k = \frac{1}{Var(\hat{\theta}_k)},$$

where $\hat{\theta}_k$ denotes the estimated treatment effect at this analysis.

Cytel


# *Group Sequential Design*

## *Case Studies*

# Example: A Trial For A Cholesterol Lowering Drug

**To Compare**

Experimental treatment vs placebo control

**Endpoint**

Reduction in serum cholesterol (mg/dL) after 4 weeks

**Responses**

Normally distributed with standard deviation $\sigma = 25$

**Treatment Effect**

$\theta =$ Difference in mean response between treatments

Cytel

# Example: A trial for a cholesterol lowering drug

**First, consider how you would design a fixed sample trial.**

We wish to $H_0$: $\theta \leq 0$ vs $\theta > 0$ with type I error rate $\alpha = 0.025$

We observe responses from $n$ patients on each treatment arm,

compute the $Z$-statistic,

and reject $H_0$ if this is significant at level 0.025.

(For simplicity we assume $\sigma$ known: in practice, we would use a $t$-test.)

In order to achieve power $0.9$ when $\theta = 10$, we find we need a sample size of

$$n_{fix} = 132$$

subjects on each treatment.

Cytel

# Example: A Trial For A Cholesterol Lowering Drug

**In creating a group sequential design, we specify**

- 3 analyses (2 interim analyses, 1 final analysis)
- An error spending design, spending error in proportion to Information$^2$
- A non-binding futility boundary

This design has an "Inflation factor" of $1.093$,
so we shall need up to

$$\mathbf{1.093 \times 132 = 144}$$

patients per treatment arm.

Cytel

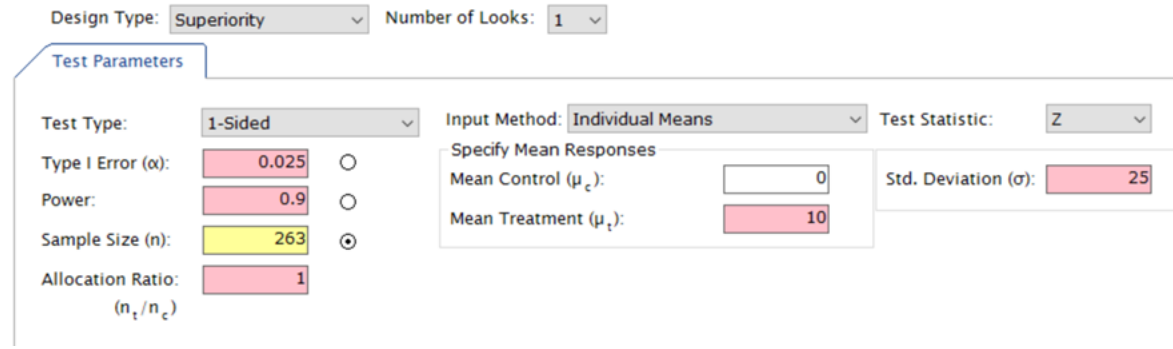
# Demonstration of East Software

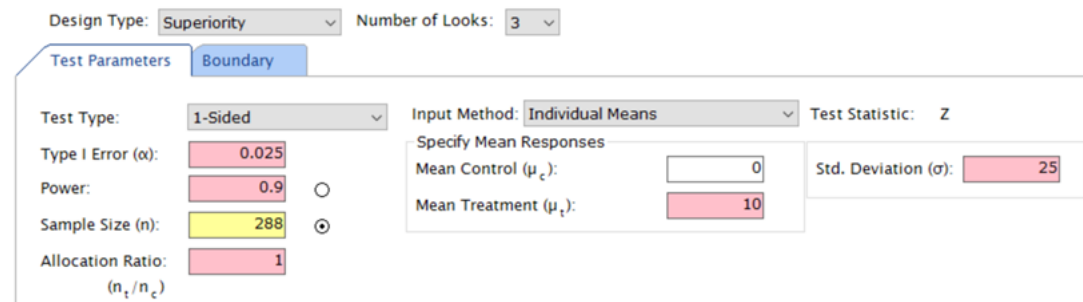
## *Cholesterol Study:*

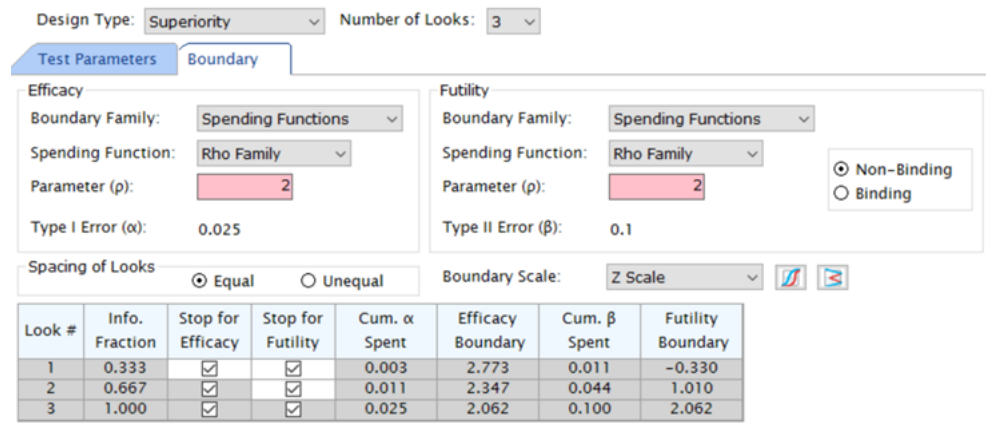
- A fixed sample size design
- Group sequential design
- Interim analyses: Data entry and boundary calculation
- Simulating the group sequential design

Cytel

| | Ex:GSD | Ex:Fixed Sample |
|---|---|---|
| Mnemonic | MN-2S-DI | MN-2S-DI |
| **Test Parameters** | | |
| Design Type | Superiority | Superiority |
| No. of Looks | 3 | 1 |
| Test Type | 1-Sided | 1-Sided |
| Specified α | 0.025 | 0.025 |
| Attained α | 0.024 | |
| Power | 0.901 | 0.9 |
| **Model Parameters** | | |
| Allocation Ratio (nt/nc) | 1 | 1 |
| Input Method | Individual Means | Individual Means |
| Diff. in Means (δ = μt - μc) | 10 | 10 |
| Mean Control (μc) | 0 | 0 |
| Mean Treatment (μt1) | 10 | 10 |
| Std. Deviation (σ) | 25 | 25 |
| Test Statistic | Z | Z |
| **Boundary Parameters** | | |
| Spacing of Looks | Equal | |
| Efficacy Boundary | Rho (2) | |
| Futility Boundary | Rho (2) (NB) | |
| **Sample Size** | | |
| Maximum | 288 | 263 |
| Expected Under H0 | 169.642 | |
| Expected Under H1 | 198.33 | |




| Look # | Info. Fraction | Stop for Efficacy | Stop for Futility | Cum. α Spent | Efficacy Boundary | Cum. β Spent | Futility Boundary |
|---|---|---|---|---|---|---|---|
| 1 | 0.333 | ☑ | ☑ | 0.003 | 2.773 | 0.011 | -0.330 |
| 2 | 0.667 | ☑ | ☑ | 0.011 | 2.347 | 0.044 | 1.010 |
| 3 | 1.000 | ☑ | ☑ | 0.025 | 2.062 | 0.100 | 2.062 |


# ***Group Sequential Design***

## *Design Options and Discussion*

# How Rapidly Should We "Spend" Alpha? The $\rho$-Family Of Error Spending Functions

**We fix the parameter $\rho$ and maximum number of analyses $K$.**

Then, at analysis $k$, with observed information $I_k$, boundaries are set so that

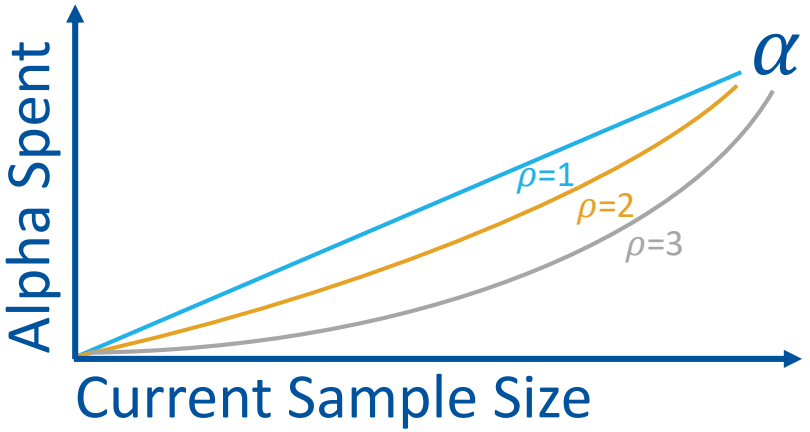
$$\text{Cumulative type I error probability} = (I_k/I_{max})^{\rho}\alpha \quad \text{(under } \theta = 0\text{)},$$

$$\text{Cumulative type II error probability} = (I_k/I_{max})^{\rho}\beta \quad \text{(under } \theta = \delta\text{)}.$$

We set $I_{max} = RI_{fix}$, so boundaries meet at analysis $K$ if we observe

$$I_1 = \frac{1}{K}I_{max}, \cdots, I_{K-1} = \frac{K-1}{K}I_{max}, \; I_K = I_{max}.$$

Here, $I_{fix}$ is the information for a fixed sample test and the "Inflation factor" $R$ depends on $\alpha$, $\beta$, $\rho$ and $K$.




Note: The more aggressively you spend alpha early on in the study, the greater potential reduction in expected sample size for the study.

Cytel

# Choosing $\rho$ & $K$

Properties of $\rho$-family designs with non-binding futility boundaries, type I error rate $\alpha = 0.025$, and power 0.9 when $\theta = \delta$.

***For ρ=2:***

| | | Expected Sample Size, E(N) | | |
|---|---|---|---|---|
| **Number of analyses, $K$** | **Inflation factor, $R$** | **When $\theta = 0$ as % of $n_{fix}$** | **When $\theta = \delta/2$ as % of $n_{fix}$** | **When $\theta = \delta$ as % of $n_{fix}$** |
| 2 | 1.06 | 70.7 | 89.2 | 80.8 |
| 3 | 1.09 | 64.4 | 84.8 | 75.3 |
| 4 | 1.12 | 61.1 | 82.4 | 72.4 |
| 5 | 1.13 | 59.2 | 80.9 | 70.6 |

Cytel

# Choosing $\rho$ & $K$

Properties of $\rho$-family designs with non-binding futility boundaries, type I error rate $\alpha = 0.025$, and power 0.9 when $\theta = \delta$.

***For ρ=2:***

Cholesterol study example

| | | Expected Sample Size, E(N) | | |
|---|---|---|---|---|
| **Number of analyses, $K$** | **Inflation factor, $R$** | **When $\theta = 0$ as % of $n_{fix}$** | **When $\theta = \delta/2$ as % of $n_{fix}$** | **When $\theta = \delta$ as % of $n_{fix}$** |
| 2 | 1.06 | 70.7 | 89.2 | 80.8 |
| 3 | 1.09 | 64.4 | 84.8 | 75.3 |
| 4 | 1.12 | 61.1 | 82.4 | 72.4 |
| 5 | 1.13 | 59.2 | 80.9 | 70.6 |

Cytel

# Choosing $\rho$ & $K$

Properties of $\rho$-family designs with non-binding futility boundaries, type I error rate $\alpha = 0.025$, and power 0.9 when $\theta = \delta$.

***For ρ=3:***

| | | Expected Sample Size, E(N) | | |
|---|---|---|---|---|
| **Number of analyses, $K$** | **Inflation factor, $R$** | **When $\theta = 0$ as % of $n_{fix}$** | **When $\theta = \delta/2$ as % of $n_{fix}$** | **When $\theta = \delta$ as % of $n_{fix}$** |
| 2 | 1.02 | 74.8 | 92.0 | 84.1 |
| 3 | 1.04 | 68.8 | 87.2 | 78.2 |
| 4 | 1.06 | 65.4 | 84.7 | 75.1 |
| 5 | 1.07 | 63.3 | 83.2 | 73.3 |

Cytel

# Discussion: Group Sequential Tests

Designs from the $\rho$-family of error spending tests are highly efficient:
Values of $E(N)$ are close to the minimum possible for a given number of analyses $K$ and inflation factor $R$.*

A lower value of $\rho$ gives greater reductions in expected sample size and time to a conclusion, but at the cost of a higher maximum sample size.

Tables of operating characteristics, as in the previous slides, can aid the choice of a suitable design.

*Barber & Jennison, *Biometrika*, 2002.

Cytel


# ***Sample Size Re-Estimation***

## *Benefits of Advanced Methods*

# Sample Size Re-estimation (SSR)

**Another way to seek the same benefits of**

- Reduced sample size
- An earlier conclusion

In a group sequential test, we set a large sample size and hope to stop early.

Some prefer the philosophy of "Start small, then ask for more".

In a two-stage design with Sample Size Re-estimation, we:

Set an initial sample size

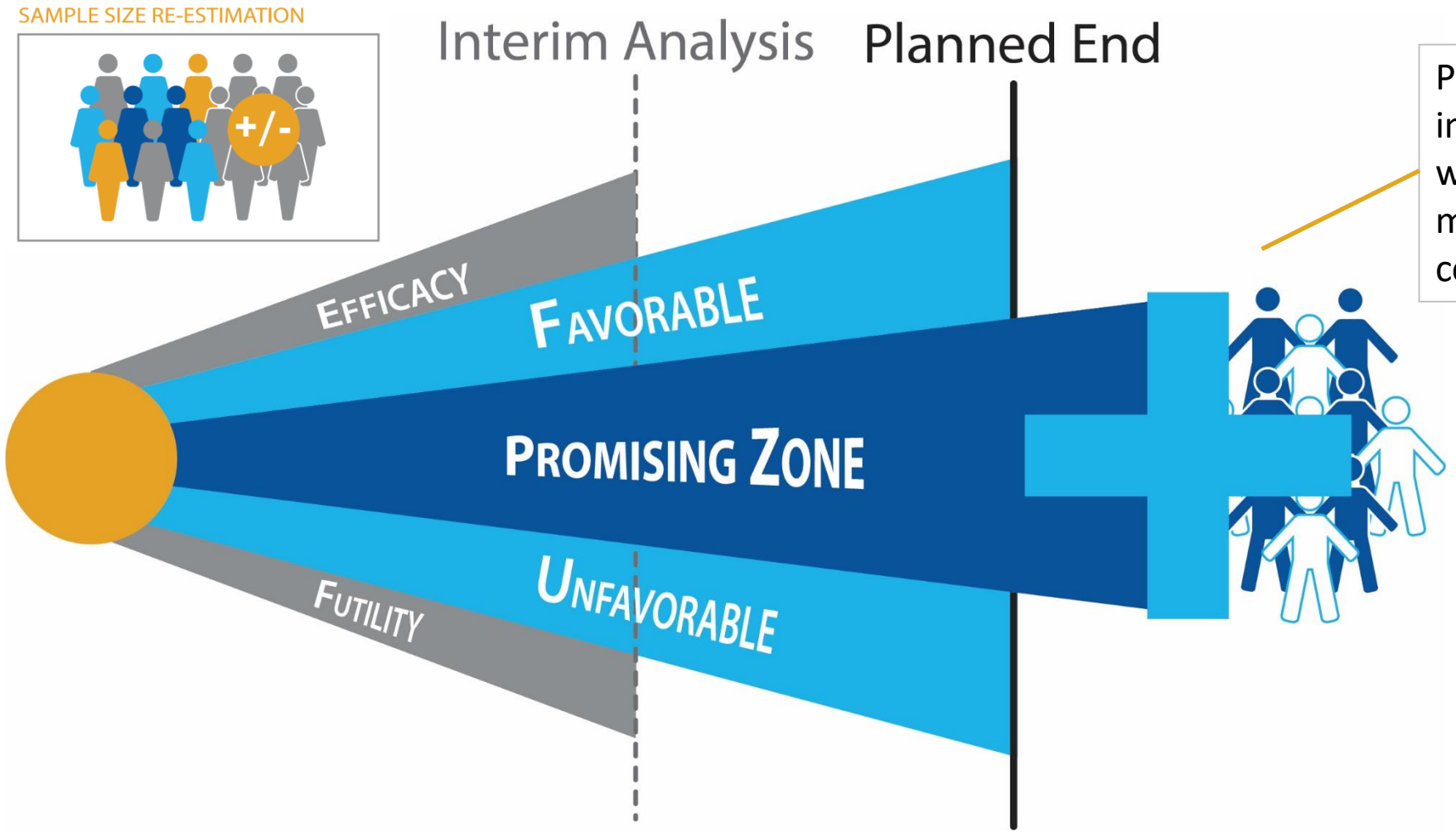
Conduct an interim analysis

Possibly increase the sample size

Analyse the final set of data

Cytel

# Adaptive Sample Size Re-estimation Can Increase Probability of Success




Pre-planned criteria for increasing enrollment when expected trial end may not each a meaningful conclusion



Cytel


# *Sample Size Re-Estimation*

## *Method Theory*

# SSR: Controlling The Type I Error Probability

## A two-stage combination test (Bauer and Köhne, *Biometrics*, 1994)

In Stage 1: Calculate $Z_{(1)}$ based on Stage 1 data

In Stage 2: Calculate $Z_{(2)}$ based on new data from Stage 2

Define

$$Z = \frac{1}{\sqrt{2}} Z_{(1)} + \frac{1}{\sqrt{2}} Z_{(2)}$$

For a level $\alpha = 0.025$ test, reject the null hypothesis if $Z > 1.96$.

Note: Z is a measure of evidence against the null hypothesis, which states the new treatment is no better than the control.

Type I error is protected even if the Stage 2 sample size depends on Stage 1 data.

Cytel

# SSR: Controlling The Type I Error Probability

**Multi-stage tests** (Cui, Hung & Wang, *Biometrics*, 1999; Lehmacher & Wassmer, *Biometrics*, 1999)

Define a $K$-stage group sequential test.

Express the cumulative $Z$-statistic at analysis $k$ as

$$Z_k = \frac{1}{\sqrt{k}} Z_{(1)} + \cdots + \frac{1}{\sqrt{k}} Z_{(k)} \qquad (*)$$

where $Z_{(1)}, \ldots, Z_{(k)}$ are the $Z$-statistics based on *new* data in each Stage $1, \ldots, k$.

If "adaptation" has occurred at an analysis $j < k$, replace $Z_{(j+1)}, \ldots, Z_{(k)}$ in (*) by the new $\tilde{Z}_{(j+1)}, \ldots, \tilde{Z}_{(k)}$ and apply the original group sequential testing boundary.

Cytel

# Poll Question

**Do you use Sample Size Re-estimation designs?**

A: No

B: Yes, but only rarely

C: Yes



Cytel


# *Sample Size Re-estimation*

## *Case Studies*

# SSR: An Example

## *Cholesterol Study:*

Has $\alpha = 0.025$, power $0.9$ when $\theta = 10$, $K = 3$ analyses, error spending test with $\rho = 2$, non-binding futility boundary.

Three groups are planned with 48 patients per treatment in each.

Suppose we observe

Analysis 1: $\hat{\theta}_1 = 4.2, \;\; Z_1 = 0.832$

Analysis 2: $\hat{\theta}_2 = 5.0, \;\; Z_2 = 1.386$ (based on cumulative data)

Cytel

# SSR: An Example

## *Cholesterol study*

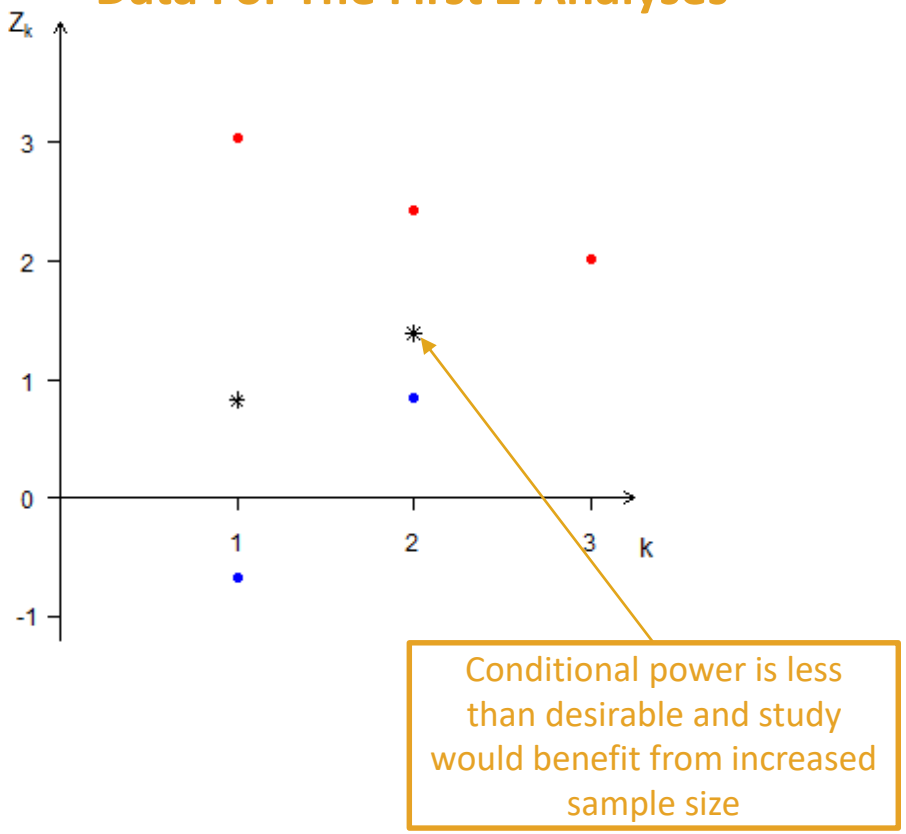
We Observe

**Analysis 1:** $\hat{\theta}_1 = 4.2,\ Z_1 = 0.823$

**Analysis 2:** $\hat{\theta}_2 = 5.0,\ Z_2 = 1.386$

(based on cumulative data)

**Data For The First 2 Analyses**




Can we increase the final sample size?

Cytel

# Demonstration of East software

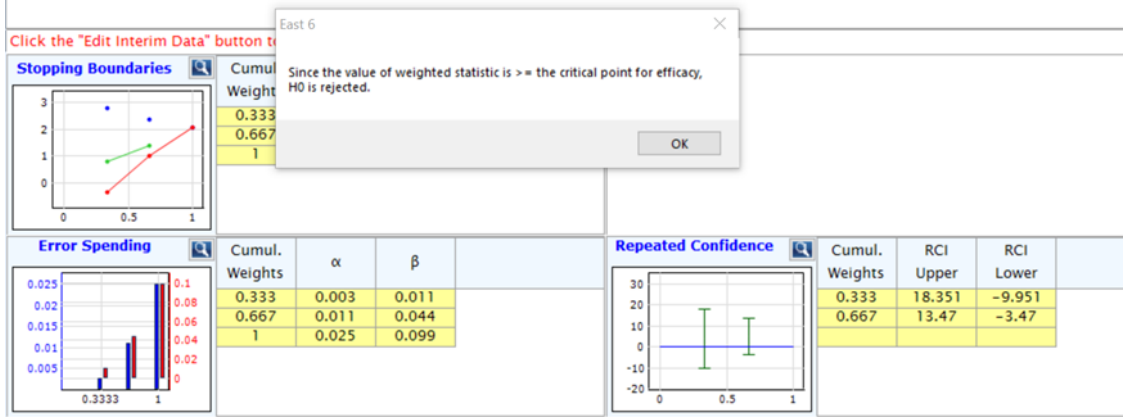
***Cholesterol Study:***

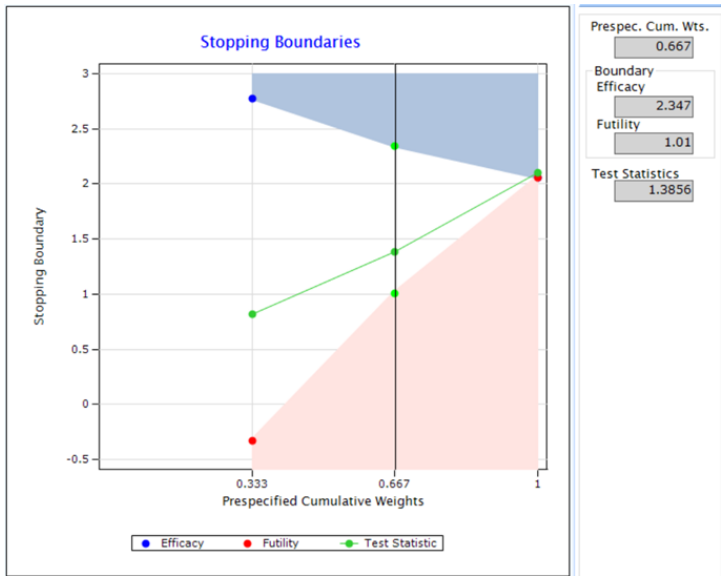
- Error Spending Group Sequential Design
- Conditional Power
- A “CHW” Design

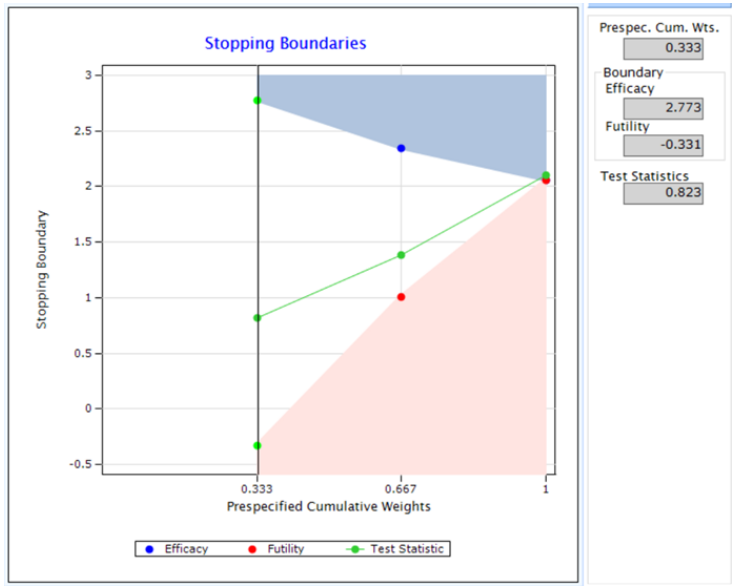
Cytel

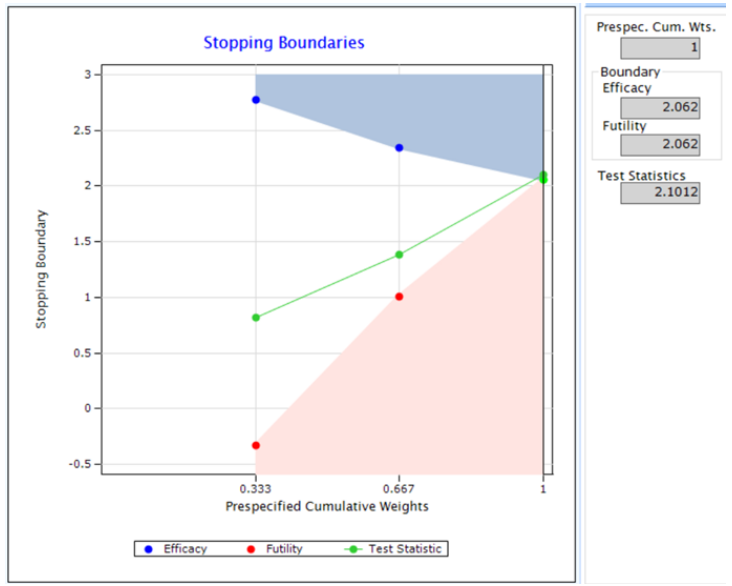
# Takeaways

| Look # | Incremental Sample Size | Cumulative Sample Size | Incremental Statistic | Prespecified Weights | Weighted Statistic | δ | Standard Error | Efficacy | Futility | 95% RCI for δ Upper | 95% RCI for δ Lower | Repeated p-value |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 96 | 96 | 0.823 | 0.333 | 0.823 | 4.2 | 5.103 | 2.773 | -0.331 | 18.351 | -9.951 | 1 |
| 2 | 96 | 192 | 1.137 | 0.333 | 1.386 | 5.8 | 5.103 | 2.347 | 1.01 | 13.47 | -3.47 | 0.204 |
| 3 | 144 | 336 | 1.68 | 0.333 | 2.101 | 7 | 4.167 | 2.062 | 2.062 | 11.411 | 0.108 | 0.023 |




# *Sample Size Re-estimation*

## *Discussion*

# SSR: Discussion

**The role of sample size re-estimation**

- Rescuing an under-powered trial
- In a prospectively designed trial (FDA require *pre-specified* adaptive designs)

**Assessing the performance of SSR designs**

- The role of the sample size rule
- Compare a group sequential design with the same maximum sample size

**Staging of investment**

- Is this different in an SSR design?

Cytel

# SSR: Handling "Pipeline" Data

**In a group sequential trial, data may still arrive after deciding to "stop"**

- The primary endpoint has not yet been observed for recently treated patients
- Observations are recorded after data are "locked" for the interim analysis.

---

Hampson and Jennison (*J. Royal Statist. Soc., B*, 2013) proposed **"Delayed response group sequential designs"** to deal with this issue.

In a 2-stage design with a large number of pipeline patients, a design based on sample size re-estimation can be an attractive proposition.

Cytel

# SSR: Handling "Pipeline" Data

**Example: Schizophrenia Study**

(Mehta and Pocock, *Statistics in Medicine*, 2010)

Endpoint: Improvement in NSA at 26 weeks

Initial sample size = 442

**At The Interim Analysis**

416 patients enrolled

208 observed responses,

208 "pipeline" patients

**Sample Size Decision**

Continue to the original target of 442 patients or increase the final sample size

Cytel

# SSR: Mehta and Pocock's "Promising zone" design

See Mehta & Pocock (2010) for more on their **"Promising zone" design**.

---

Jennison & Turnbull (*Statistics in Medicine*, 2015) discuss this design and **propose a cost-benefit approach, in which gains in conditional power are set against increased patient numbers.**

Hsiao, Liu & Mehta (*Biometrical Journal*, 2018) propose **a sample size rule that combines Jennison & Turnbull's approach with a requirement of a minimum conditional power for sample size to increase.**

Cytel

Cytel
Shaping the Future
of Drug Development

# *Factors to Consider & Summary*

# Factors To Consider

The stopping boundary and sample size rule must be pre-specified

The Data Monitoring Committee should discuss the rationale for study design with the sponsor before the trial is under way and a “firewall” is in place.

We need to increase maximum sample size slightly to account for interim looks

Cytel

# Summary

**Group sequential designs help reduce patient numbers and reach early conclusions.**

Designs with 2 or 3 analyses and a maximum sample size 5% or 10% greater than the fixed sample test can make savings of around 30%.

Sample size re-estimation offers an alternative approach, and can be an attractive option when there is a large amount of "pipeline" data.

Error spending tests are efficient and flexible. Current software makes these methods straightforward to apply.

Cytel

East

# Easy Access to the Adaptive Designs That Matter

Delivered by the Thought Leaders Behind the Methods

Software that is Faster & Easier to Use

Popular Fixed and Adaptive Designs at your Fingertips

Cytel

# Global Products and Services

## Statistical Software

Industry standard for trial design, including CID adaptive (East, EOD)

Leader in exact statistical solutions (Xact: StatXact, LogXact, Procs)

Operations software (e.g. ACES, EnForeSys, FlexRandomizer)

All 25 top biopharma companies, the FDA, EMA & PMDA use our software

## Strategic Consulting

PhD statisticians expert in innovative design & complex statistical questions

Experts in Data Science, PK/PD, Enrolment & Event Forecasting, Portfolio/Program Optimization (NPV)

## Project-Based Services

Reliable Biometrics service provider delivering high quality, on time

Lead staff with over 15 years industry experience on average

Including biostatistics & programming, ISC, data management, PK/PD analysis, medical writing

## Functional Services Provision (FSP)

Creation of dedicated teams operating within/as an extension of the client's own biostatistics & programming, data management and PK/PD teams

Leader in offshoring of Biometrics competencies

Cytel

Cytel
Shaping the Future
of Drug Development

# Q&A


# Conclusion

# Upcoming Webinars

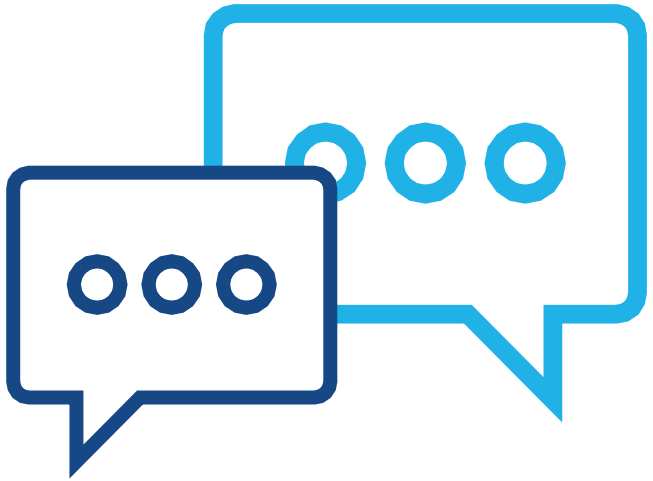
| Topic | Date | Time | Speaker | |
|---|---|---|---|---|
| Complex Innovative Trial Designs at a Glance – The Concepts, the Promise, and the Factors to Consider | Wednesday, May 20, 2020 | 11:00AM EDT \| 16:00 GMT | Zoran Antonijevic | ✓ |
| Group Sequential Designs and Sample Size Re-estimation – Modern Uses | Wednesday, June 3, 2020 | 11:00AM EDT \| 16:00 GMT | Christopher Jennison | |
| Practical Model-based Approaches for Phase I Oncology Trials | Wednesday, June 17, 2020 | 11:00AM EDT \| 16:00 GMT | Satrajit Roychoudhury | |
| Introduction to Population Enrichment | Wednesday, July 15, 2020 | 11:00AM EDT \| 16:00 GMT | Thomas Burnett | |

**Other Topics Planned for Series:** Introduction to Adaptive Dose Finding, Seamless Phase 2/3 Trial Designs, Basket Trial Designs, Umbrella Trial Design, Multi-arm Multi-stage Trial Design, and Program/Portfolio Designs

**Recordings will be posted to www.cytel.com.**



Cytel



# Thank you


## Professor Chris Jennison

*University of Bath, UK*




Cytel

Cytel
Shaping the Future
of Drug Development

# Thank you

www.cytel.com