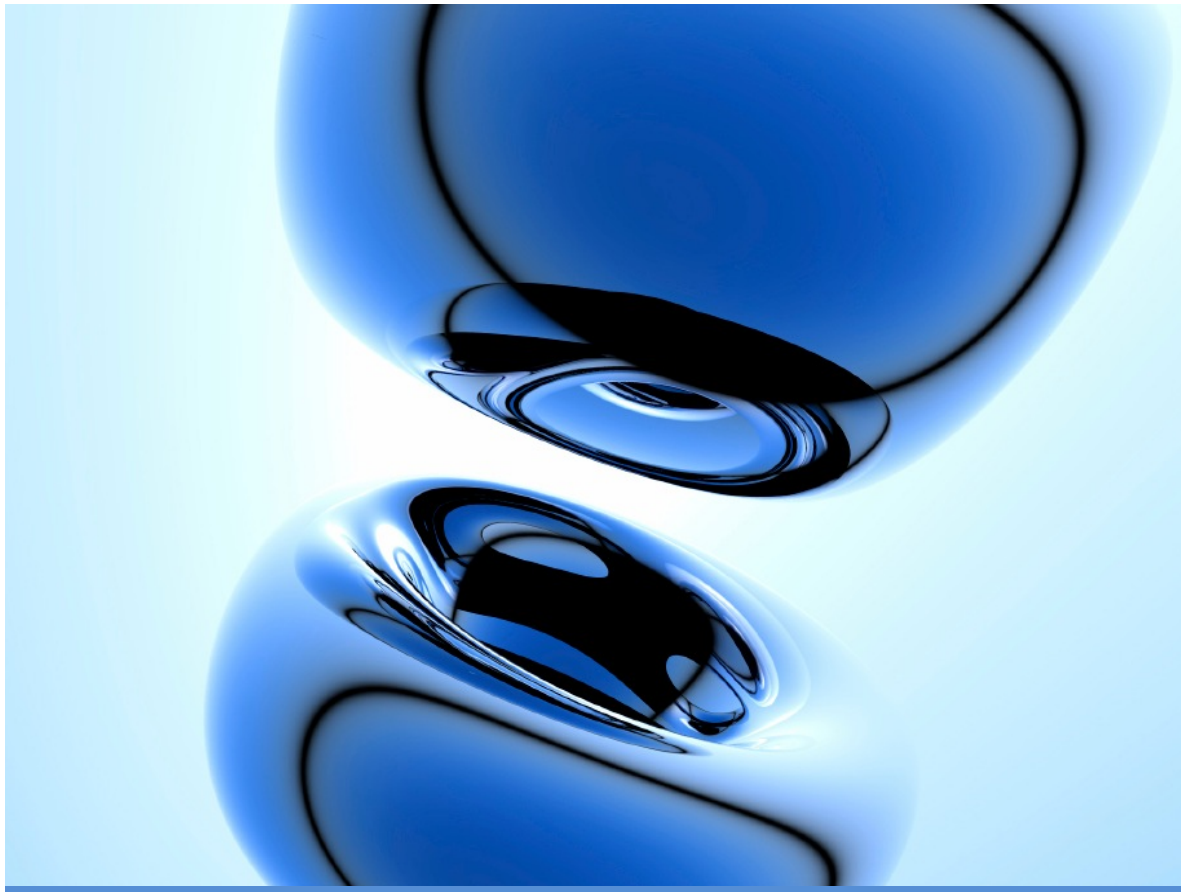


Cytel

Sample Size Re-Estimation: Controlling the Type-1 Error

Yannis Jemiai, Ph.D.

26 September 2017

ASA Biopharmaceutical Section

Regulatory-Industry Statistics Workshop

unblinded sample size re-estimation is an essential design tool

Addresses uncertainty in trial design assumptions

One of the most popular adaptations, especially when using a Promising Zone approach

21st Century Cures Act, PDUFA VI, encourage the use of adaptive designs

Regulatory guidance documents exist from EMA (2007), FDA CDER / CBER (2010), and CDRH (2016)

Increasingly many examples of regulatory acceptance

Cytel

So what are some of the issues concerning uSSR designs?

Can type-1 error be controlled?

Can sound adaptive decision rules be developed?

How do we get a point estimate and confidence intervals for the treatment effect?

How do we avoid operational bias during trial conduct?

We focus here on type-1 error control

Cytel

Why does type-1 error get inflated?

Consider a two-stage design without sample size increase

STAGE I	STAGE II
sample size n_1	sample size $n^{(2)}$
estimate $\hat{\delta}_1$	estimate $\hat{\delta}^{(2)}$
compute $z_1 = \hat{\delta}_1 / \text{se}(\hat{\delta}_1)$	compute $z^{(2)} = \hat{\delta}^{(2)} / \text{se}(\hat{\delta}^{(2)})$

Reject H_0 if $\sqrt{\frac{n_1}{n_1+n^{(2)}}} z_1 + \sqrt{\frac{n^{(2)}}{n_1+n^{(2)}}} z^{(2)} \geq C_\alpha$

Suppose now that we increase the sample size in stage II from $n^{(2)}$ to $n^{*(2)}$, but we do not change the critical value

$$\sqrt{\frac{n_1}{n_1 + n^{*(2)}}} z_1 + \sqrt{\frac{n^{*(2)}}{n_1 + n^{*(2)}}} z^{*(2)} \geq C_\alpha$$

This will lead to type-1 error inflation

Cytel

How can we control type-1 error then?

1. Use CHW statistic with pre-specified weighting of data from each stage (Cui, Hung & Wang, 1999)
2. Use conventional Wald test if promising interim result are obtained (Chen, DeMets, Lan, 2004; Gao, Ware, Mehta, 2008) (Only valid for two-stage designs)
3. Preserve the conditional type-1 error that would have been obtained had there been no adaptation (Muller & Schafer, 2001)

Cytel

Chronology of development (partial list)

Year	Journal	Authors	Contribution
1994	Biometrics	Bauer & Köhne	Combining p-values from two stages
1995	Biometrics	Proschan & Hunsberger	Conditional error rate function
1998	Statist. Med.	LD Fisher	Variance spending
1999	Biometrics	Cui, Hung & Wang	Weighted combination of Z-statistics
1999	Biometrics	Lehmacher & Wassmer	Weighted combination of p-values
2001	Biometrics	Müller & Schäfer	Conditional rejection probability principle
2003	Biometrika	Tsiatis & Mehta	Indiscriminate SSR is inefficient
2004	Statist. Med.	Chen, DeMets, Lan	Sample size increase only if CP $>$ 50%
2007	Statist. Med.	Mehta et. al.	RCIs for adaptive GSDs
2011	Statist. Med	Mehta & Pocock	Promising zone designs
2013	Statist. Med	Gao, Liu, Mehta	Exact estimation for adaptive GSDs

Cytel

1. Use a weighted statistic with pre-specified weights

1. Hypothesis Testing Without Sample Size Increase

STAGE I	STAGE II
sample size n_1	sample size $n^{(2)}$
estimate $\hat{\delta}_1$	estimate $\hat{\delta}^{(2)}$
compute $z_1 = \hat{\delta}_1 / \text{se}(\hat{\delta}_1)$	compute $z^{(2)} = \hat{\delta}^{(2)} / \text{se}(\hat{\delta}^{(2)})$

Reject H_0 if $\sqrt{\frac{n_1}{n_1+n^{(2)}}} z_1 + \sqrt{\frac{n^{(2)}}{n_1+n^{(2)}}} z^{(2)} \geq C_\alpha$

2. Hypothesis Testing With Sample Size Increase

STAGE I	STAGE II WITH EXTENSION
sample size n_1	sample size $n^{*(2)} > n^{(2)}$
estimate $\hat{\delta}_1$	estimate $\hat{\delta}^{*(2)}$
compute $z_1 = \hat{\delta}_1 / \text{se}(\hat{\delta}_1)$	compute $z^{*(2)} = \hat{\delta}^{*(2)} / \text{se}(\hat{\delta}^{*(2)})$

Reject H_0 if $\sqrt{\frac{n_1}{n_1+n^{(2)}}} z_1 + \sqrt{\frac{n^{(2)}}{n_1+n^{(2)}}} z^{*(2)} \geq C_\alpha$

Cytel

Also called the p-value combination approach

- Although we have extended the sample size from $n^{(2)}$ to $n^{*(2)}$ at stage-II, the criterion for declaring statistical significance is

$$\sqrt{\frac{n_1}{n_1+n^{(2)}}}z_1+\sqrt{\frac{n^{(2)}}{n_1+n^{(2)}}}z^{*(2)}\geq C_\alpha$$

instead of

$$\sqrt{\frac{n_1}{n_1+n^{*(2)}}}z_1+\sqrt{\frac{n^{*(2)}}{n_1+n^{*(2)}}}z^{*(2)}\geq C_\alpha$$

- Contribution of the second cohort of patients has been down-weighted
- Also known as the method of "inverse normal weighting of p-values" because significance criterion can be expressed in the form

$$\sqrt{\frac{n_1}{n_1+n^{(2)}}}\Phi^{-1}(1-p_1)+\sqrt{\frac{n^{(2)}}{n_1+n^{(2)}}}\Phi^{-1}(1-p^{*(2)})\leq\Phi^{-1}(1-p_\alpha)$$

Cytel

2. Use the Conventional Wald Statistic

- Result is due to Chen, Demets and Lan (2004) (CDL method)
- Valid only for two-stage designs in which the sample size may be increased, but not decreased at the interim look
- Use conventional Wald statistic for the final analysis even if the sample size was increased from n_2 to n_{2*}, provided the interim results were promising
- Interim result is considered promising if $\text{CP}_{\hat{\delta}}(z_1, n_2) \geq 0.5$

Cytel

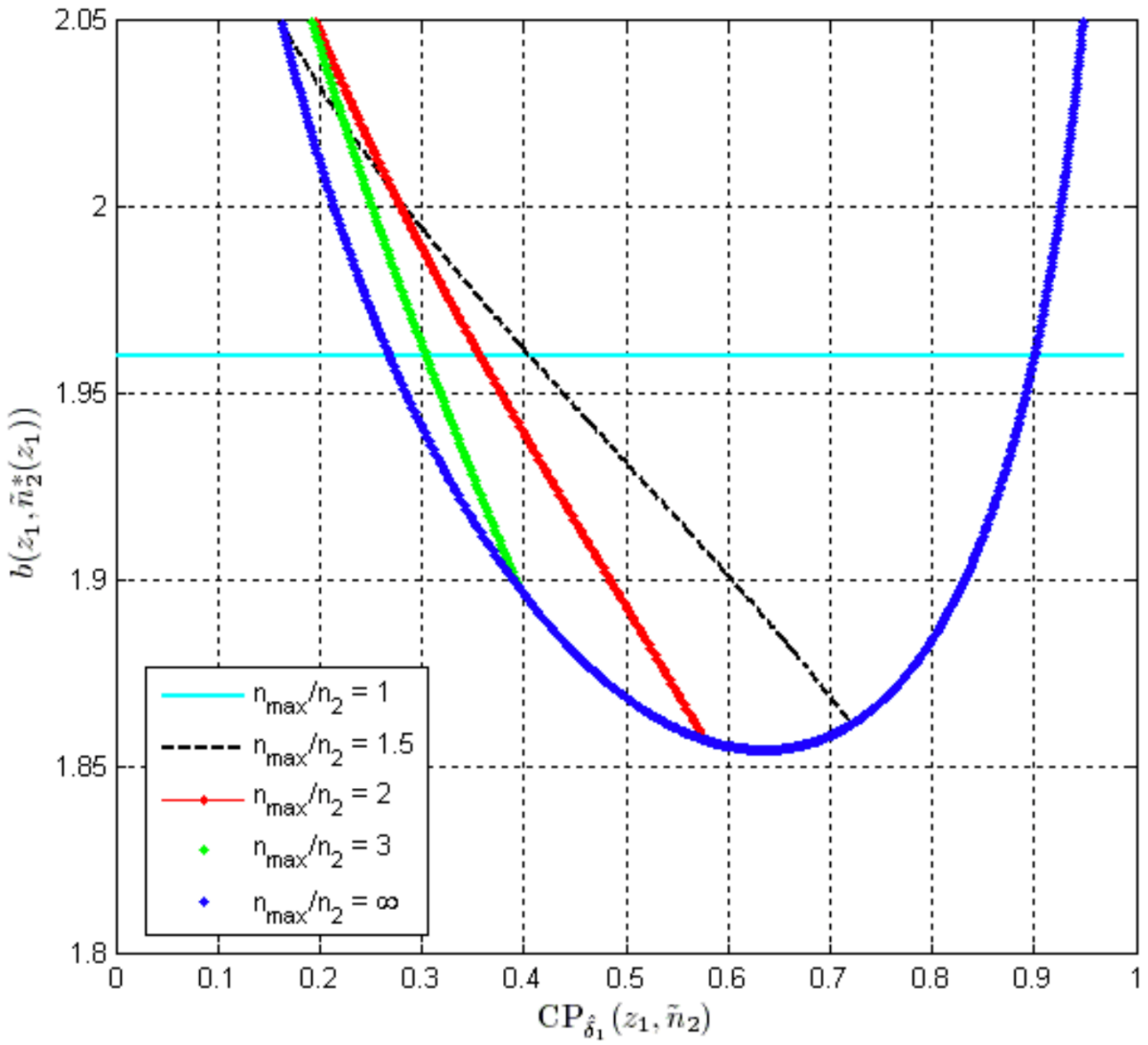
Extended CDL Method

- Due to Gao, Ware and Mehta (2008) and Mehta and Pocock (2010)
- Can relax the criterion for a using conventional Wald statistic if $CP_{\hat{\delta}_1}(z_1, n_2) \geq CP_{min}$ as tabulated below:

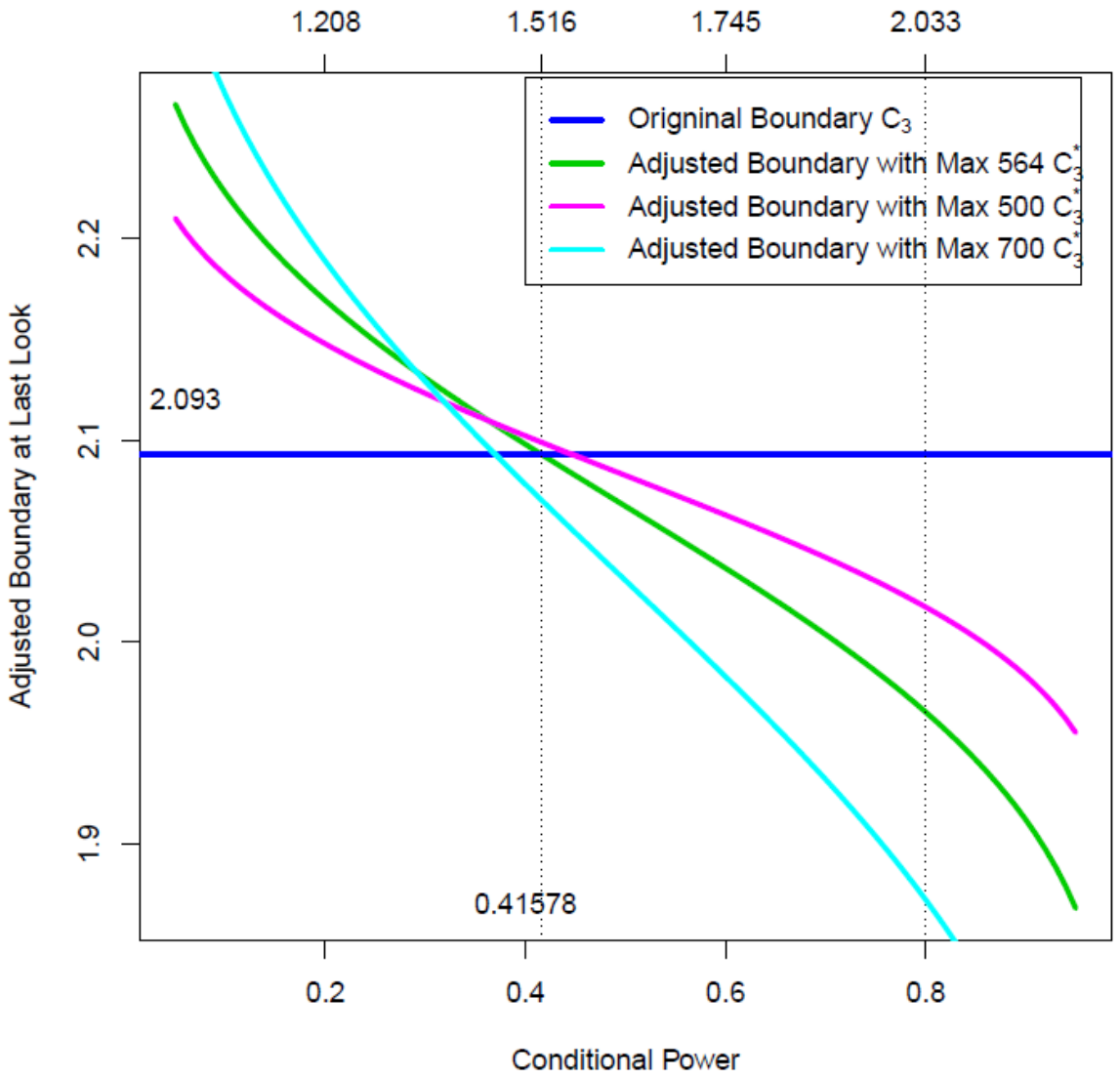
Sample Size Ratios		CP_{min} Values for Targeted Conditional Powers		
Maximum Allowed (N^*_{max}/n_2)	At Interim Look (n_1/n_2)	80%	90%	95%
1.5	0.25	0.42	0.42	0.42
1.5	0.5	0.41	0.41	0.41
1.5	0.75	0.38	0.38	0.38
2	0.25	0.37	0.37	0.37
2	0.5	0.36	0.36	0.36
2	0.75	0.33	0.33	0.33
3	0.25	0.32	0.32	0.32
3	0.5	0.31	0.31	0.30
3	0.75	0.30	0.27	0.27
∞	0.25	0.32	0.28	0.26
∞	0.5	0.31	0.27	0.25
∞	0.75	0.30	0.25	0.23

Cytel

Why does it work?



… and what are the concerns?



3. Preserve Conditional type-1 error rate

- Due to Muller and Schafer (2001)
- This method is the most flexible of all
- It gives full freedom to completely re-design a group sequential trial at any interim look. You could:
  - increase the sample size
  - change the spending function
  - alter the number and spacing of future interim looks
- **Only Requirement:** Preserve the conditional type-1 error computed at the time of the design modification

Cytel

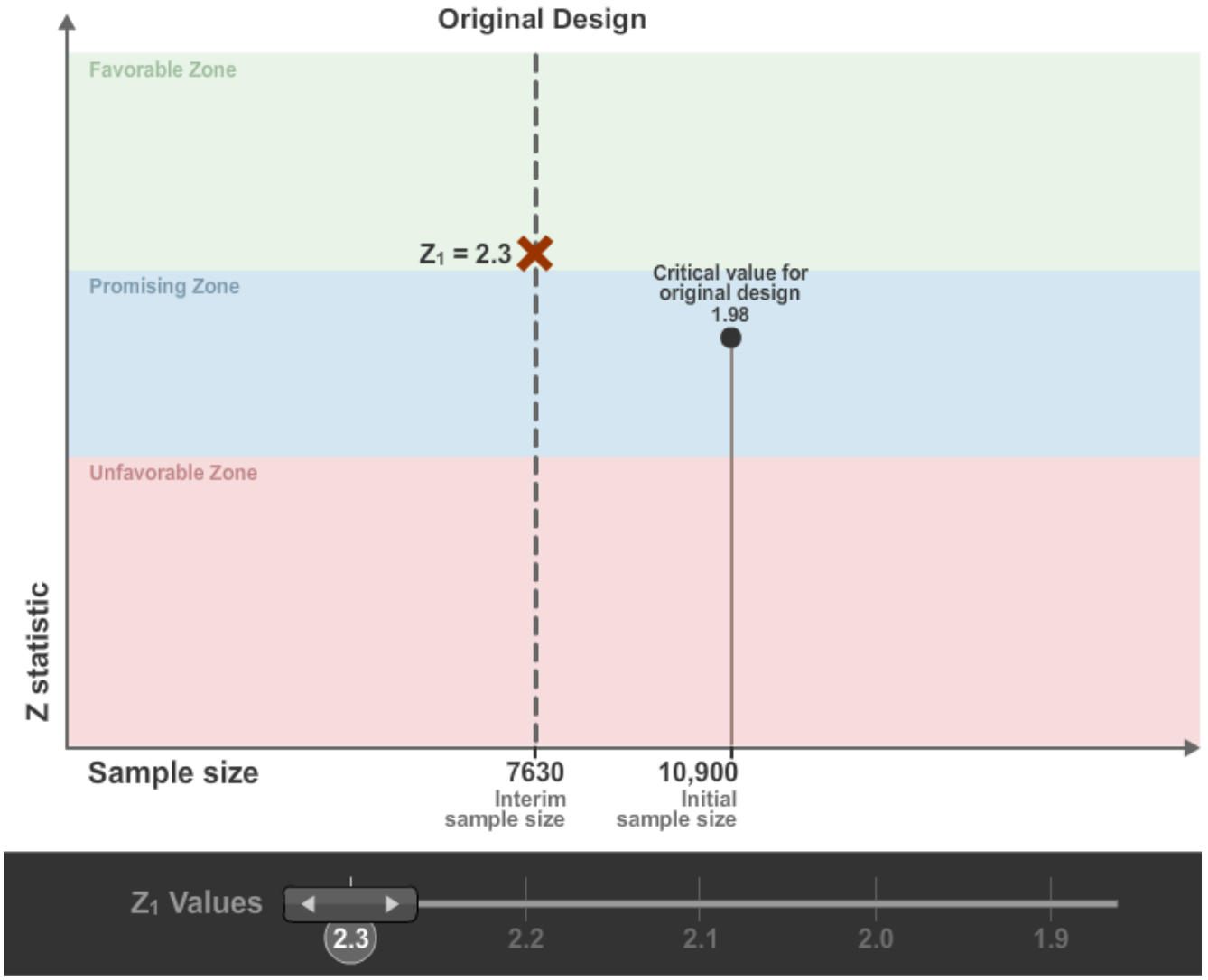
Preserving the overall type-1 error rate

In order to preserve the overall type-1 error of this procedure:

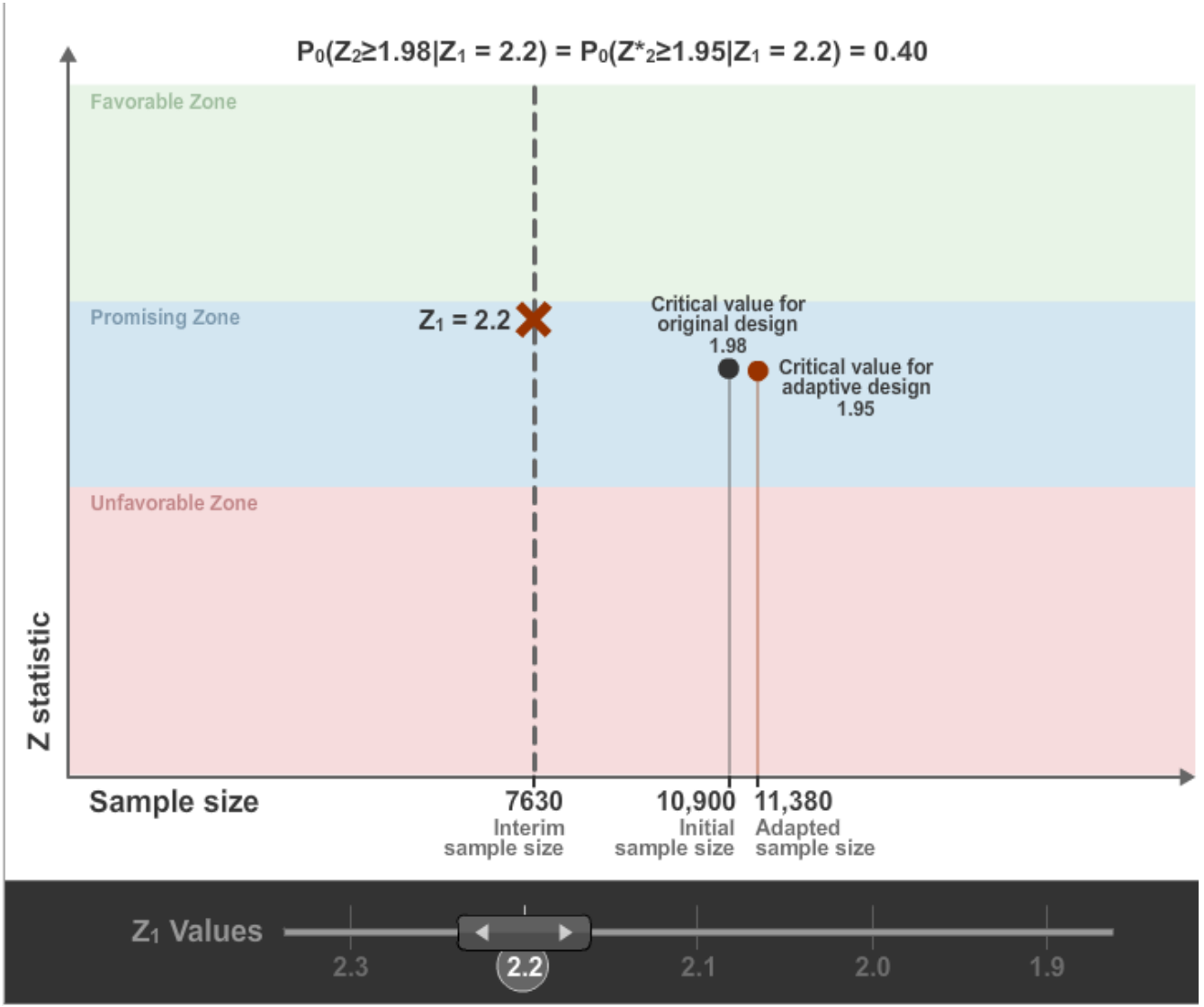
1. **Compute what the conditional type-1 error would be if you were to go to the end of the trial without re-designing**
2. **Use this conditional type-1 error as the significance level for the re-designed trial**

Cytel

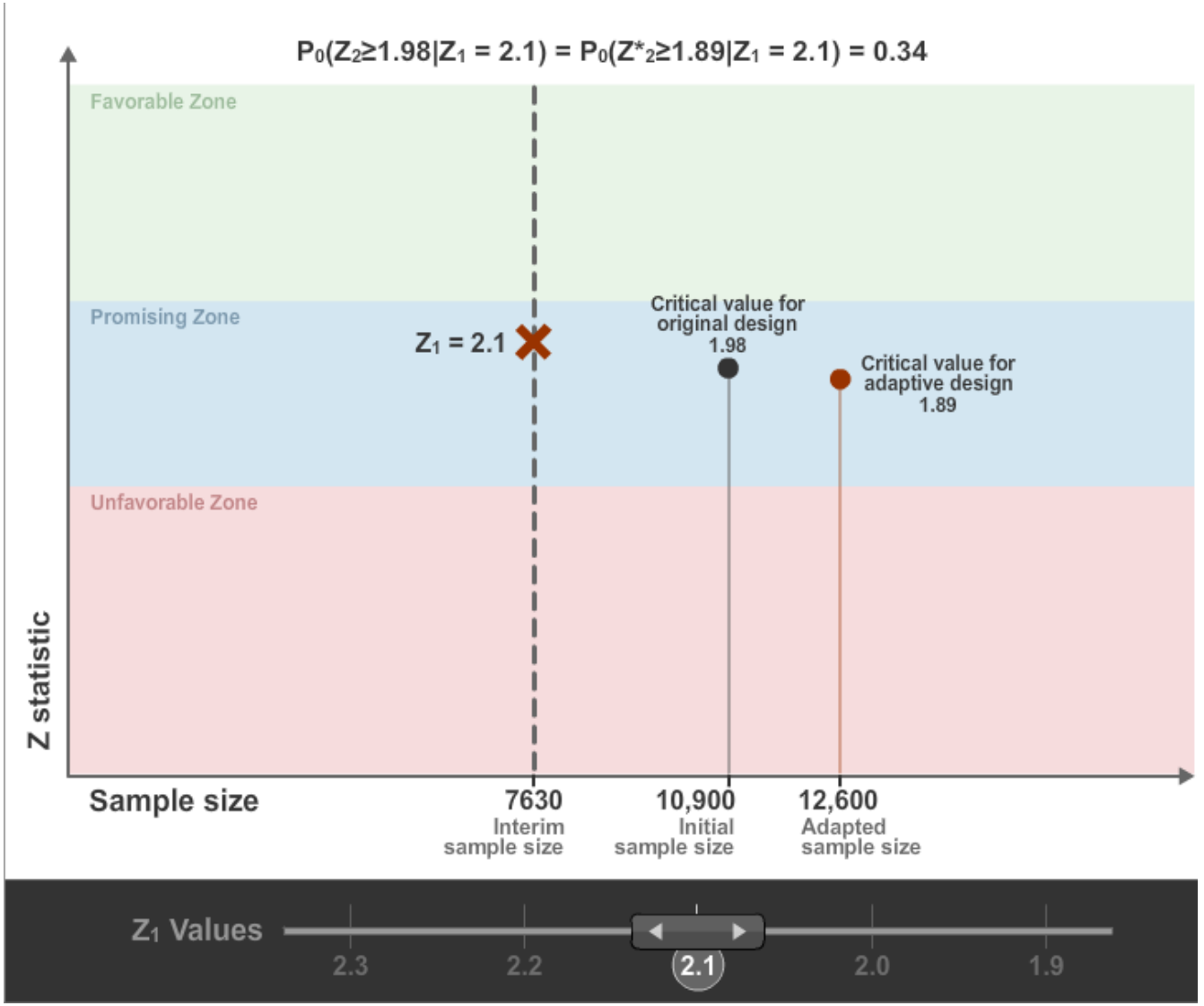
Conditional type-1 error rate



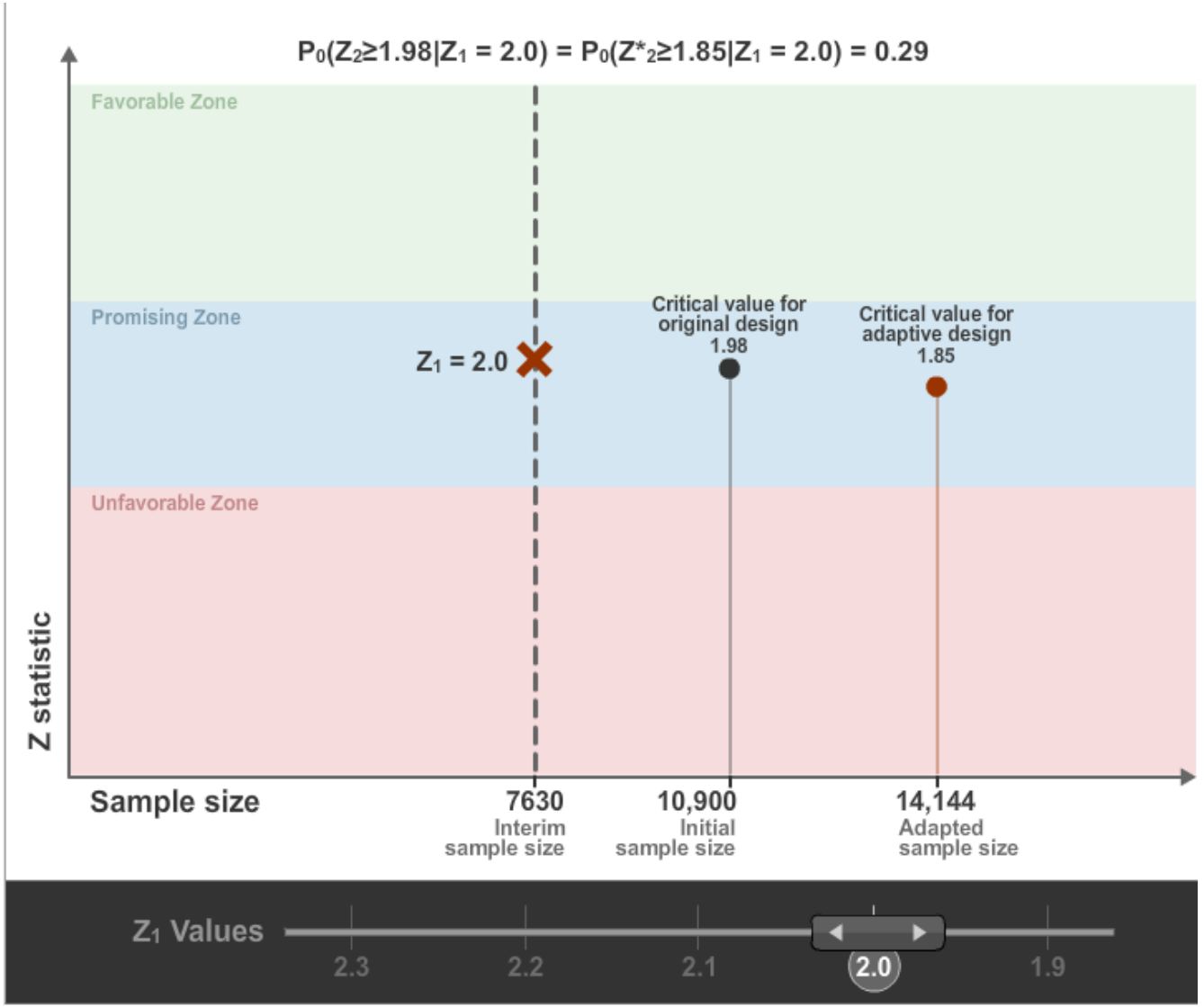
Conditional type-1 error rate



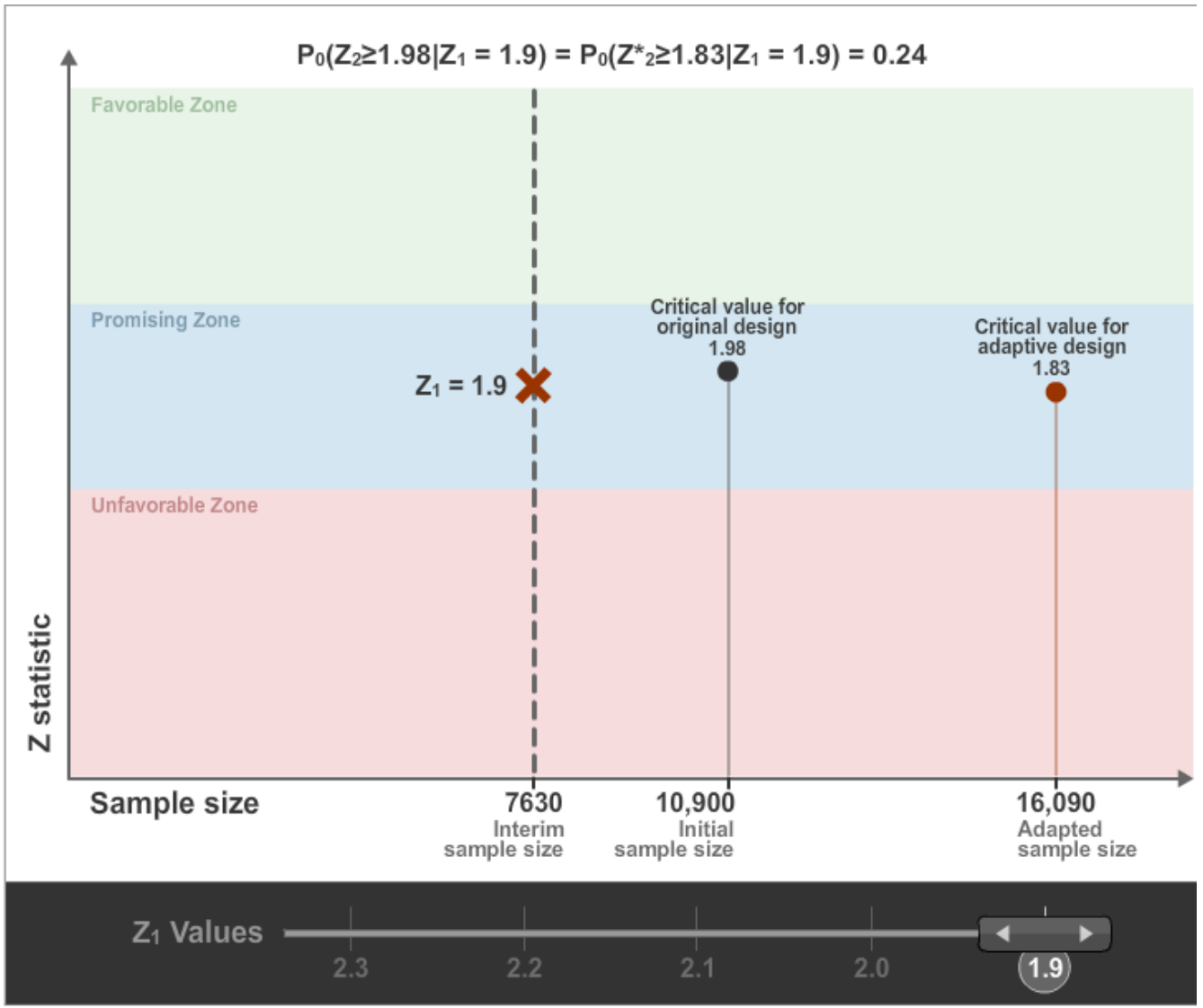
Conditional type-1 error rate



Conditional type-1 error rate



Conditional type-1 error rate



Points to consider

Handling survival endpoints

Usable information at interim analysis

Non-inferiority & equivalence settings

Independent increments

Small samples

Cytel

Recap: challenges in unblinded SSR trials

Type-1 error control is not an obstacle. Methods exist to ensure strong control

Inference remains a challenge, but making some progress

Decision-making algorithm can be optimized using simulations and latest research

Operational bias can be addressed/minimized by using iDMCs, putting in place proper processes, and making use of technology

Cytel

“By failing to prepare, you are preparing to fail.”

- Benjamin Franklin

Cytel

Thank you