Exact Inference for Adaptive Group Sequential Clinical Trials

Cyrus Mehta, Ph.D.

Cytel Inc.

Acknowledgements

- This is joint work with
  - Ping Gao, The Medicines Company
  - Lingyun Liu, Cytel Inc
- Expert computing help from Pranab Ghosh

Statistics in Medicine

Research Article

Received 26 February 2013, Accepted 16 April 2013 Published online 19 May 2013 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.5847

Exact inference for adaptive group sequential designs

Ping Gao,[a] Lingyun Liu[b] and Cyrus Mehta[b,c*†]

Background

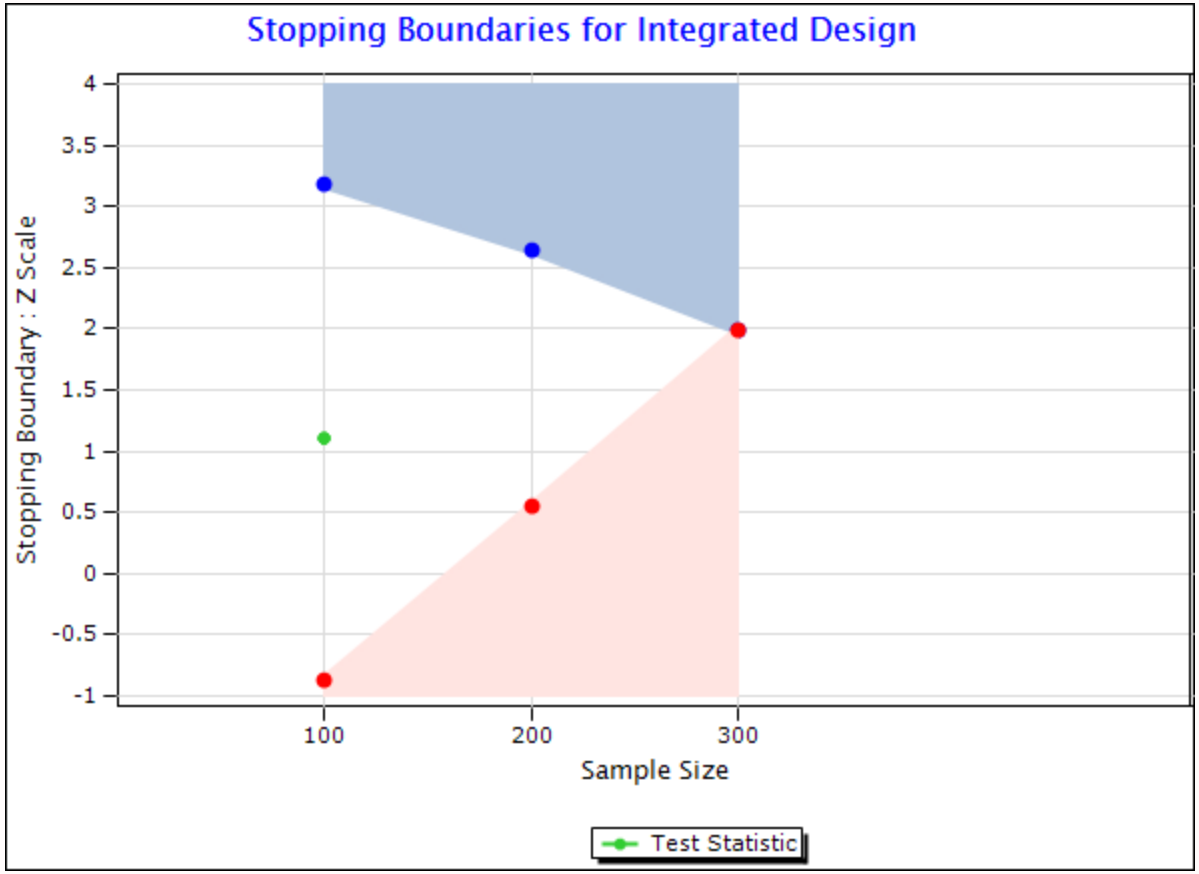
- Methods for type-1 error control in adaptive trials are well developed
- Methods for p-values, confidence intervals and parameter estimates not well developed
- Available Methods:
  - Extension of RCI (Jennison &Turnbull , 1989): do not exhaust the α, hence conservative
  - Extension of SWCI (Tsiatis, Rosner, Mehta, 1984): do exhaust the α, hence provide exact coverage

Adaptive Methods for Inference on δ

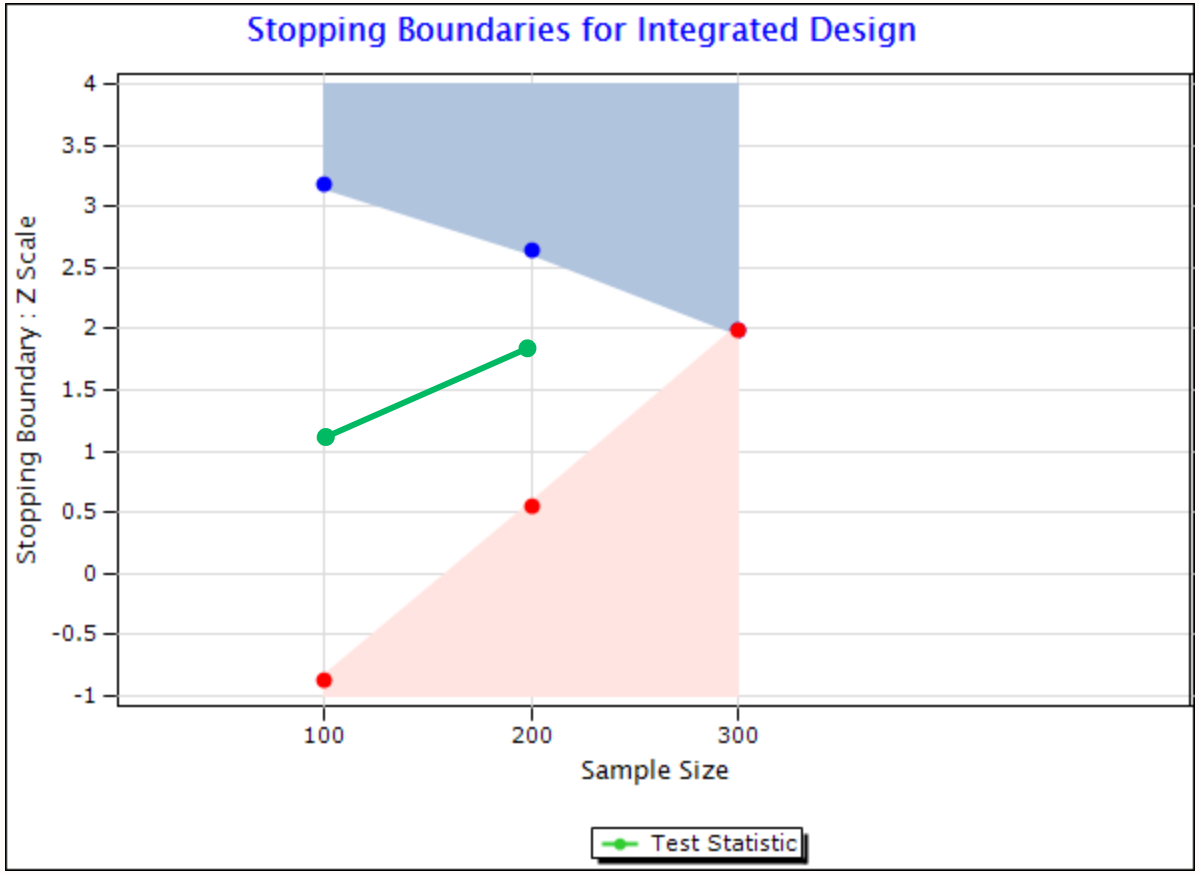
Cytel

1. RCI Method with weighted combination test
   - Lehmacher and Wassmer (Biometrics, 1999)
   - Conservative asymmetric coverage
   - Only applicable to sample size re-estimation
2. RCI Method with conditional error function
   - Mehta, Bauer, Posch, Brannath (Stat.Med, 2007)
   - Conservative asymmetric coverage
   - Applicable to any type of design change
3. SWCI Method with conditional error function
   - Brannath, Mehta, Posch (Biometrics, 2007)
   - Exact coverage
   - Applicable to any type of design change
   - Limited to one sided confidence intervals
4. BWCI Method with conditional error function
   - Gao, Liu, Mehta (Stat. Med, 2013)
   - Extends SWCI so as to handle two sided confidence intervals

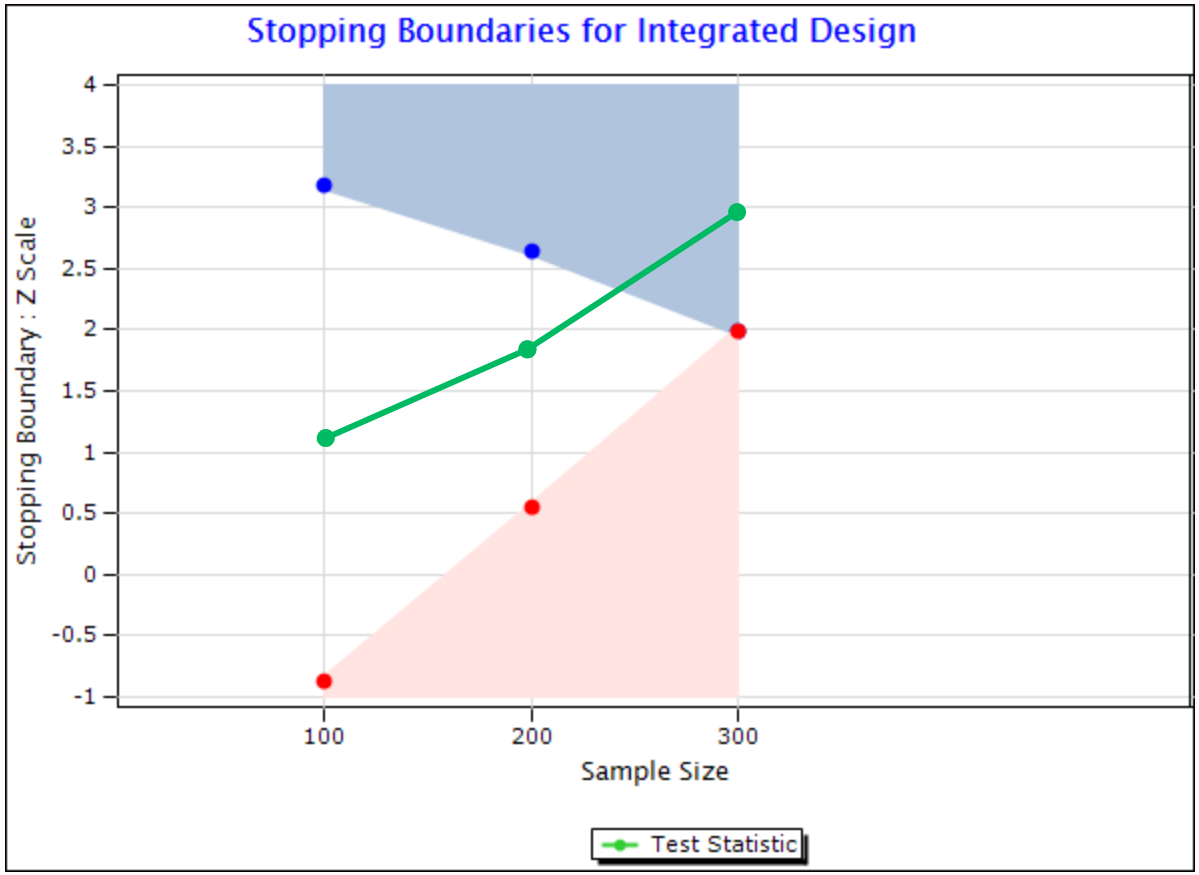
Look 1 of Classical Three-Look Design



Look 2 of Classical Three-Look Design



Look 3 of Classical Three-Look Design



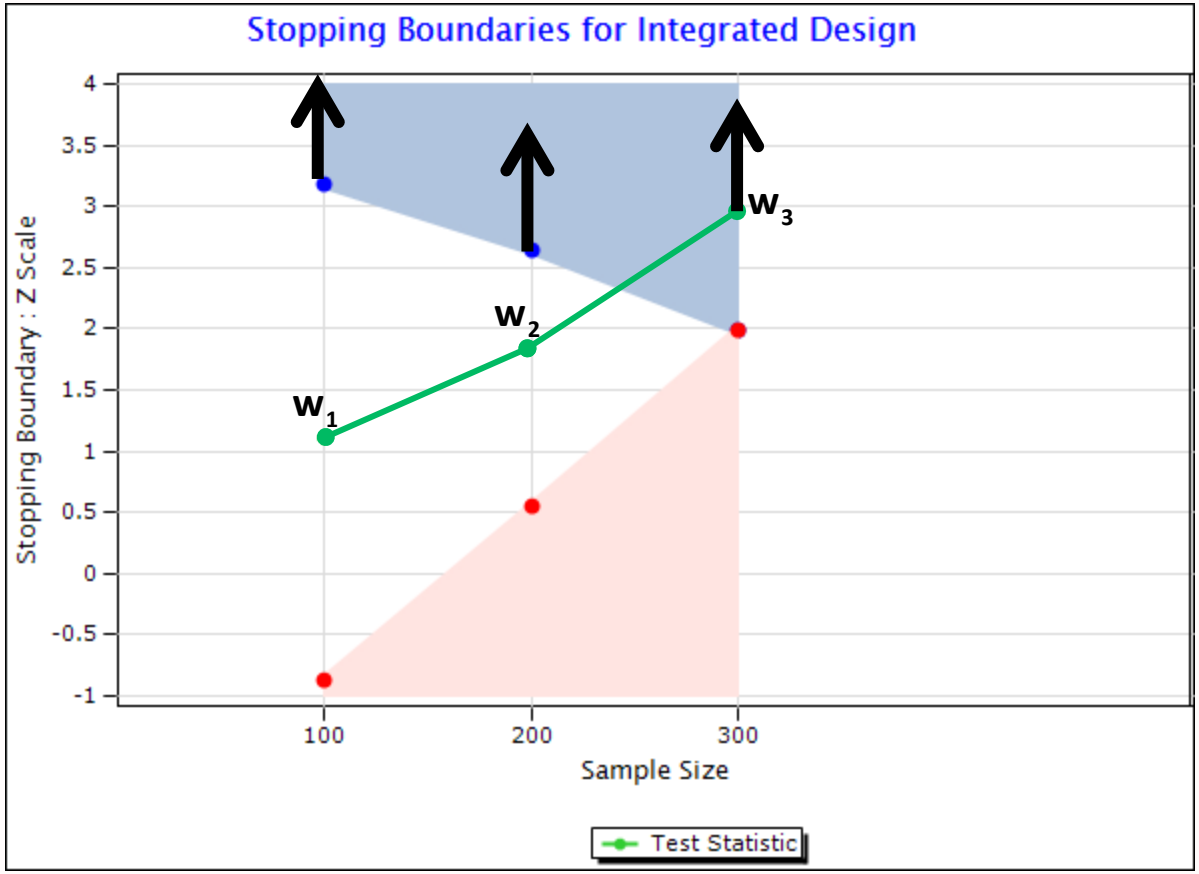
Null hypothesis that $\delta=0$ is rejected. But what about P-value and CI?

Classical P-value and Confidence Interval

- Stage wise ordering of sample space
  - Stopping at the same look with larger value of test statistic is more extreme
  - Stopping at an earlier look is more extreme than stopping at a later look
- Compute the probability of the more-extreme region based on this ordering

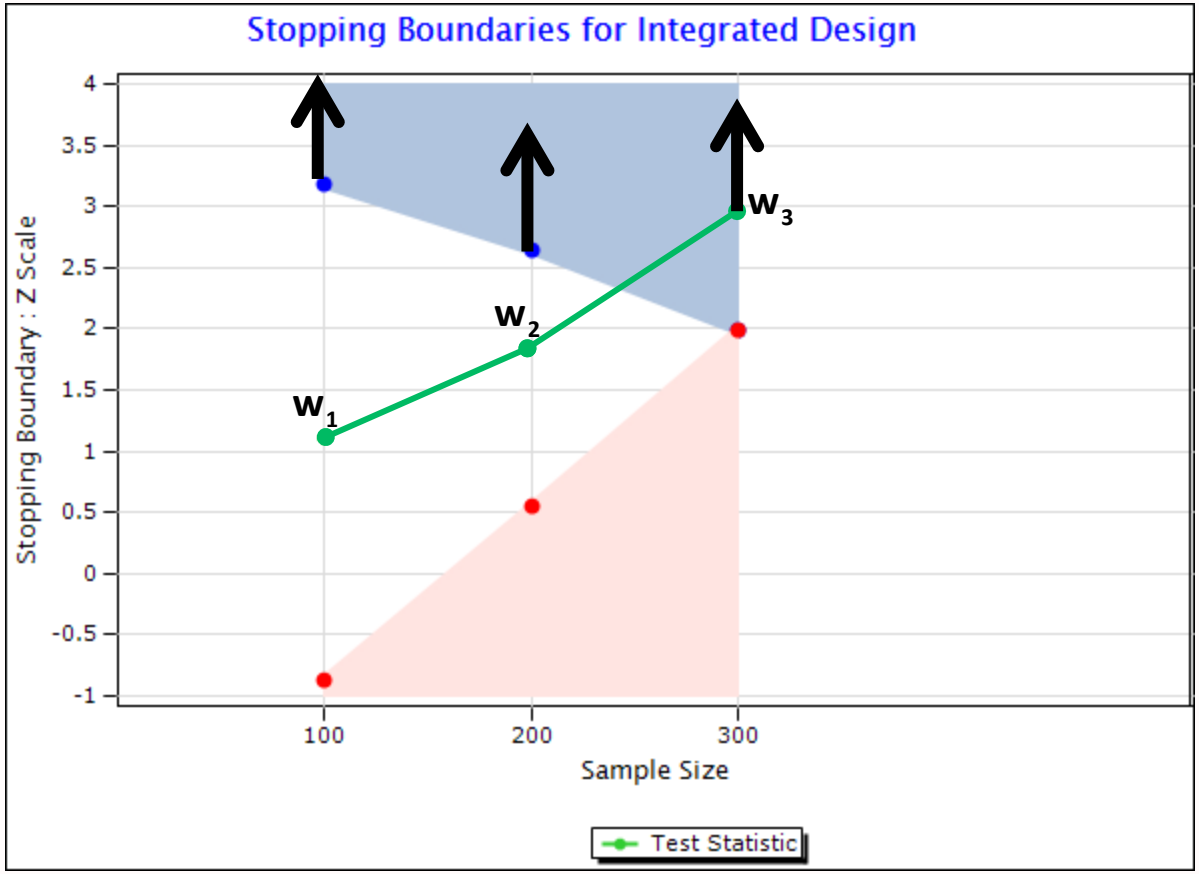
Classical SWCI Method: Tsiatis, Rosner, Mehta (*Biometrics*, 1984)

Classical SWCI: P-Value Computation



$$P_{value} = P_0(W_1 \geq b_1 \text{ or } W_2 \geq b_2 \text{ or } W_3 \geq w_3)$$

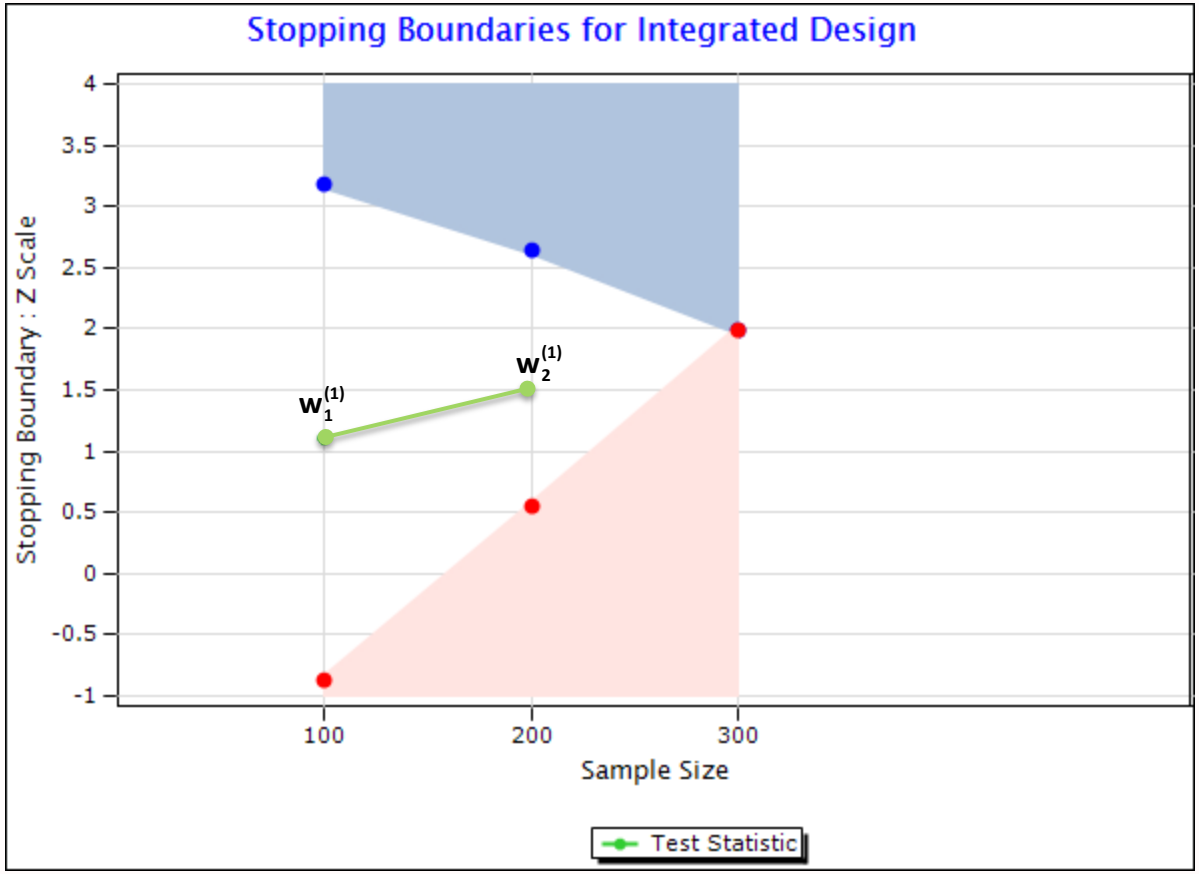
Classical SWCI: Confidence Interval



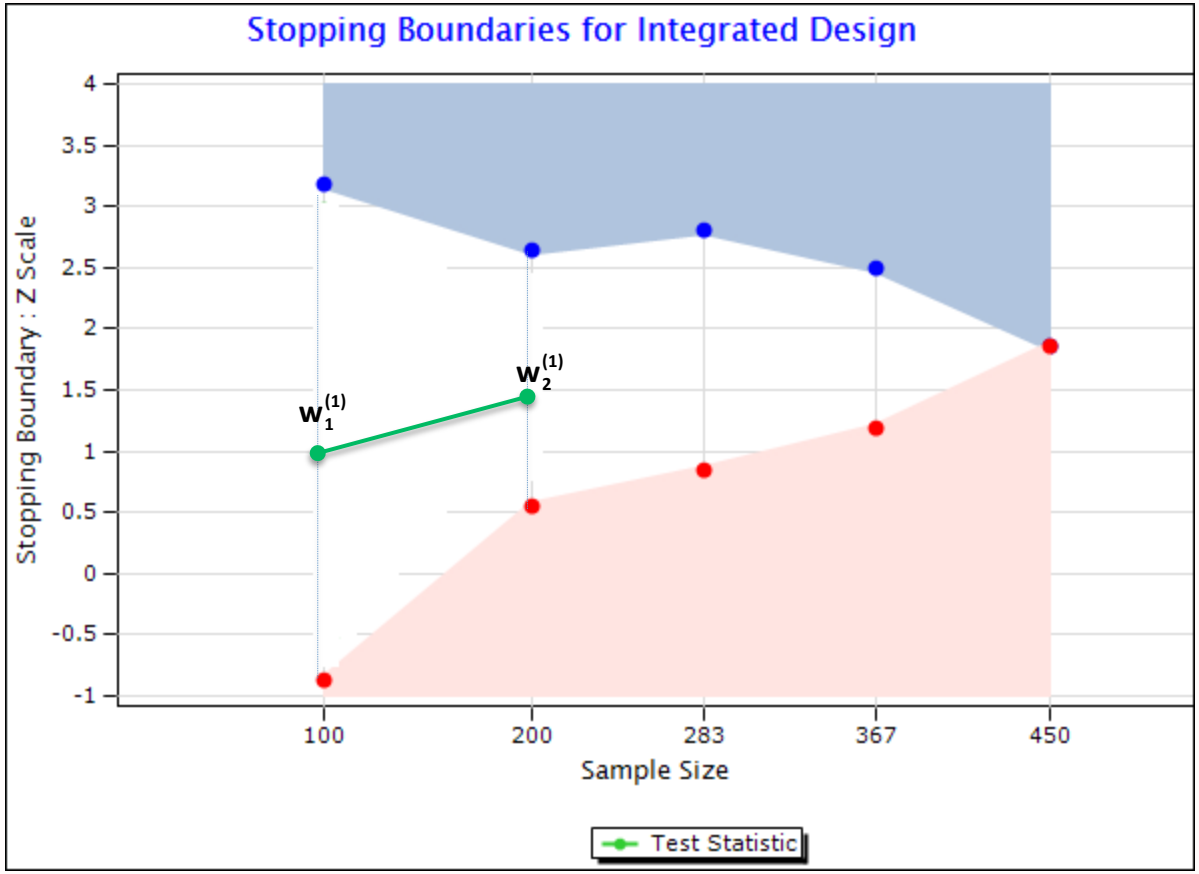
Find δ_{low} such that: $P_{\delta low}(W_1 \geq b_1 \text{ or } W_2 \geq b_2 \text{ or } W_3 \geq w_3) = 0.975$

Find δ_{up} such that: $P_{\delta up}(W_1 \geq b_1 \text{ or } W_2 \geq b_2 \text{ or } W_3 \geq w_3) = 0.025$

What if we adapt at look 2?



Adaptive Design at Look 2

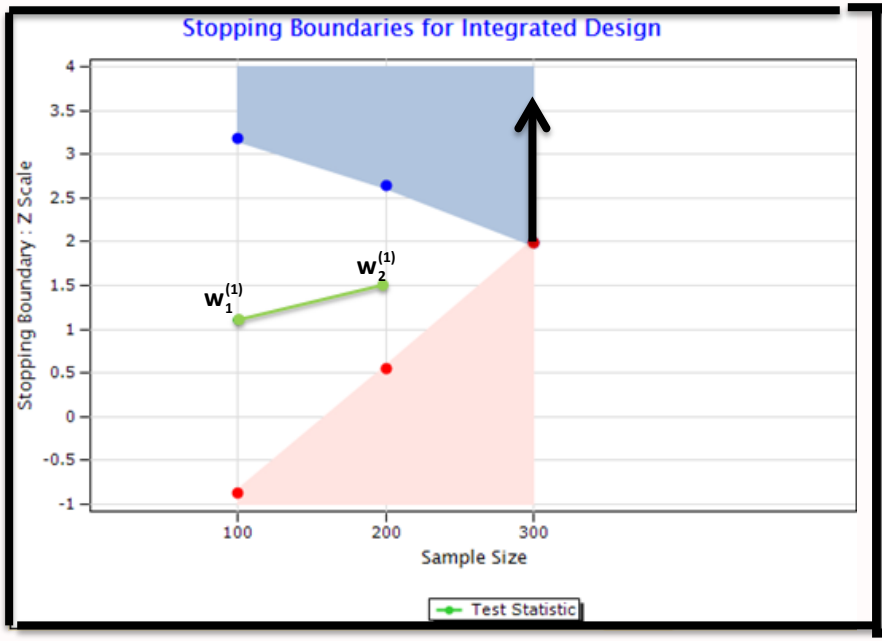


Created two additional looks, changed the spending function and increase sample size

How Type-1 Error is Preserved

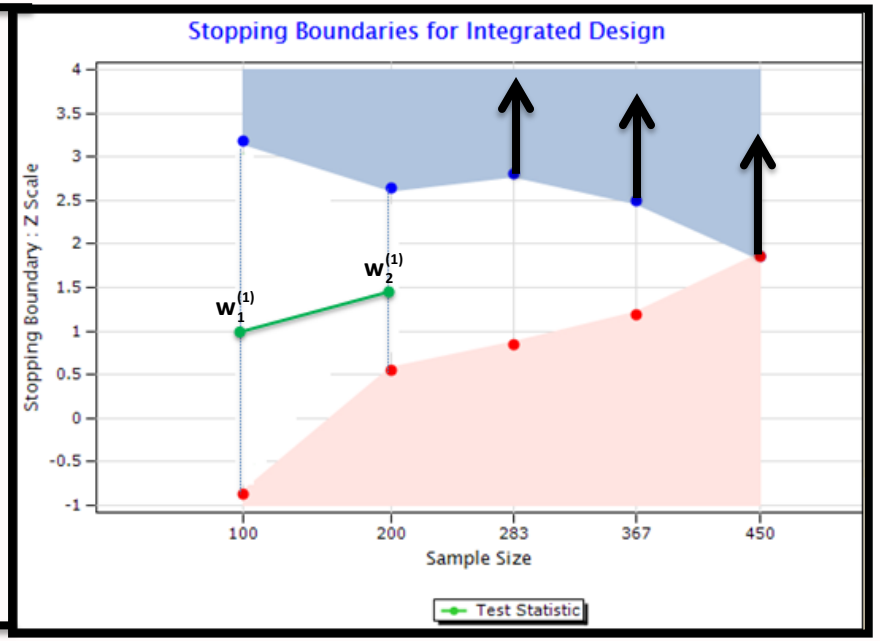
Design (1): Non-Adaptive

$$P_0(W_3^{(1)} \geq b_3^{(1)} \mid w_2^{(1)}) = \varepsilon_0^{(1)}$$



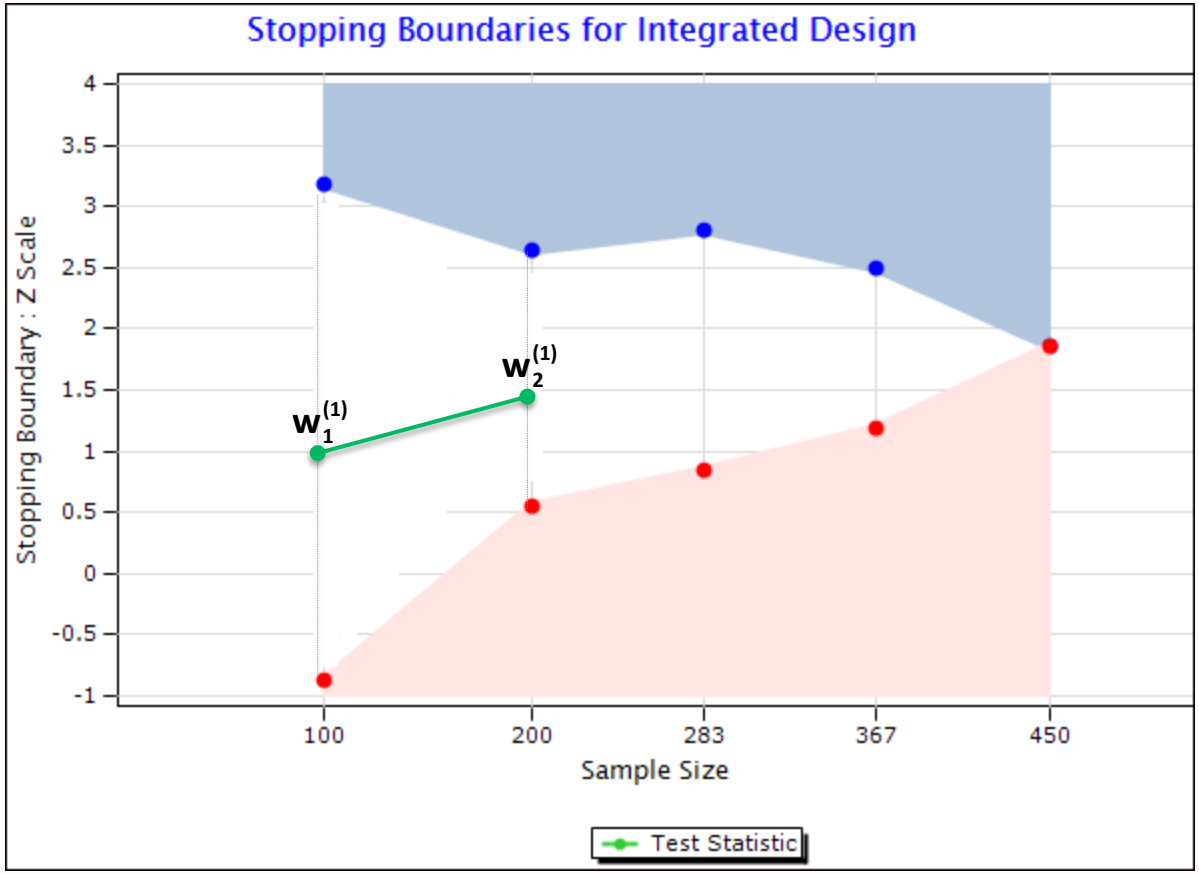
Design (2): Adaptive

$$P_0 \cup (W_3^{(2)} \geq b_3^{(2)}, W_4^{(2)} \geq b_4^{(2)}, W_5^{(2)} \geq b_5^{(2)} \mid w_2^{(1)}) = \varepsilon_0^{(2)}$$

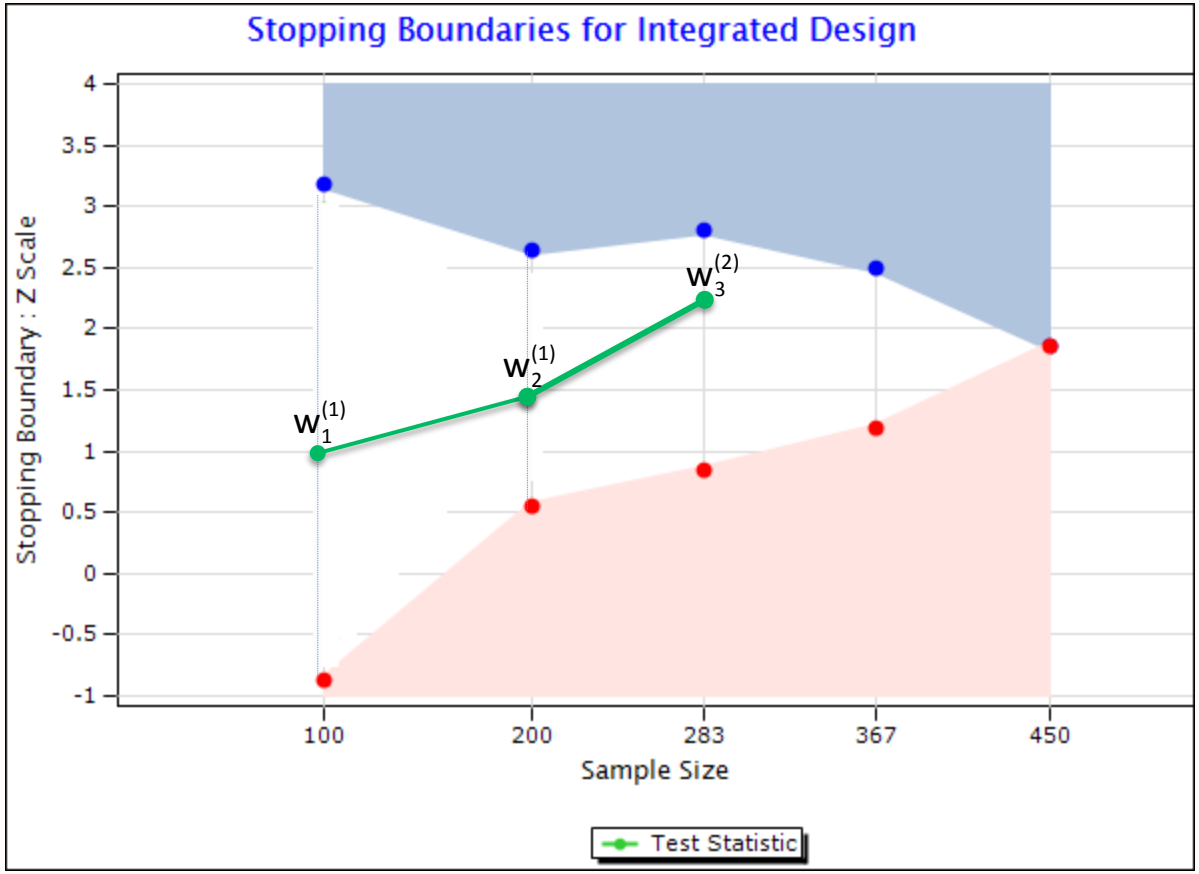


Identical Conditional Errors: $\varepsilon_0^{(1)} = \varepsilon_0^{(2)}$

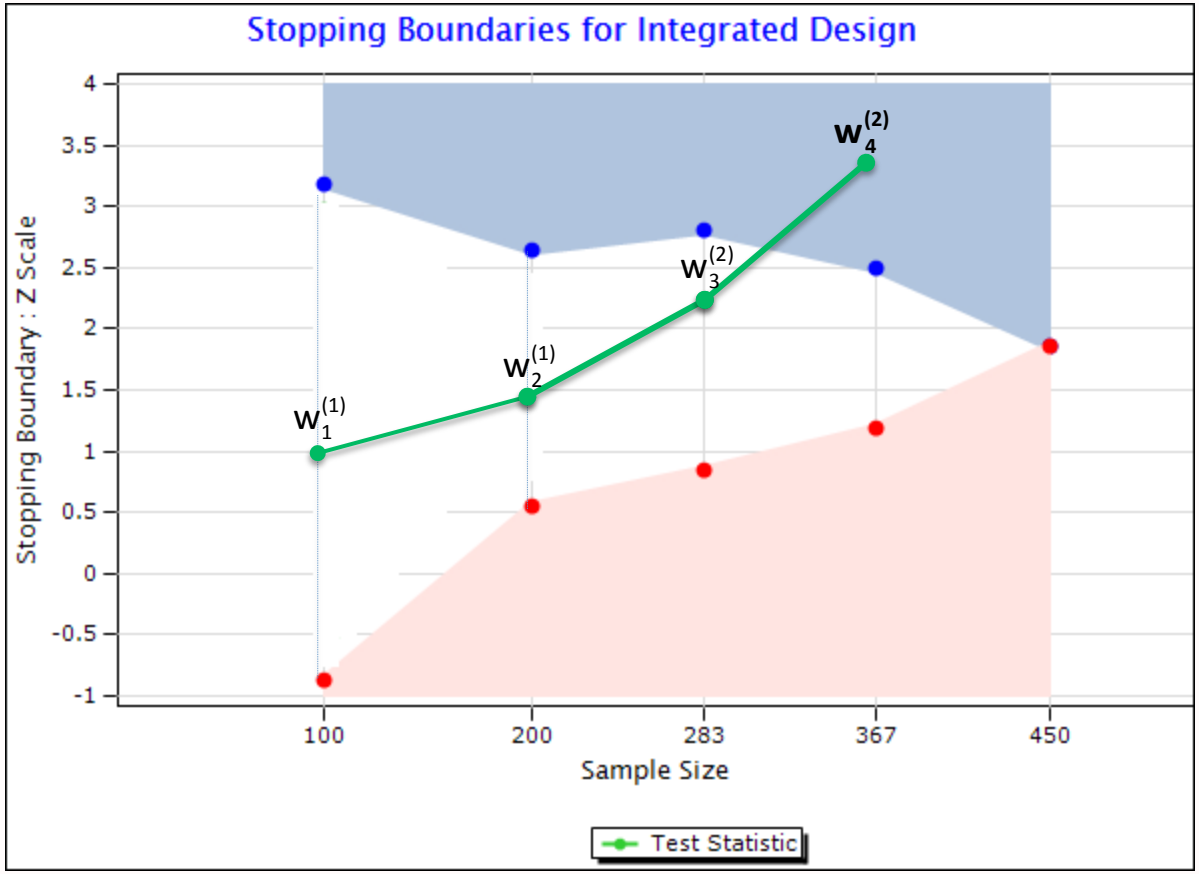
Monitoring the Adaptive Trial: Look 2



Monitoring the Adaptive Trial: Look 3

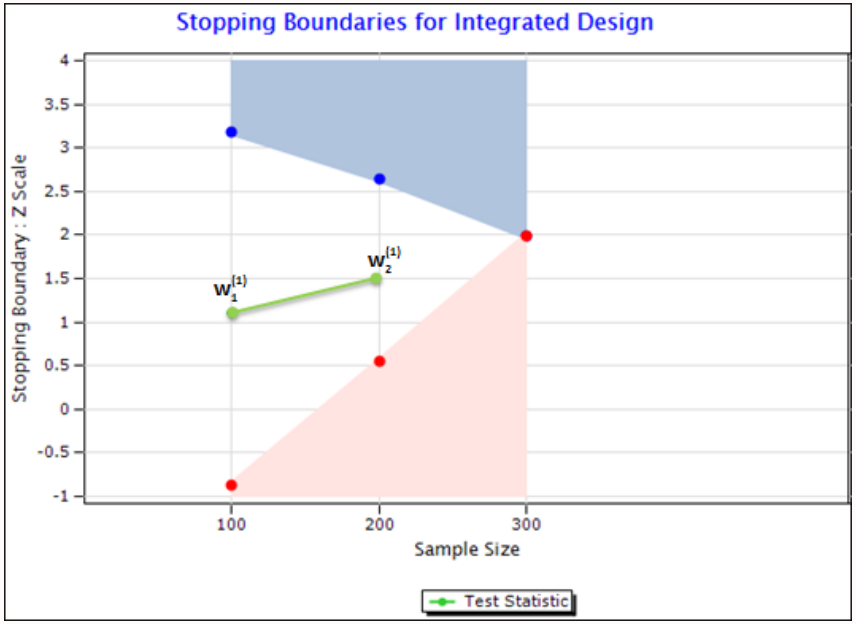


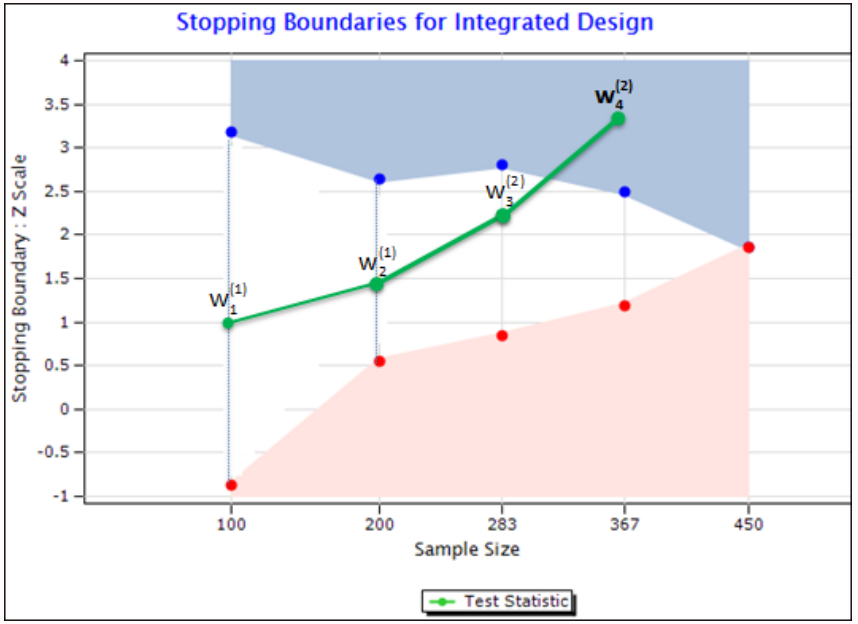
Monitoring the Adaptive Trial: Look 4



- **Null hypothesis that δ=0 is rejected**
- **What about P-value and confidence interval?**

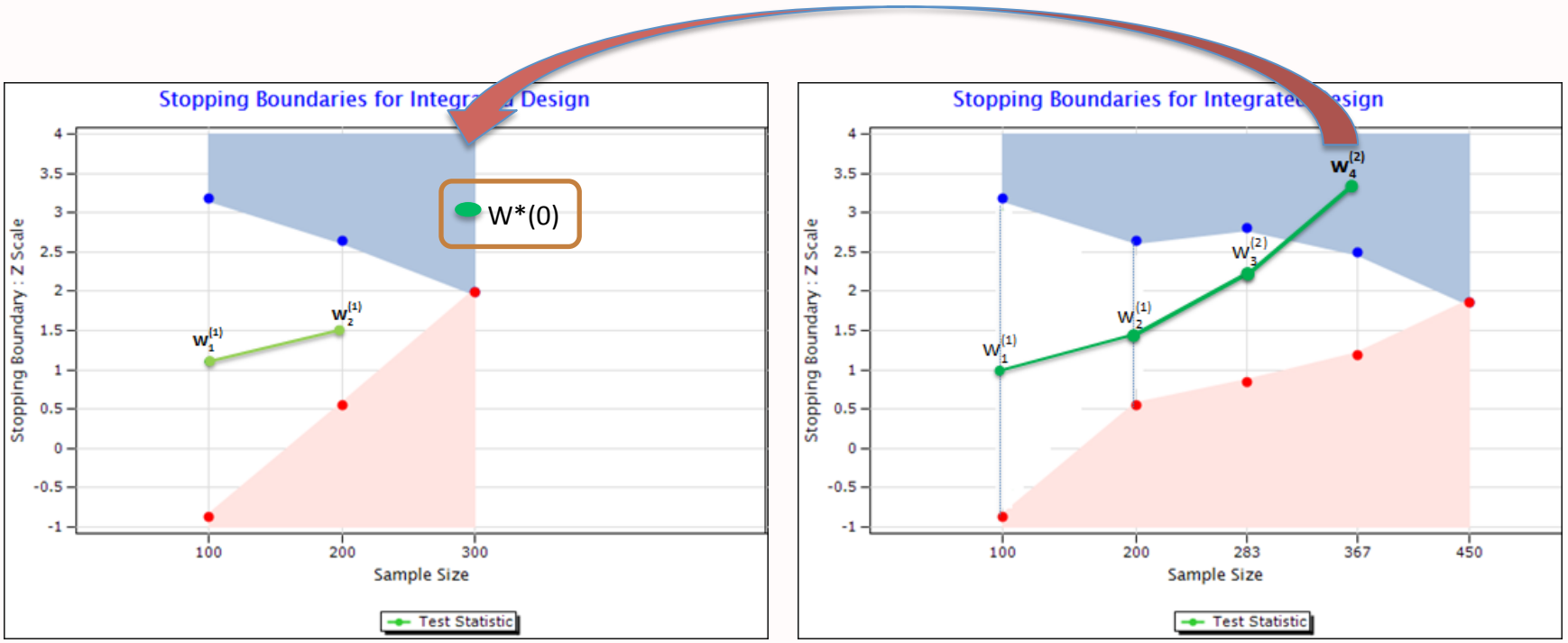
P-value and Confidence Interval Computations





1. **Find the backward image, in the classical sample space, of the statistic $w_4^{(2)}$ obtained in the adaptive sample space**
2. **Use the classical SWCI method to compute P-value and CI for the backward image**

Backward Image, w*(0), for Obtaining the P-Value

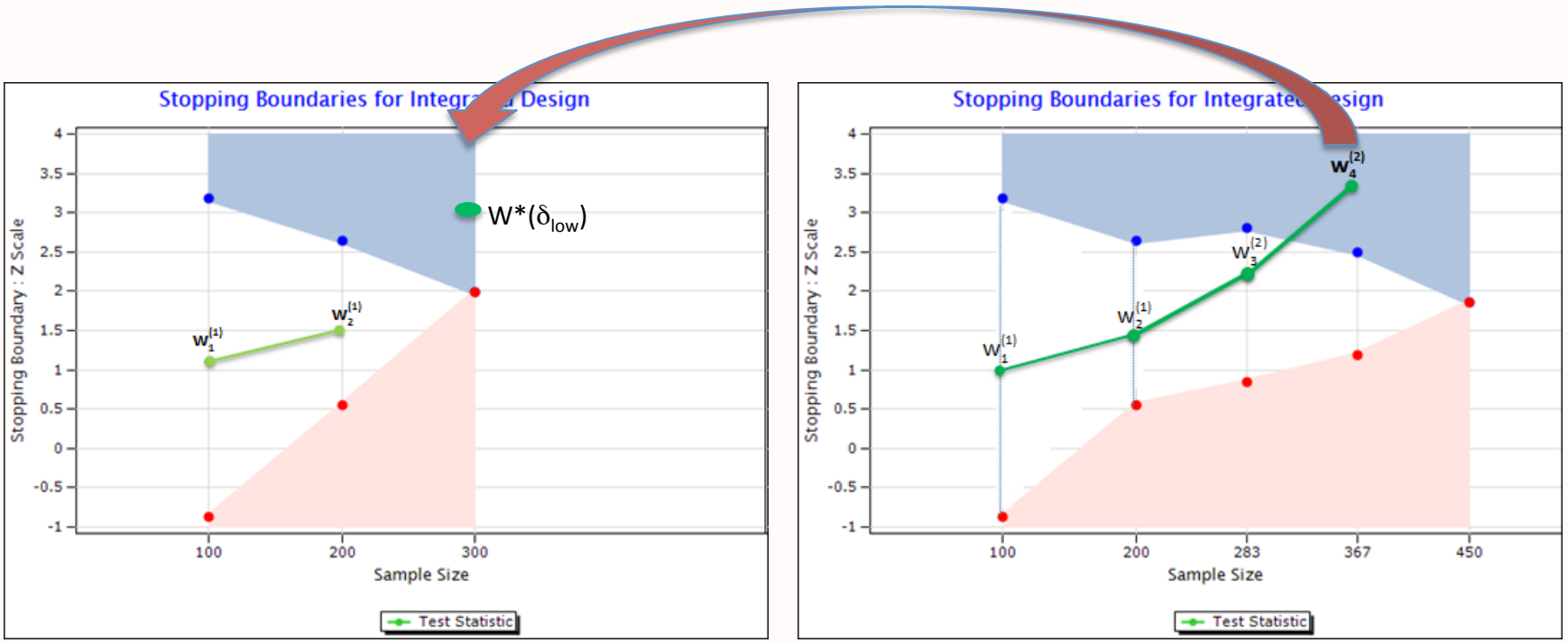


- Find w*(0) such that:

$$P_0(W_3^{(1)} \geq w^*(0) \mid w_2^{(1)}) = P_0(W_3^{(2)} \geq b_3^{(2)} \text{ or } W_4^{(3)} \geq w_4^{(3)} \mid w_2^{(1)})$$

- Then find the p-value corresponding to w*(0) by classical SWCI method

Backward Image, $w^*(\delta_{low})$, for Lower Confidence Bound, δ_{low}



1. $P_{\delta low}(W_3^{(1)} \geq w^*(\delta_{low}) \mid w_2^{(1)}) = P_{\delta low}(W_3^{(2)} \geq b_3^{(2)}, W_4^{(3)} \geq w_4^{(3)} \mid w_2^{(1)})$

2. $P_{\delta low}(W_1^{(1)} \geq b_1^{(1)} \text{ or } W_2^{(1)} \geq b_2^{(1)} \text{ or } W_3^{(1)} \geq w^*(\delta_{low})) = 0.975$

Results for Backward Image (BWCI) Method

Cytel

Simulation of a 3-look LD(OF) GSD with adaption at look 2 to a 3-look LD(PK) GSD (100,000 simulations)

True Value of δ	Median of 100,000 estimates of δ	Prop. of 90% CI's containing δ	Proportion of 90% CI's excluding δ	
			From below	From above
-0.15	-0.14972	0.90007	0.05022	0.04971
0.00	0.00027	0.90073	0.04920	0.05007
0.15	0.14986	0.89866	0.04955	0.05179
0.30	0.29999	0.90087	0.04940	0.04973
0.45	0.44963	0.89929	0.05083	0.04988

Comparison with other adaptive methods

- Simulation of a 3-look LD(OF) GSD with adaption at look 1 to a 2-look LD(OF) GSD (100,000 simulations)

True value of δ	Probability of lower 95% CL > δ			Probability of upper 95% CL < δ		
	BWCI	SWCI	RCI	BWCI	SWCI	RCI
-0.15	0.02505	0.0256	0.01905	0.02529	NA	0.02324
0.0	0.02462	0.0251	0.02448	0.02524	NA	0.02339
0.15	0.02473	0.0256	0.02585	0.02511	NA	0.02238
0.3	0.02411	0.0253	0.00654	0.02527	NA	0.01749
0.45	0.02470	0.0259	0.00075	0.02594	NA	0.01050

- SWCI Method: Brannnath, Mehta, Posch (*Biometrics*, 2009)
- RCI Method: Mehta, Bauer, Posch, Brannath (*Statist. Med.*, 2007)

Summary of Results

Cytel

- BWCI method is the only published method with exact two-sided coverage
- Applicable to SSR, alteration of number and spacing of interim looks and spending function
- SWCI method (Brannath et. al. 2009) provides comparable results but only for one-sided case
- RCI methods:
  - Lehmacher & Wassmer (1999) restricted to SSR
  - Mehta et. al. (2007) permits any type of adaptation
  - Both methods are conservative and asymmetric

Future Plans: Extend to MAMS

Cytel

- Special Case:
  - Select the arm that has the highest response at stage 1 (Todd & Stallard; hypothesis test only)
  - Inference by backward image is under review
- General Case:
  - Drop losers at each stage and continue (Gao, Liu, Mehta, 2014; hypothesis test only)
  - Inference method is under investigation