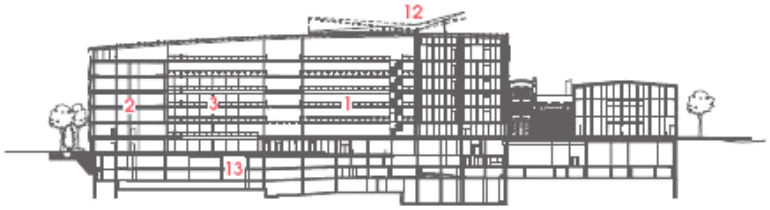
Group Sequential and Adaptive Design for Confirmatory Trials using East

Pantelis Vlachos

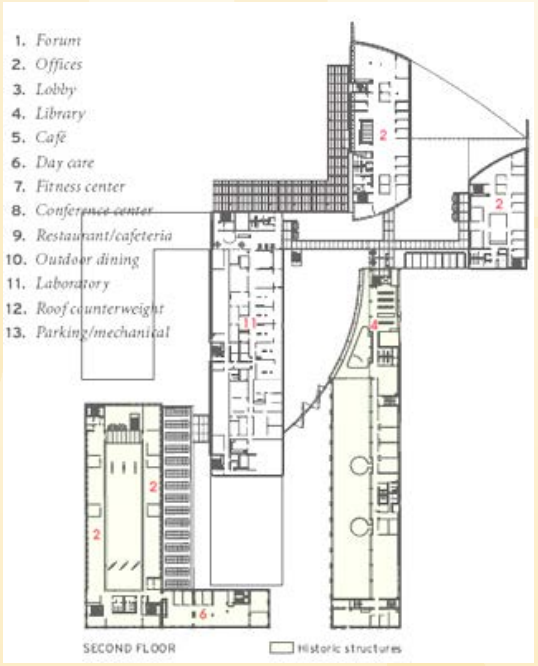
Pantelis.Vlachos@cytel.com

ROeS 2013, Dornbirn

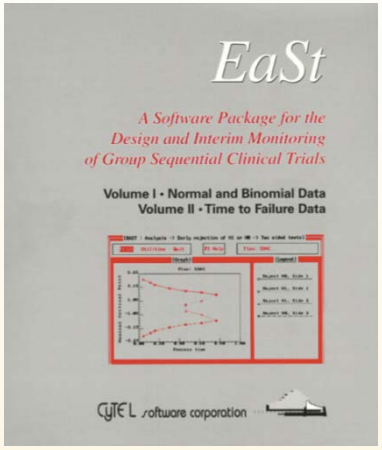


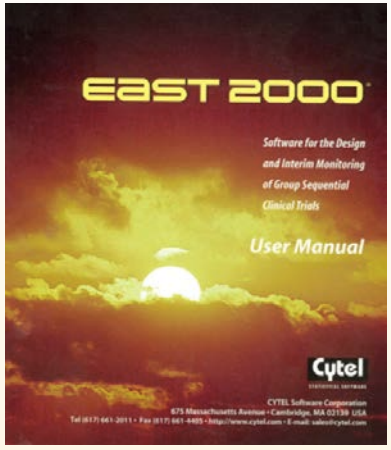
SECTION A-A

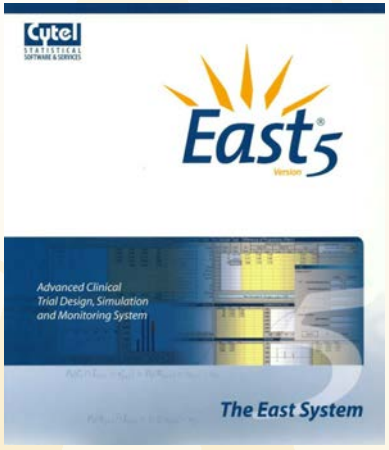
The design of a clinical trial is the **blueprint** for its eventual success



Where are we today?

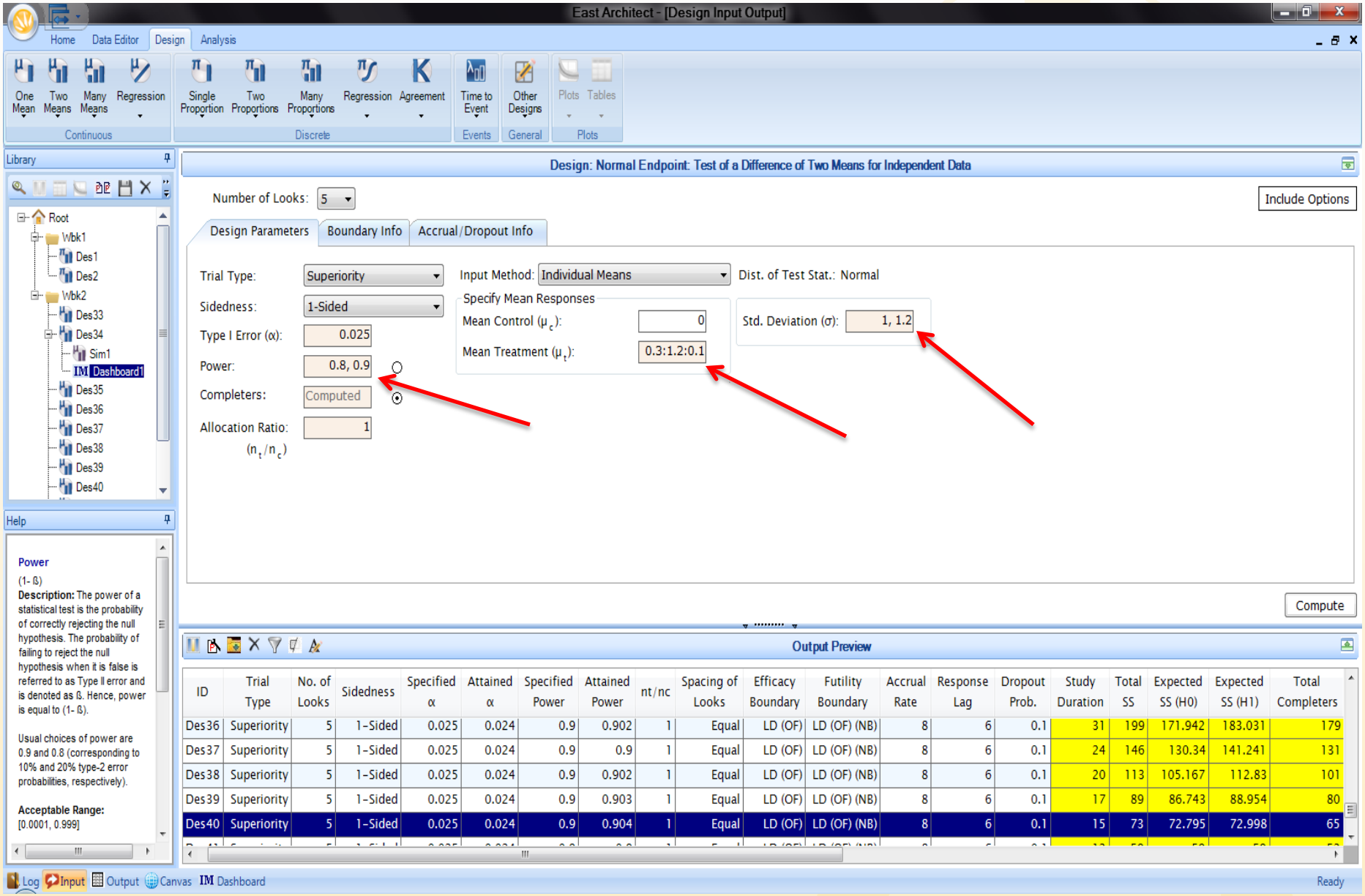






The Industry Standard

Quickly create multiple designs



ID	Trial Type	No. of Looks	Sidedness	Specified α	Attained α	Specified Power	Attained Power	nt/nc	Spacing of Looks	Efficacy Boundary	Futility Boundary	Accrual Rate	Response Lag	Dropout Prob.	Study Duration	Total SS	Expected SS (H0)	Expected SS (H1)	Total Completers
Des36	Superiority	5	1-Sided	0.025	0.024	0.9	0.902	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	31	199	171.942	183.031	179
Des37	Superiority	5	1-Sided	0.025	0.024	0.9	0.9	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	24	146	130.34	141.241	131
Des38	Superiority	5	1-Sided	0.025	0.024	0.9	0.902	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	20	113	105.167	112.83	101
Des39	Superiority	5	1-Sided	0.025	0.024	0.9	0.903	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	17	89	86.743	88.954	80
Des40	Superiority	5	1-Sided	0.025	0.024	0.9	0.904	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	15	73	72.795	72.998	65

Preview, sort, and filter designs

East Architect - [Design Input Output]

Home | Data Editor | Design | Analysis

One Mean | Two Means | Many Means | Regression — Continuous
Single Proportion | Two Proportions | Many Proportions | Regression | Agreement — Discrete
Time to Event — Events
Other Designs — General
Plots | Tables — Plots

Library

- Root
  - Wbk1
    - Des1
    - Des2
  - Wbk2
    - Des33
    - Des34
      - Sim1
      - Dashboard1
    - Des35
    - Des36
    - Des37
    - Des38
    - Des39
    - Des40

Help

Boundary Family

Description: Boundaries are the generalization of the critical values of a test to group sequential methods, when the test statistic value crosses the boundaries, it is indicative of enough evidence to reject H0 or H1. There are various families of boundaries specified in the literature which allow the user to specify how conservatively or aggressively tests are to be performed at each interim look and in the final analysis while preserving the type-1 error.

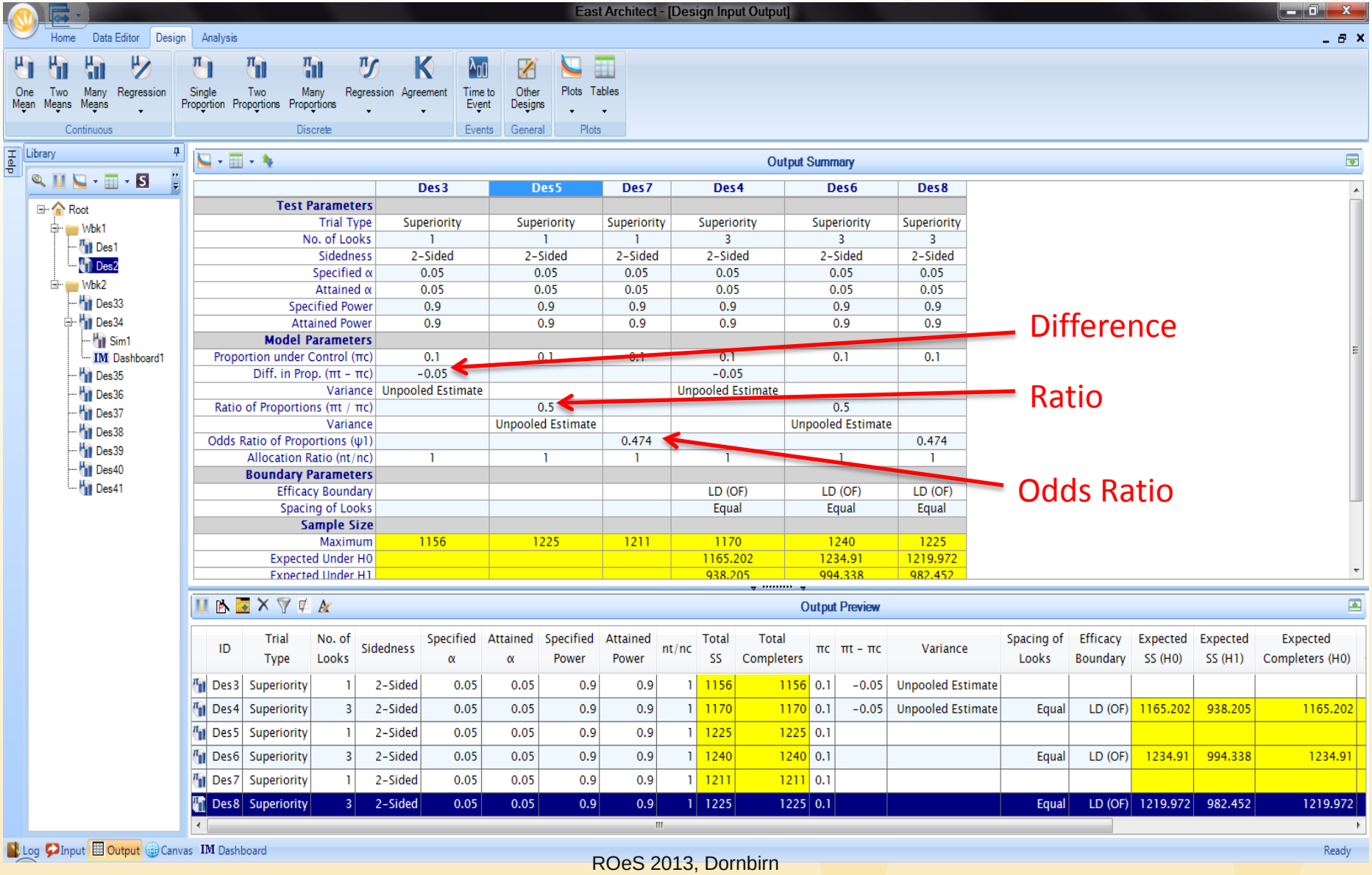
Acceptable Range: Haybittle-Peto (p-value), Wang-Tsiatis, Published Spending

Output Preview

ID	Trial Type	No. of Looks	Sidedness	Specified α	Attained α	Specified Power	Attained Power	nt/nc	Spacing of Looks	Efficacy Boundary	Futility Boundary	Accrual Rate	Response Lag	Dropout Prob.	Study Duration	Total SS	Expected SS (H0)	Expected SS (H1)	Total Completers
Des11	Superiority	5	1-Sided	0.025	0.023	0.8	0.808	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	10	33	33	33	29
Des12	Superiority	5	1-Sided	0.025	0.023	0.8	0.818	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	10	28	28	28	25
Des13	Superiority	5	1-Sided	0.025	0.023	0.8	0.801	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	83	616	513.116	539.54	554
Des14	Superiority	5	1-Sided	0.025	0.023	0.8	0.801	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	49	347	289.201	309.685	312
Des15	Superiority	5	1-Sided	0.025	0.023	0.8	0.802	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	34	223	187.161	204.575	200
Des16	Superiority	5	1-Sided	0.025	0.023	0.8	0.802	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	25	155	134.423	148.152	139
Des17	Superiority	5	1-Sided	0.025	0.023	0.8	0.802	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	20	114	102.649	113.14	102
Des18	Superiority	5	1-Sided	0.025	0.023	0.8	0.801	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	17	87	83.496	86.774	78
Des19	Superiority	5	1-Sided	0.025	0.023	0.8	0.804	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	15	69	68.454	68.977	62
Des20	Superiority	5	1-Sided	0.025	0.023	0.8	0.802	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	13	56	56	56	50
Des21	Superiority	5	1-Sided	0.025	0.023	0.8	0.81	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	12	47	47	47	42
Des22	Superiority	5	1-Sided	0.025	0.023	0.8	0.806	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	11	39	39	39	35
Des23	Superiority	5	1-Sided	0.025	0.024	0.9	0.9	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	75	550	464.489	476.661	495
Des24	Superiority	5	1-Sided	0.025	0.024	0.9	0.901	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	45	310	262.524	274.969	279
Des25	Superiority	5	1-Sided	0.025	0.024	0.9	0.902	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	31	199	171.942	183.031	179
Des26	Superiority	5	1-Sided	0.025	0.024	0.9	0.901	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	23	138	124.427	134.351	124
Des27	Superiority	5	1-Sided	0.025	0.024	0.9	0.9	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	19	102	96.748	101.883	91
Des28	Superiority	5	1-Sided	0.025	0.024	0.9	0.902	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	16	78	77.698	77.997	70
Des29	Superiority	5	1-Sided	0.025	0.024	0.9	0.9	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	14	62	61.978	62	55
Des30	Superiority	5	1-Sided	0.025	0.024	0.9	0.903	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	12	50	50	50	45
Des31	Superiority	5	1-Sided	0.025	0.024	0.9	0.902	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	11	42	42	42	37
Des32	Superiority	5	1-Sided	0.025	0.024	0.9	0.9	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	10	35	35	35	31
Des33	Superiority	5	1-Sided	0.025	0.024	0.9	0.9	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	105	793	669.582	681.678	713
Des34	Superiority	5	1-Sided	0.025	0.024	0.9	0.9	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	62	446	376.015	389.512	401
Des35	Superiority	5	1-Sided	0.025	0.024	0.9	0.901	1	Equal	LD (OF)	LD (OF) (NB)	8	6	0.1	42	286	241.586	254.664	257

Log | Input | Output | Canvas | Dashboard — Ready

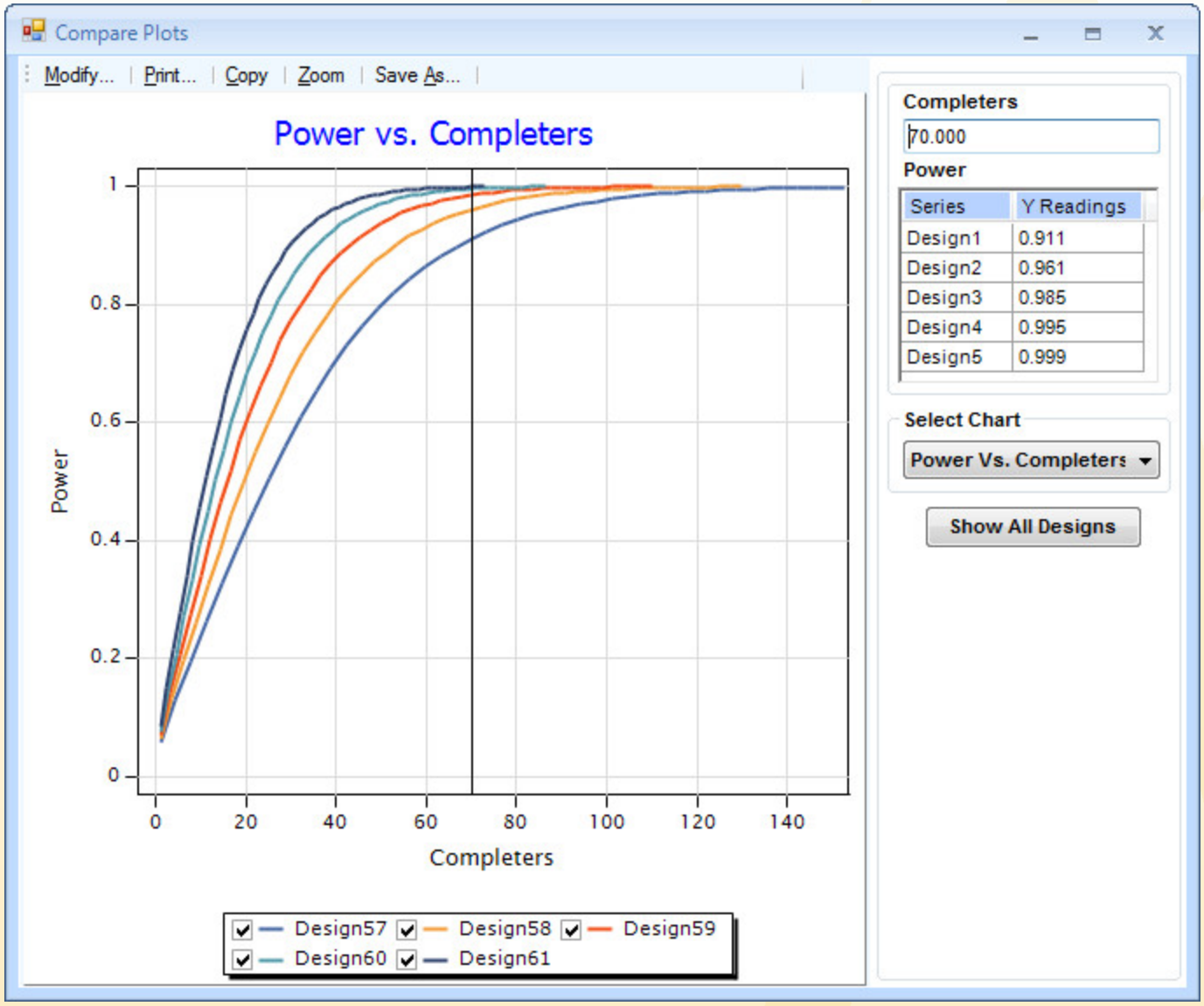
Compare trials of different types and endpoints



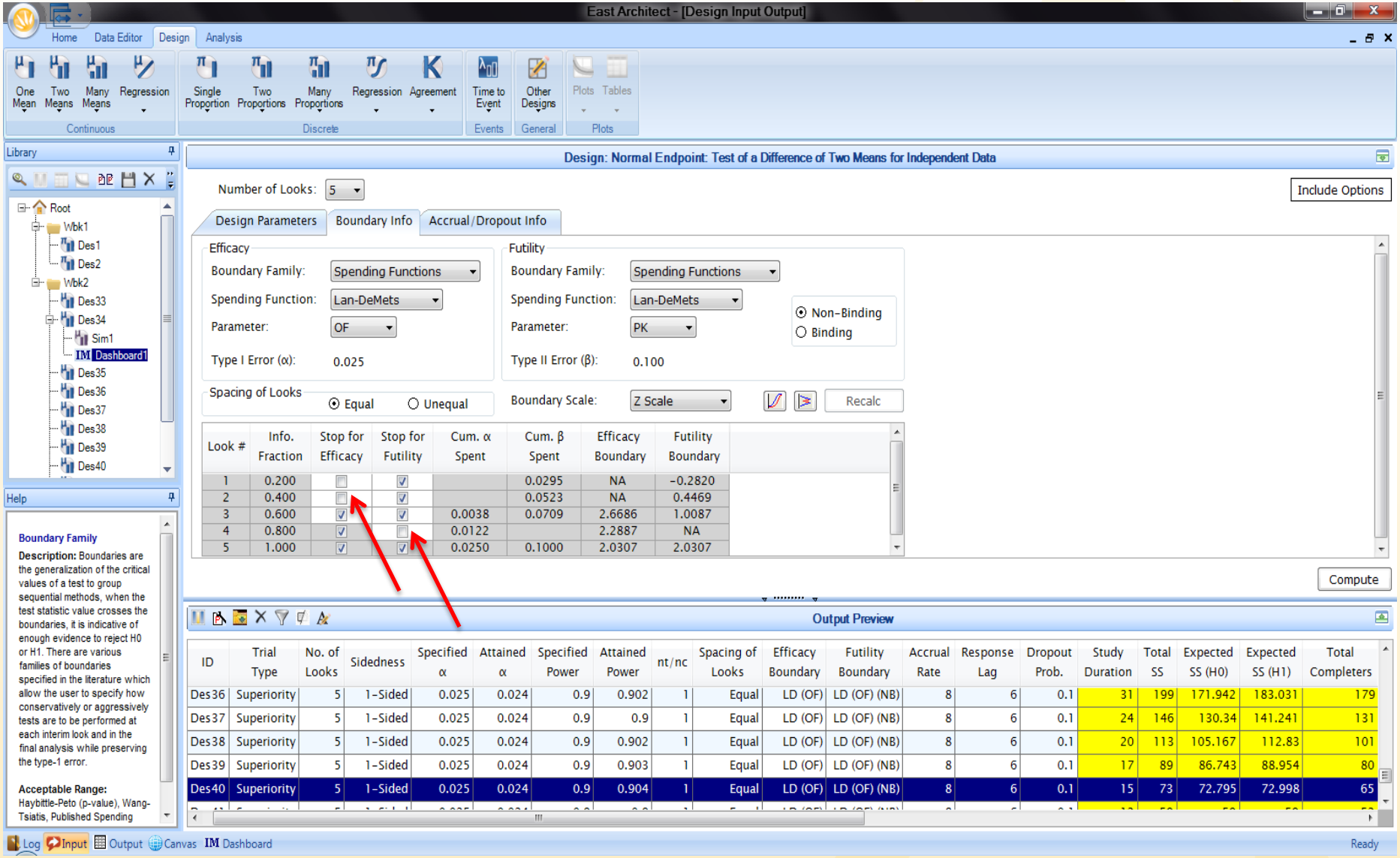
	Des3	Des5	Des7	Des4	Des6	Des8
Test Parameters						
Trial Type	Superiority	Superiority	Superiority	Superiority	Superiority	Superiority
No. of Looks	1	1	1	3	3	3
Sidedness	2-Sided	2-Sided	2-Sided	2-Sided	2-Sided	2-Sided
Specified α	0.05	0.05	0.05	0.05	0.05	0.05
Attained α	0.05	0.05	0.05	0.05	0.05	0.05
Specified Power	0.9	0.9	0.9	0.9	0.9	0.9
Attained Power	0.9	0.9	0.9	0.9	0.9	0.9
Model Parameters						
Proportion under Control (πc)	0.1	0.1	0.1	0.1	0.1	0.1
Diff. in Prop. (πt − πc)	-0.05			-0.05		
Variance	Unpooled Estimate			Unpooled Estimate		
Ratio of Proportions (πt / πc)		0.5			0.5	
Variance		Unpooled Estimate			Unpooled Estimate	
Odds Ratio of Proportions (ψ1)			0.474			0.474
Allocation Ratio (nt/nc)	1	1	1	1	1	1
Boundary Parameters						
Efficacy Boundary				LD (OF)	LD (OF)	LD (OF)
Spacing of Looks				Equal	Equal	Equal
Sample Size						
Maximum	1156	1225	1211	1170	1240	1225
Expected Under H0				1165.202	1234.91	1219.972
Expected Under H1				938.205	994.338	982.452

ID	Trial Type	No. of Looks	Sidedness	Specified α	Attained α	Specified Power	Attained Power	nt/nc	Total SS	Total Completers	πc	πt − πc	Variance	Spacing of Looks	Efficacy Boundary	Expected SS (H0)	Expected SS (H1)	Expected Completers (H0)
Des3	Superiority	1	2-Sided	0.05	0.05	0.9	0.9	1	1156	1156	0.1	-0.05	Unpooled Estimate					
Des4	Superiority	3	2-Sided	0.05	0.05	0.9	0.9	1	1170	1170	0.1	-0.05	Unpooled Estimate	Equal	LD (OF)	1165.202	938.205	1165.202
Des5	Superiority	1	2-Sided	0.05	0.05	0.9	0.9	1	1225	1225	0.1							
Des6	Superiority	3	2-Sided	0.05	0.05	0.9	0.9	1	1240	1240	0.1			Equal	LD (OF)	1234.91	994.338	1234.91
Des7	Superiority	1	2-Sided	0.05	0.05	0.9	0.9	1	1211	1211	0.1							
Des8	Superiority	3	2-Sided	0.05	0.05	0.9	0.9	1	1225	1225	0.1			Equal	LD (OF)	1219.972	982.452	1219.972

Visualize multiple designs on a single chart



Disable boundaries at selected interim looks



A few new features

1. Response lag
2. Bayesian power
3. R functions
4. Stratification
5. Multiple endpoints

Example 1: Schizophrenia Trial

Two arms: asenapine vs. olanzapine (control)

Endpoint: negative symptoms assessment (NSA)

<u>Design Parameters</u>

$\alpha = 0.025$ (1-sided)

$1-\beta = 80\%$

$\delta = 1.6$ to 2

$\sigma = 7.5$

<u>Accrual / Dropout Info</u>

Accrual 8 patients / week

Response lag = 26 weeks

Dropout = 8%

Example 2: Adaptive Re-estimation of Events in a CV Outcomes Trial

- New agents for type-2 diabetes must demonstrate safety to cardiovascular risk
- Primary outcome is major adverse cardiac event (MACE) -- death, MI, stroke
- Non-inferiority trial at one-sided α=0.025. If upper bound of 95% CI for HR is:
  - ≤ 1.8 – file for provisional approval
  - ≤ 1.3 – file for final approval

1.3 Non-Inferiority Design: Assumptions

- Enroll 6500 patients over 3 years
- Complete the study within 5 years
- 90% power to reject H_0: HR≤1.3 at H_1: HR=1 with a 1-sided test at α=0.025
- 3% annualized event rate for control arm
- Interim analysis at 80% of information:

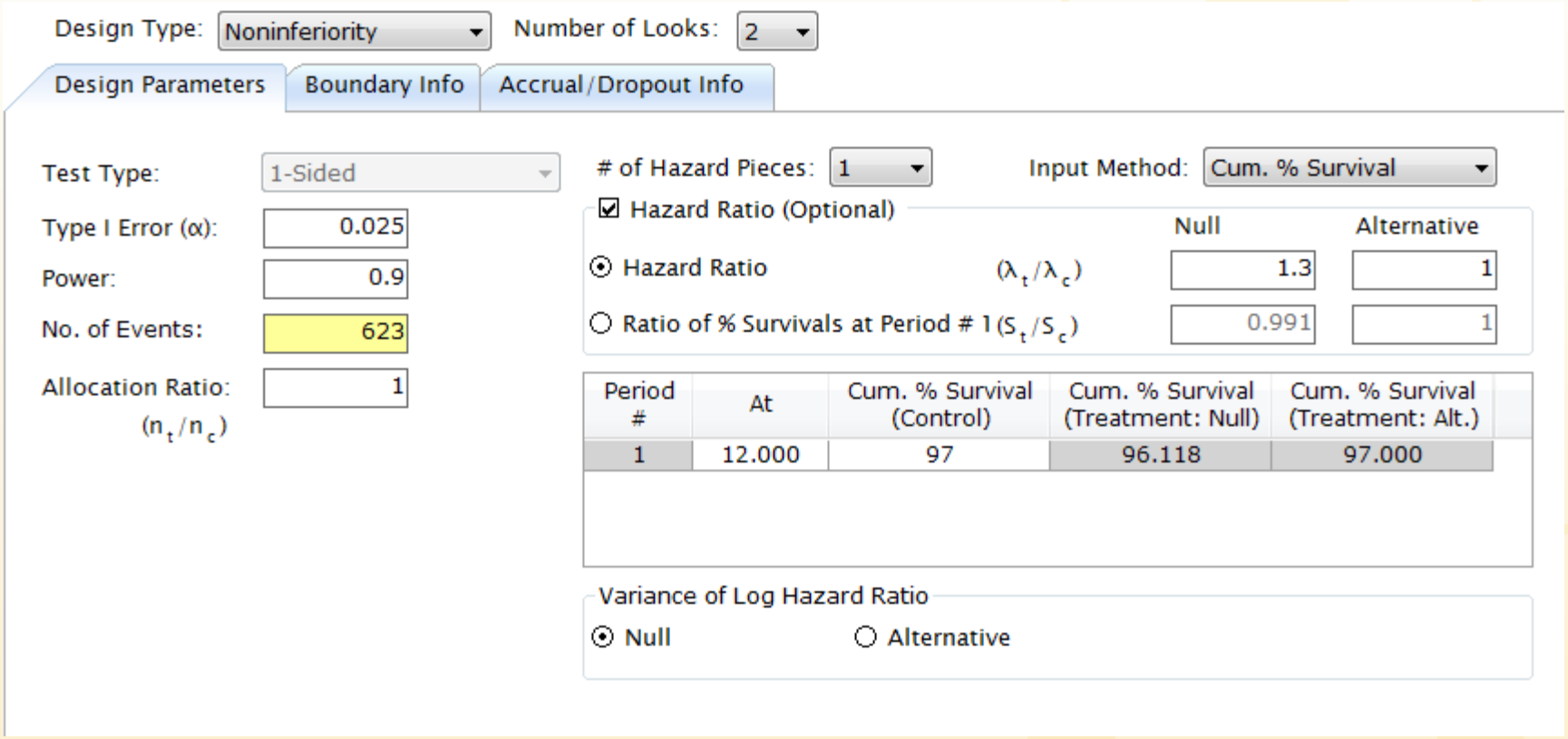
Options available at interim analysis

- If upper 97.5% CI > 1.3, proceed to final analysis
- If upper 97.5% CI is between 1 and 1.3, either:
  1. Stop and file claim for non-inferiority
  2. Continue and test for superiority (HR = 1) at the planned final analysis
  3. Adaptively extend number of events. Test for superiority (HR=1) at the extended final analysis
- If upper 95% CI <1, stop and claim superiority

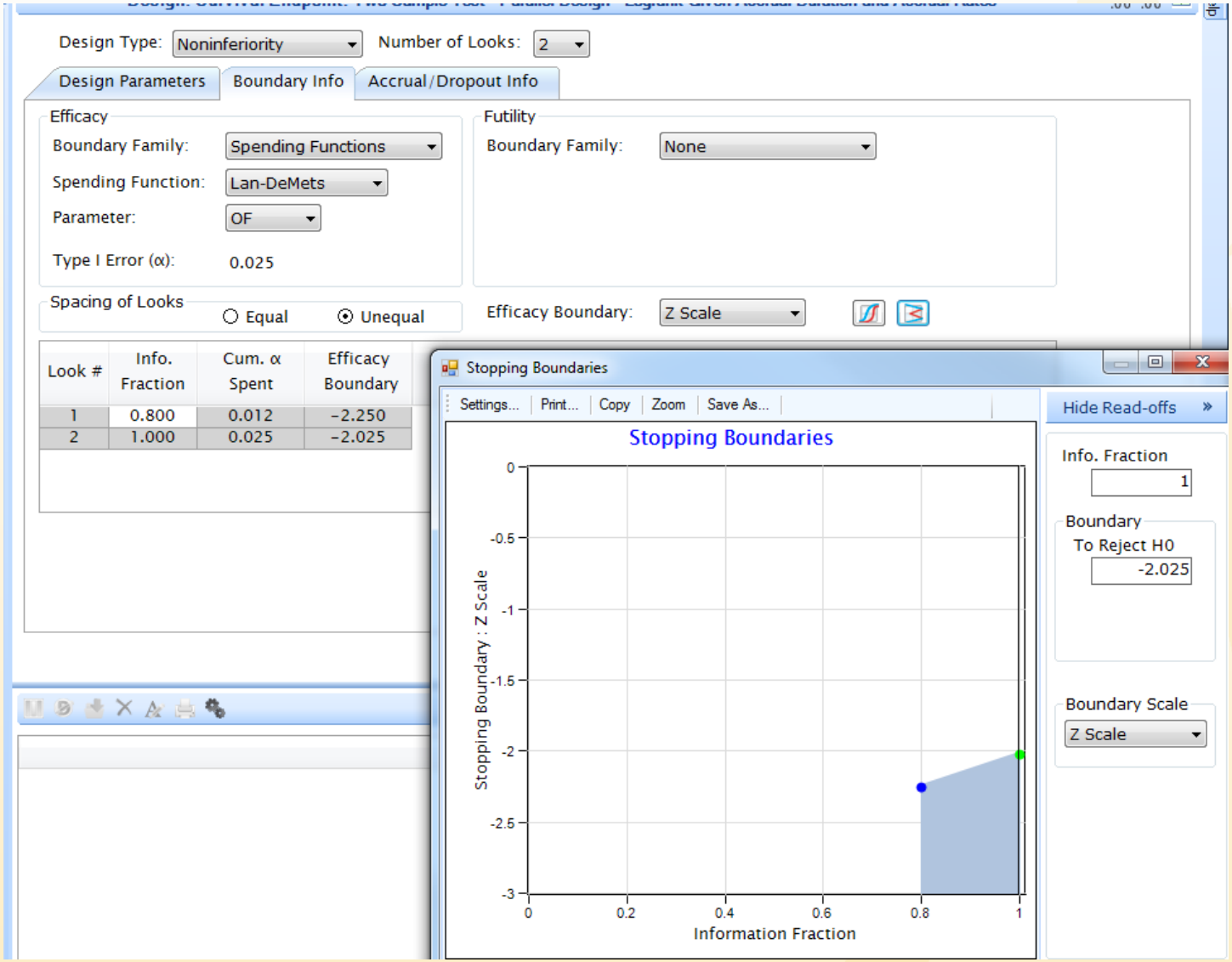
Rationale for Adaptive Option

- Easier to show that new anti-diabetic agents (e.g. the class of DPP4 inhibitors) do not increase CV outcomes than to show that they actually decrease CV outcomes relative to current standard of care (e.g. Insulin, Metformin, Sulfonylureas)
- CV outcome trials powered for superiority require more follow-up and/or more patients than CV outcome trials powered for NI
- **Design up-front for NI. If interim results are promising, then extend follow-up to show Sup.**

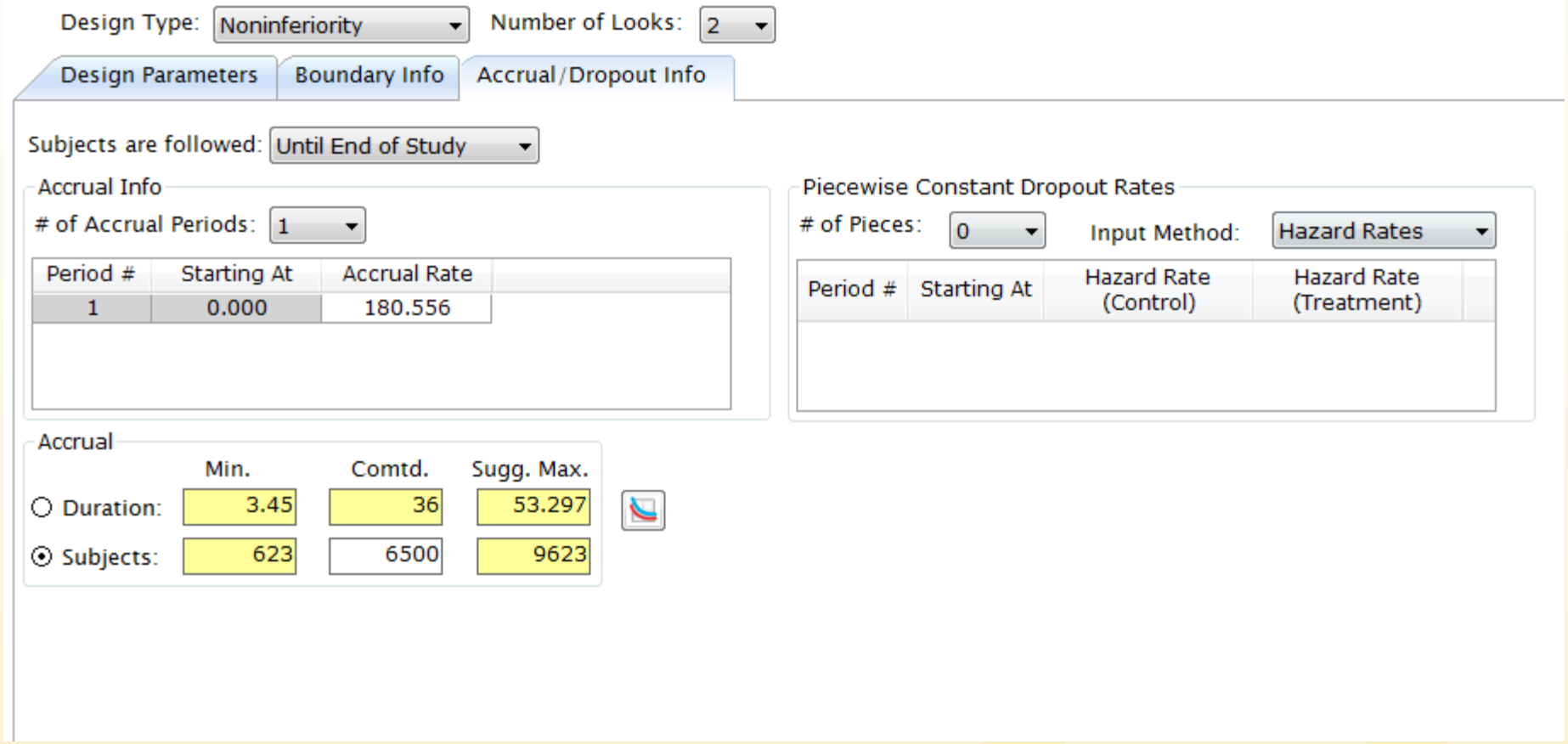
Design Tab for the NI Design



Boundaries Tab for the NI Design



Accrual Tab for NI Design



Design Summary of NI Design

	NI-SUP:NI
Mnemonic	SU-2S-LRAR
Test Parameters	
Design Type	Noninferiority
No. of Looks	2
Test Type	1-Sided
Specified α	0.025
Power	0.9
Model Parameters	
Hazard Ratio (Null)	1.3
Hazard Ratio (Alt.)	1
Var (Log HR)	Null
Allocation Ratio (nt/nc)	1
Boundary Parameters	
Spacing of Looks	Unequal
Efficacy Boundary	LD (OF)
Accrual & Dropout Parameters	
Accrual Rate	180.556
Subjects are Followed	Until End of Study
No. of Accrual Periods	1
No. of Dropout Pieces	0
Sample Size	
Maximum	6500
Expected Under H0	6500
Expected Under H1	6500
Events	
Maximum	623
Expected Under H0	621.478
Expected Under H1	529.192
Accrual Duration	
Maximum	36
Expected Under H0	36
Expected Under H1	36
Study Duration	
Maximum	57.832
Expected Under H0	52.619
Expected Under H1	51.609

- 623 events
- 6500 patients
- 58 months max study duration

Design Details of NI Design

Design: Survival Endpoint: Two-Sample Test - Parallel Design - Logrank Given Accrual Duration

Test Parameters	
Design ID:	NI
Design Type:	Noninferiority
Number of Looks:	2
Test Type:	1-Sided
Specified α:	0.025
Power:	0.9

Model Parameters	
HR = λ_t/λ_c	
Under H0:	1.3
Under H1:	1
Ratio of % Surv. at Period #1:	1
Var (Log HR):	Null
Allocation Ratio (n_t/n_c):	1

Boundary Parameters	
Spacing of Looks:	Unequal
Efficacy Boundary:	LD (OF)

Accrual/Dropout Parameters	
Accrual Duration:	36
Max Study Duration:	57.832
Dropout:	No

Stopping Boundaries: Look by Look

Look #	Info. Fraction (s/s_max)	Events (s)	Cumulative α Spent	Boundaries
				Efficacy Z
1	0.799	498	0.012	-2.251
2	1	623	0.025	-2.025

Events, Sample Size, Pipeline and Analysis Times: Look by Look (Under H0)

Look #	Info. Fraction (s/s_max)	Sample Size (n)	Events (s)	Pipeline (n-s)	Analysis Time	Incr. Boundary Crossing Prob.
						Under H0
						Efficacy
1	0.799	6500	498	6002	45.486	0.012
2	1	6500	623	5877	52.707	0.013

Events, Sample Size, 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
1	0.799	6500	498	6002	49.54	0.75
2	1	6500	623	5877	57.832	0.15

Survival Info. : Cum. % Survival

Period #	At	Cum. % Survival			Hazard Ratio	
		Control (λc)	Treatment (λt0)	Treatment (λt1)	Null (λt0/λc)	Alt. (λt1/λc)
1	12	97	96.118	97	1.3	1

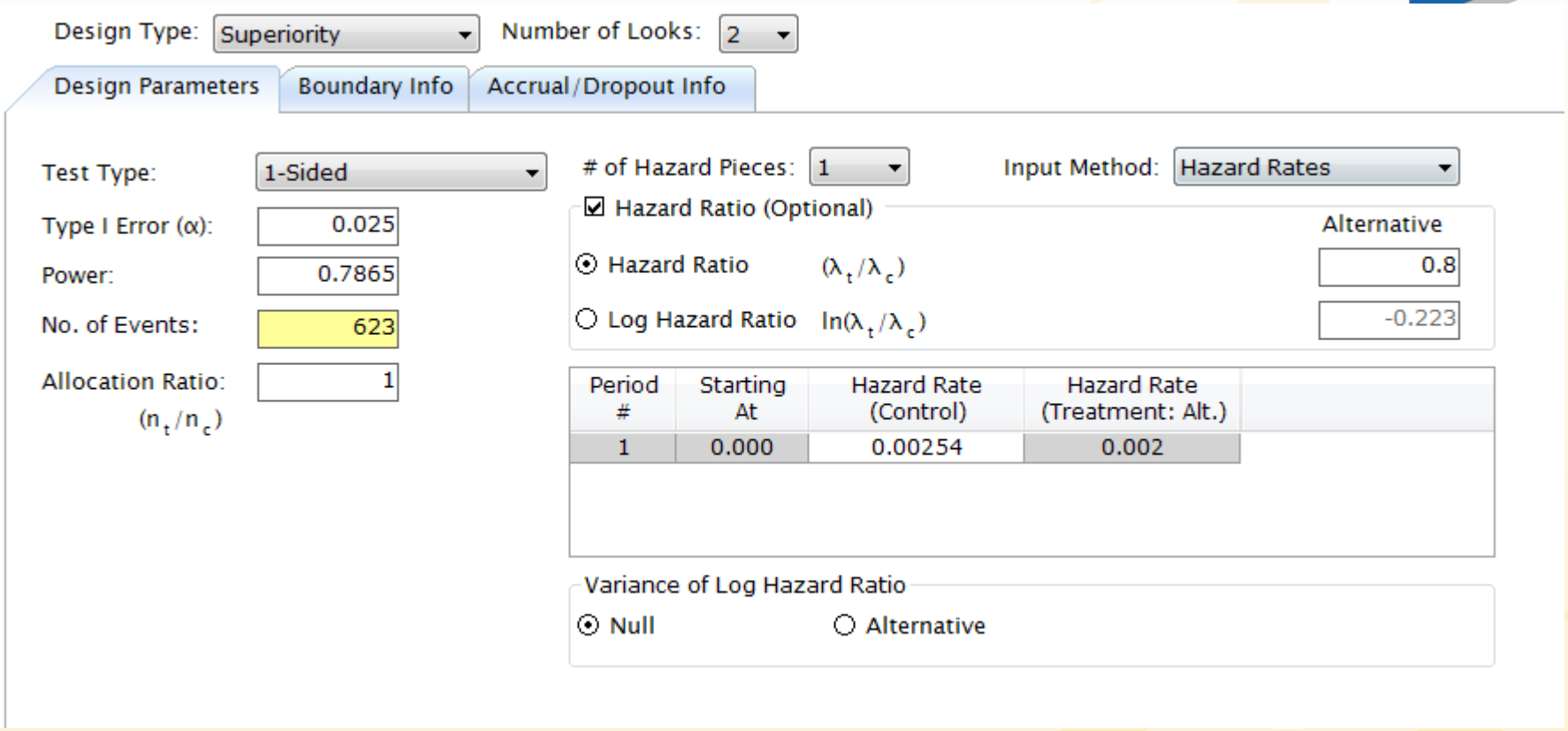
Switch to Superiority Trial

Create a superiority design in East so as to have:

- 623 events (like NI)
- 6500 patients enrolled over 36 months (like NI)
- 3% annualized event rate for control arm (like NI)
- Interim analysis at 80% of information (like NI)
- One-sided α=0.025 to reject H_0=1
- Powered at HR=0.8

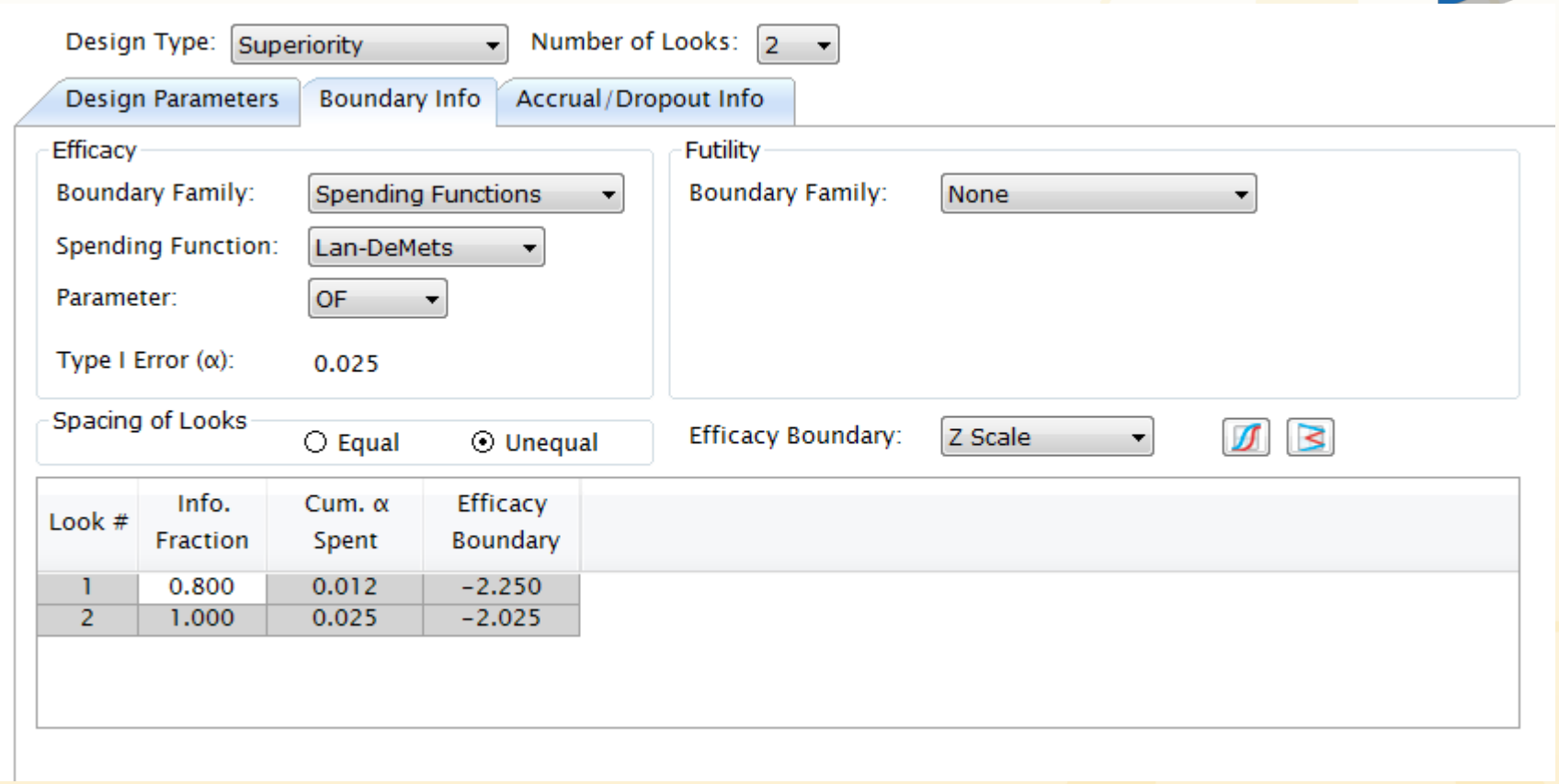
(Decrease power by trial & error until 623 events)

Design Tab for Superiority Trial



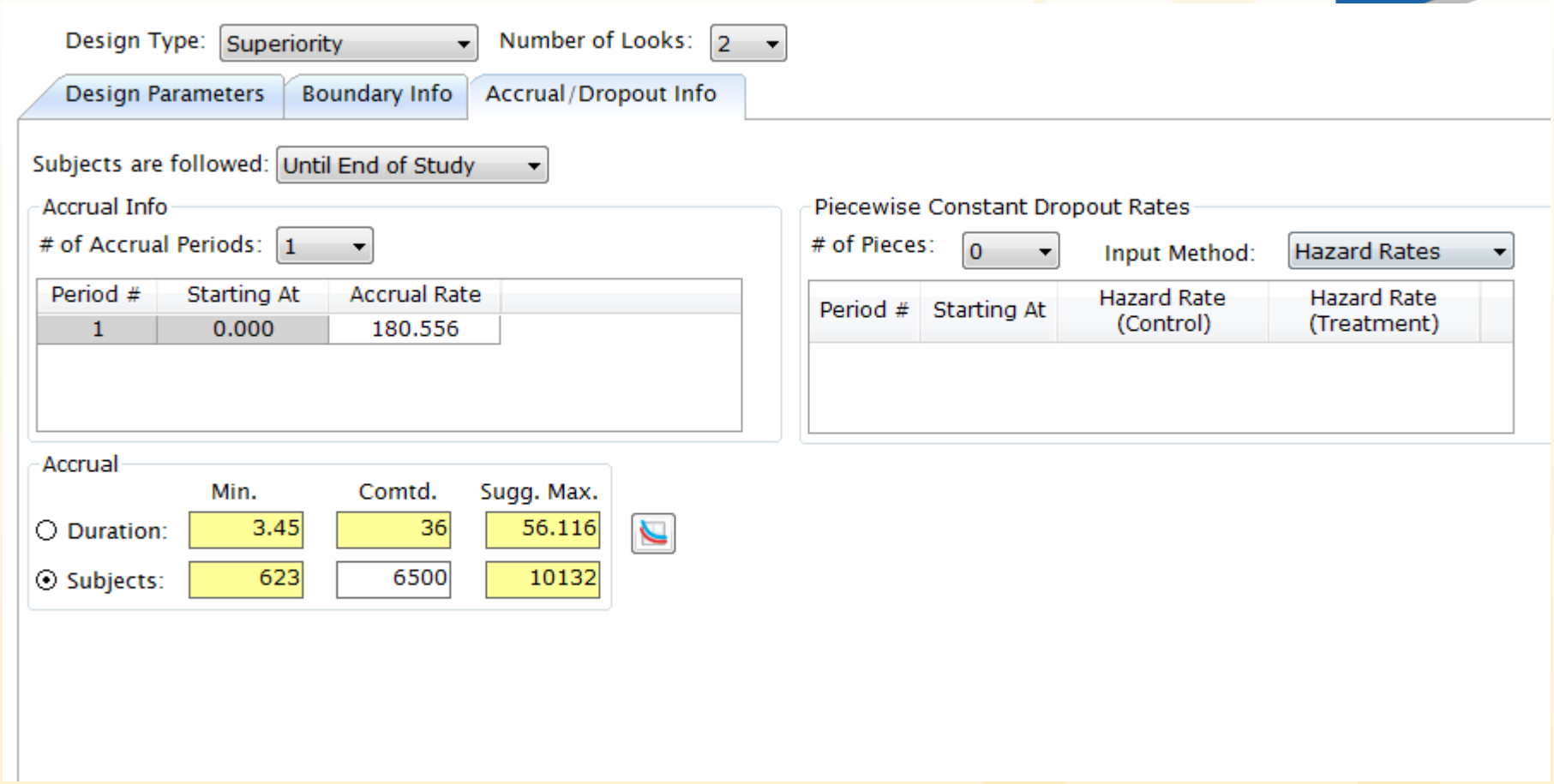
623 events, α= 0.025 and 3% annualized event rate -- just like NI design

Boundary Tab for Superiority Design



Interim analysis at 80% of information. Same as NI design

Accrual Tab for Superiority Design



6500 patients enrolled over 3 years. Same as for NI design

Summary Comparison of NI and Sup

	NI-SUP:NI	NI-SUP:SUP
Mnemonic	SU-2S-LRAR	SU-2S-LRAR
Test Parameters		
Design Type	Noninferiority	Superiority
No. of Looks	2	2
Test Type	1-Sided	1-Sided
Specified α	0.025	0.025
Power	0.9	0.787
Model Parameters		
Hazard Ratio (Null)	1.3	
Hazard Ratio (Alt.)	1	0.8
Var (Log HR)	Null	Null
Allocation Ratio (nt/nc)	1	1
Boundary Parameters		
Spacing of Looks	Unequal	Unequal
Efficacy Boundary	LD (OF)	LD (OF)
Accrual & Dropout Parameters		
Accrual Rate	180.556	180.556
Subjects are Followed	Until End of Study	Until End of Study
No. of Accrual Periods	1	1
No. of Dropout Pieces	0	0
Sample Size		
Maximum	6500	6500
Expected Under H0	6500	6500
Expected Under H1	6500	6500
Events		
Maximum	623	623
Expected Under H0	621.478	621.478
Expected Under H1	529.192	548.727
Accrual Duration		
Maximum	36	36
Expected Under H0	36	36
Expected Under H1	36	36
Study Duration		
Maximum	57.832	62.257
Expected Under H0	52.619	57.731
Expected Under H1	51.609	56.777

Comparison of Sup and NI

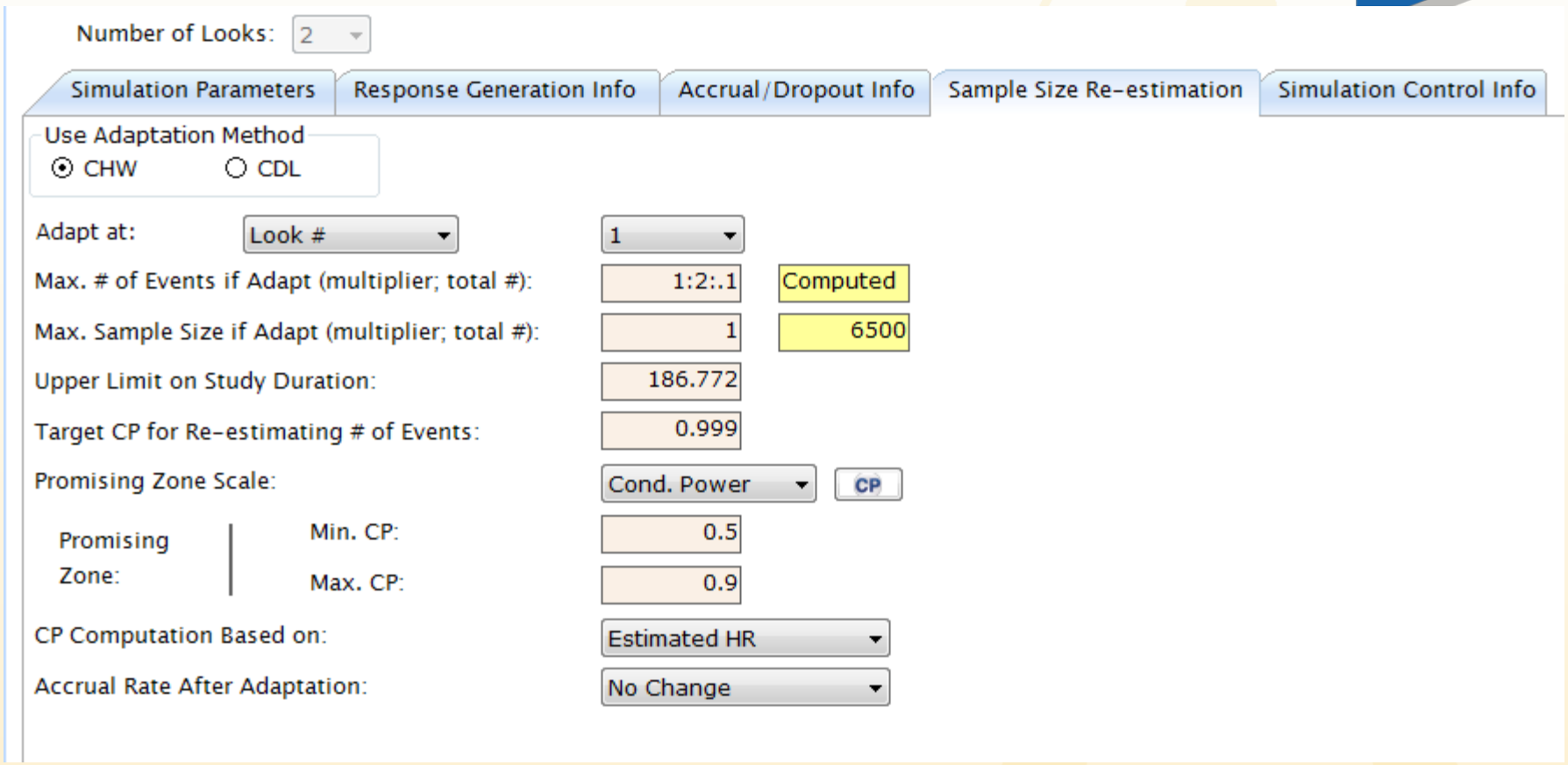
- Same number of events
- Same sample size
- Hence longer maximum and expected study durations for Sup
- Only 79% power at HR = 0.8 for Sup
- Permit adaptive increase of events to boost power

Adaptive Strategy

- Perform interim analysis at 600 events
- Is 1 < (upper 97.5% confidence bound) < 1.3?
  - compute the conditional power for superiority at 650 events
  - if CP is in promising zone, permit a one-time increase in number of events
- Explore by simulation the impact on study duration of various decision rules

Note: Sample size increase is not an option

Select a range of Adaptive Options



Advantage: can submit multiple inputs as a single batch job (11 designs with 10,000 simulations/design took 15 minutes

Summary results from 11 trials

	NI-SUP:SUP:CHWSim1	NI-SUP:SUP:CHWSim6	NI-SUP:SUP:CHWSim7	NI-SUP:SUP:CHWSim8	NI-SUP:SUP:CHWSim9
Mnemonic	SU-2S-LRAR	SU-2S-LRAR	SU-2S-LRAR	SU-2S-LRAR	SU-2S-LRAR
Test Parameters					
Design Type	Superiority	Superiority	Superiority	Superiority	Superiority
Test Type	1-Sided	1-Sided	1-Sided	1-Sided	1-Sided
Test Statistic	Logrank	Logrank	Logrank	Logrank	Logrank
Power	0.784	0.794	0.805	0.806	0.814
Power (Promising)	0.77	0.857	0.881	0.926	0.949
No. of Looks	2	2	2	2	2
Model Parameters					
No. of Hazard Pieces	1	1	1	1	1
Boundary Parameters					
Efficacy Boundary	User Specified	User Specified	User Specified	User Specified	User Specified
Spacing of Looks	User Specified	User Specified	User Specified	User Specified	User Specified
Accrual & Dropout Parameters					
Followup Duration	Until End of Study	Until End of Study	Until End of Study	Until End of Study	Until End of Study
Accrual Rate	180.556	180.556	180.556	180.556	180.556
No. of Accrual Periods	1	1	1	1	1
Sample Size					
Maximum	6500	6500	6500	6500	6500
Events					
Maximum	623	623	623	623	623
Sample Size Re-estimation Parameters					
Method of Adaptation	Cui-Hung-Wang	Cui-Hung-Wang	Cui-Hung-Wang	Cui-Hung-Wang	Cui-Hung-Wang
Adaptation Stage	Look # : 1	Look # : 1	Look # : 1	Look # : 1	Look # : 1
Max. # of Events if Adapt	623	685	747	809	872
Max. Sample Size if Adapt	6500	6500	6500	6500	6500
Upper Limit on Study Duration	186.772	186.772	186.772	186.772	186.772
Target CP	0.999	0.999	0.999	0.999	0.999
CP Computation Based on	Estimated HR	Estimated HR	Estimated HR	Estimated HR	Estimated HR
Promising Zone	0.5 <= CP < 0.9	0.5 <= CP < 0.9	0.5 <= CP < 0.9	0.5 <= CP < 0.9	0.5 <= CP < 0.9
Study Duration					
Maximum	56.811	57.58	58.266	59.091	59.704
Simulation Results (Promising)					
Average Study Duration	62.314	66.842	71.602	76.345	81.184
Average Sample Size	6500	6500	6500	6500	6500
Average Events	623	685	747	809	872

Details of CHWsim6 (increases events by 10% in the promising zone)

Simulation: Survival Endpoint: Two-Sample Test - Parallel Design - Logrank Given Accrual Duration and Accrual Rates (CHW Sim

Simulation Parameters	
Simulation ID:	CHWSim6
Design Type:	Superiority
Number of Looks:	2
Test Type:	1-Sided
Fix at Each Look:	Total No. of Events
Test Statistic:	Logrank
Average Events:	559.277
Total Accrual Duration:	36
Avg. Power at Termination:	0.794

Sample Size Re-estimation Parameters	
Method of Adaptation:	Cui-Hung-Wang
Adapt At Look No.:	1
Max. # of Events if Adapt:	
Multiplier:	1.1
Total #:	685
Max. Sample Size if Adapt:	
Multiplier:	1
Total #:	6500
Upper Limit on Study Duration:	186.772
Target CP:	0.999
Promising Zone Scale:	Cond. Power
Min. CP:	0.5
Max. CP:	0.9
CP Computation Based on:	Estimated HR
Accrual Rate After Adaptation:	No Change

Simulation Control Parameters	
Starting Seed:	Clock
Number of Simulations:	10000

Simulation Boundaries and Boundary Crossing Probabilities:

Look #	Events	Boundaries Efficacy Lower	Early Stopping For Efficacy	Total Simulations Count	Total Simulations %
1	498	-2.251	5880	5880	58.8
2	623	-2.025	2057	4120	41.2
Total			7937	10000	
%			79.37		

Average Sample Size and Look Times:

Look #	Average Sample Size	Average Events Control	Average Events Treatment	Average Look Time	Average Follow up
1	6500	275.492	222.508	53.046	33.569
2	6500	347.613	299.118	64.035	43.583
Average	6500	309.358	249.919	57.58	37.702

Response Generation Parameters

No. of Hazard Pieces: 1

Input Method: Hazard Rates

Piece #	Starting At	Control	Treatment	Hazard Ratio
1	0	0.00254	0.00203	0.8

Simulation Results by Zone:

Zone	Simulation Rejecting H0 Count	Simulation Rejecting H0 Row %	Simulation Not Rejecting H0 Count	Simulation Not Rejecting H0 Row %	Total Simulations Count	Total Simulations Column %	Average Number of Events	Average Sample Size	Average Accrual Duration	Average Study Duration
Futility	0	0	0	0	0	0	0	0	0	0
Unfavorable	706	27.762	1837	72.238	2543	25.43	623	6500	35.994	62.294
Promising	1351	85.669	226	14.331	1577	15.77	685	6500	35.994	66.842
Favorable	0	0	0	0	0	0	0	0	0	0
Efficacy	5880	100	0	0	5880	58.8	498	6500	35.994	53.057
All Trials	7937	79.37	2063	20.63	10000	100	559.277	6500	35.994	57.58

Some conclusions

- Adaptive change only affects the study duration, not the sample size
- Additional investment of study duration only made if entering the promising zone
- Probability enter promising zone is about 16%
- If entering the promising zone:
  - 6 month study prolongation (685 events) yields 86% power instead of 77% power in the PZ
  - 12 month study prolongation (747 events) yields 88% power instead of 77% power in the PZ

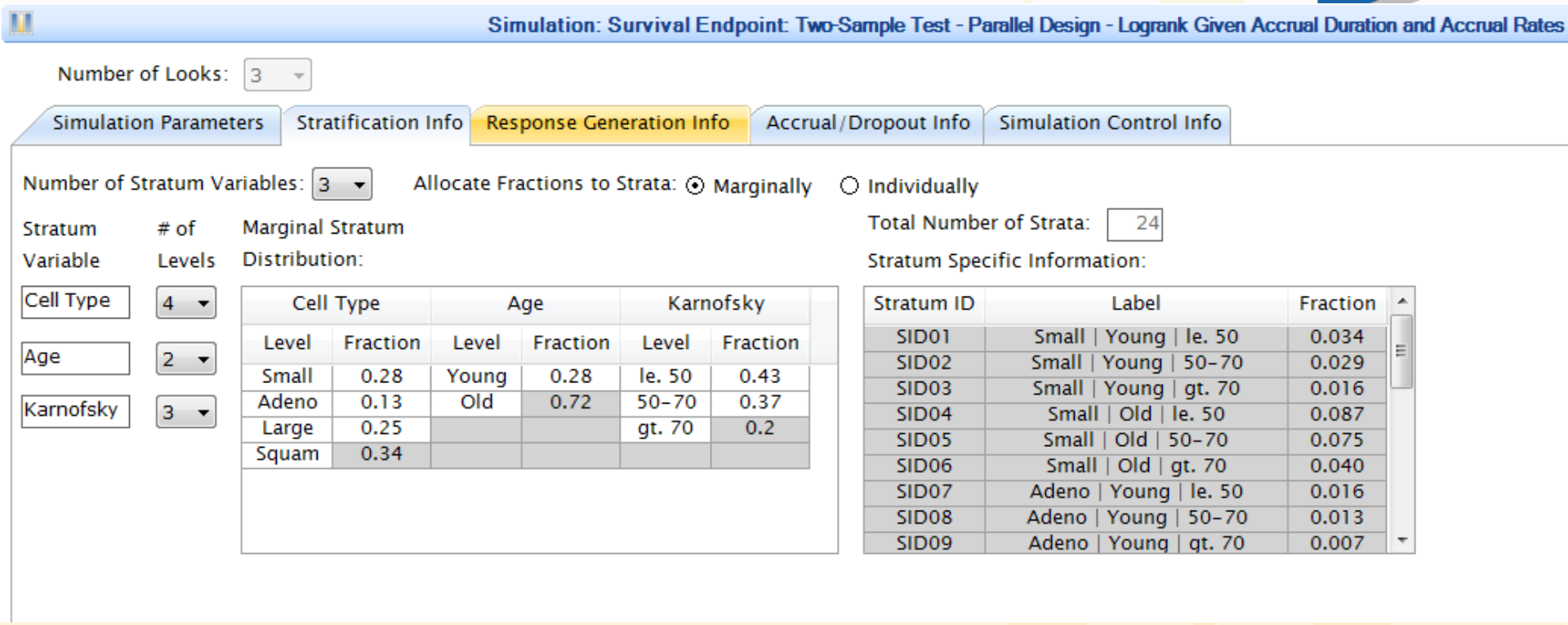
Example 3: Impact of Stratification in a Lung Cancer Clinical Trial

- The base-case design has:
  - 2-sided level 0.05 logrank test
  - 90% power to detect HR = 0.447
  - Baseline hazard (control arm) = 0.009211 (corresponds to 9% five-year survival)
  - 3 equally spaced OBF stopping boundaries
  - Uniform enrollment at 12/week for 24 weeks
- Consider stratification by cell type (4), age (2) and performance status (2)

Kalbfleisch and Prentice (2002) Data

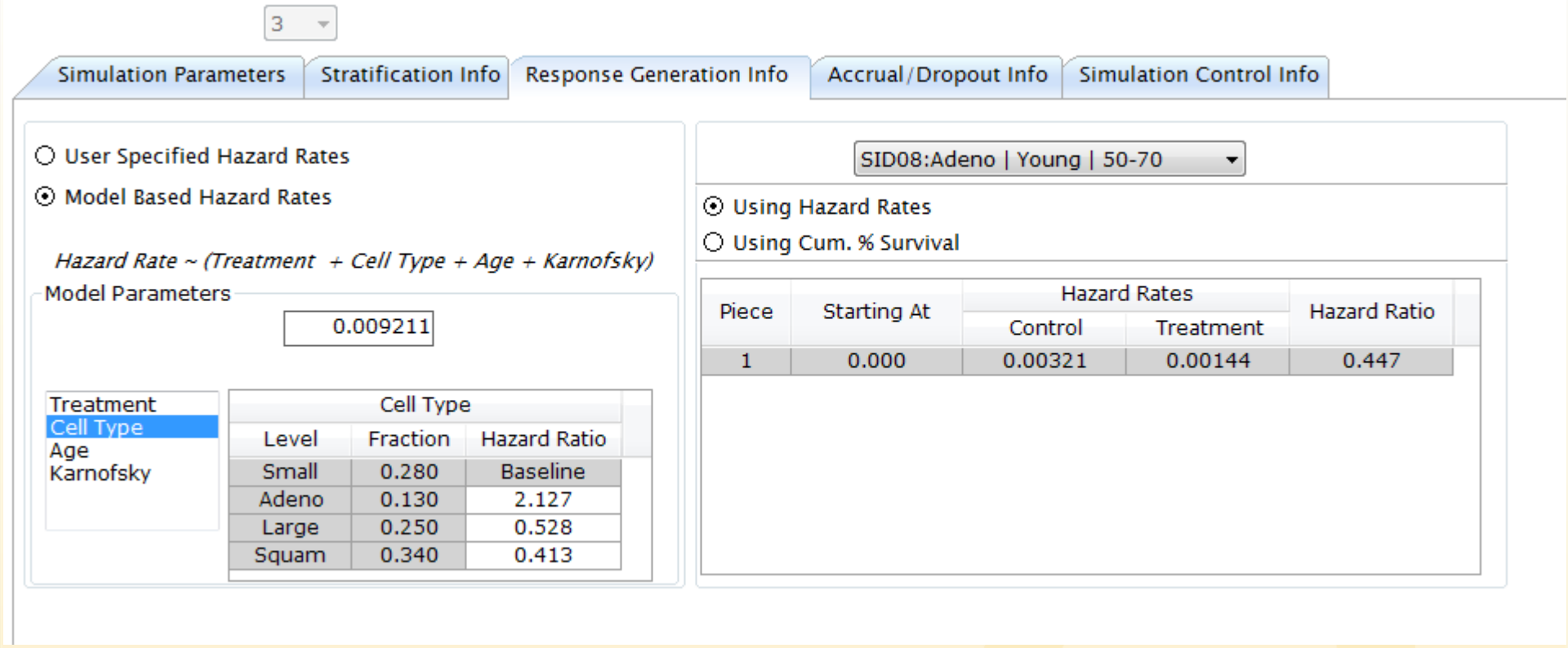
Cell type	Proportion	Hazard Ratio
Small cell	0.28	Baseline
Adenocarcinoma	0.13	2.127
Large cell	0.25	0.528
Squamous	0.34	0.413
Age Group	**Proportion**	**Hazard Ratio**
≤ 50	0.28	Baseline
> 50	0.72	0.438
Karnofsky P.S.	**Proportion**	**Hazard Ratio**
≤ 50	0.43	Baseline
50 to 70	0.37	0.164
>70	0.20	0.159

Entering the Stratification Information



- There are 24 strata: Cell(4) x Age(2) x Karnofsky(3)
- East provides two ways to enter the hazard rate for each stratum
  - enter individual hazard rates stratum by stratum
  - enter hazard rates implicitly through the Cox model

Entering Hazard Ratios through a Cox Model



Comparing Stratified and Unstratified Analyses

	Stratification:no strat:Sim1	Stratification:no strat:Sim2	Stratification:no strat:Sim3
Mnemonic	SU-2S-LRAR	SU-2S-LRAR	SU-2S-LRAR
Test Parameters			
Design Type	Superiority	Superiority	Superiority
Test Type	2-Sided	2-Sided	2-Sided
Test Statistic	Logrank	Stratified Logrank	Logrank
Power	0.887	0.849	0.771
No. of Looks	3	3	3
Model Parameters			
No. of Hazard Pieces	1		
Boundary Parameters			
Efficacy Boundary	User Specified	User Specified	User Specified
Spacing of Looks	User Specified	User Specified	User Specified
Accrual & Dropout Parameters			
Followup Duration	Until End of Study	Until End of Study	Until End of Study
Accrual Rate	12	12	12
No. of Accrual Periods	1	1	1
Sample Size			
Maximum	289	289	289
Events			
Maximum	66	66	66
Study Duration			
Maximum	43.9	171.115	177.394
Stratum Information			
No. of Stratum Variables		3	3
No. of Strata		24	24
Allocate Fractions to Strata		Marginally	Marginally
Specification of Hazard Rates		Model based	Model based

Example 4: Alzheimer's Disease Trial with Multiple Endpoints

- There are four efficacy variables measuring **change from baseline**
  - CS: Clinician's Score
  - ADAS-cog: Alzheimer's disease assessment scale of cognitive function
  - SIB: Severe Impairment Battery
  - MMSE: Mini-Mental State Examination
- Must win on **both** CS and ADAS-cog in order to claim efficacy
- SIB and MMSE are secondary endpoints. Can be included in claim if strong control of type-1 error

Serial Gatekeeping Strategy

First Family:

Test H_{CS} and $H_{ADAS\text{-}cog}$

Must reject both H_{CS} and $H_{ADAS\text{-}cog}$ to claim efficacy

Second Family:

Test H_{SIB} and H_{MMSE}

If either H_{SIB} or H_{MMSE} is rejected it can be included in claim

Require strong control of the FWER at one-sided level $\alpha = 0.025$

Means and Covariances from Historical Data

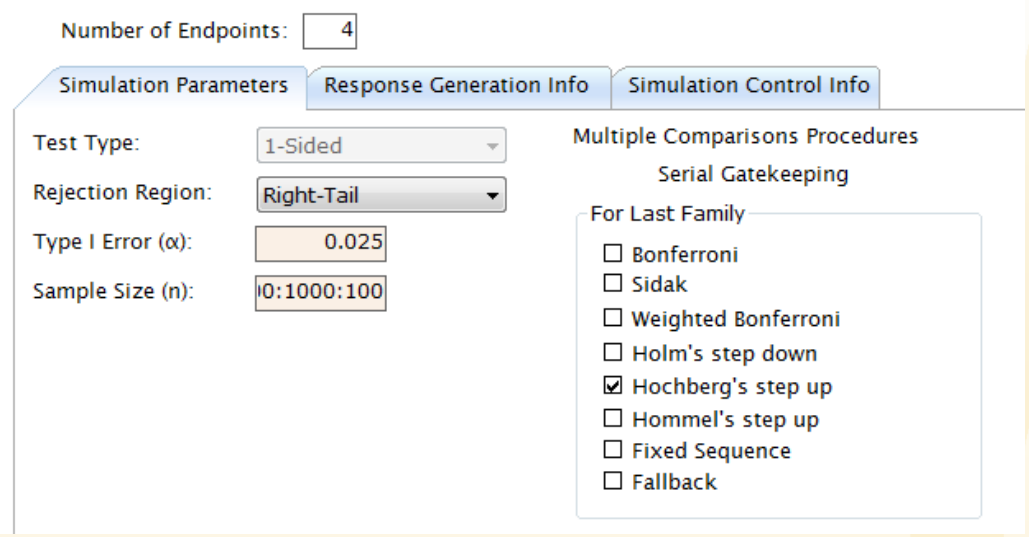
Week 28 change from baseline in control arm and targeted change in treatment arm

	Control Mean	Treatment Mean
CS	2.3	2.6
ADAS-cog	-4.5	-2.5
SIB	-10	-6.5
MMSE	-1.2	-0.4

Variance Covariance matrix for the four endpoints

CS	ADAS-cog	SIB	MMSE
1.2	3.6	6.8	1.6
3.6	42	38	9.3
6.8	38	145	17
1.6	9.3	17	8

Enter the design parameters in East



Number of Endpoints: 4

Simulation Parameters | Response Generation Info | Simulation Control Info

Endpoint Information:

EndPoint ID	Family Rank	Mean Ctrl	Mean Trmt
CS	1	2.3	2.6
ADAScog	1	-4.5	-2.5
SIB	2	-10	-6.5
MMSE	2	-1.2	-0.4

Covariance Matrix:

CS	ADAScog	SIB	MMSE
1.2	3.6	6.8	1.6
3.6	42	38	9.3
6.8	38	145	17
1.6	9.3	17	8

Simulation based Sample Size Assessment

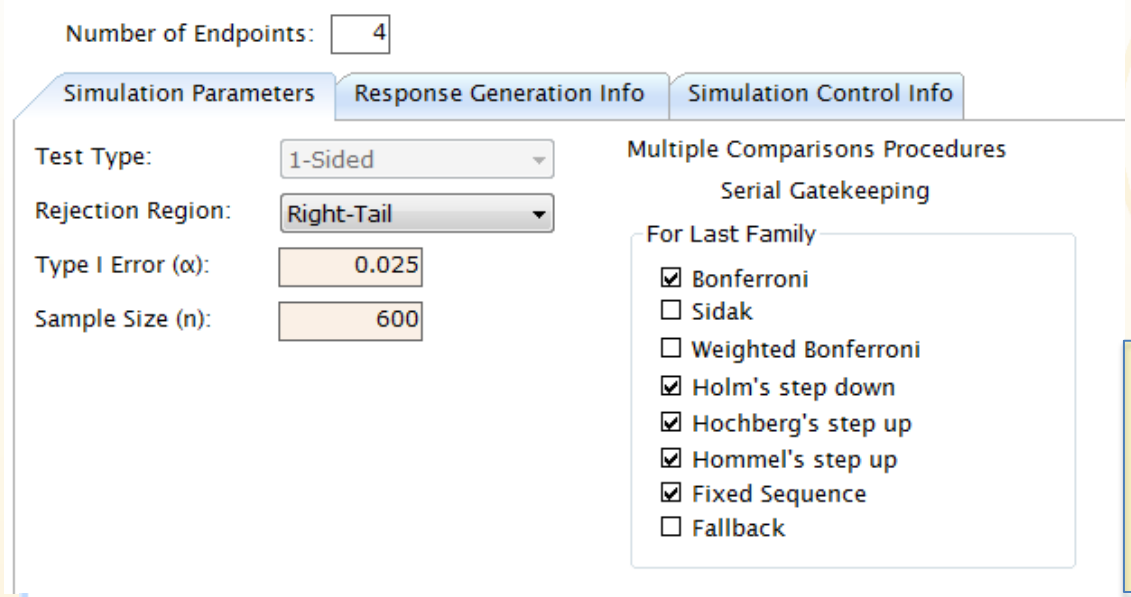
Output Summary

	Sim1	Sim2	Sim3	Sim4	Sim5	Sim6
Mnemonic	MN-2S-ME	MN-2S-ME	MN-2S-ME	MN-2S-ME	MN-2S-ME	MN-2S-ME
Test Parameters						
Multiple Comparisons Procedure	Hochberg's step up	Hochberg's step up	Hochberg's step up	Hochberg's step up	Hochberg's step up	Hochberg's step up
Test Type	1-Sided	1-Sided	1-Sided	1-Sided	1-Sided	1-Sided
Design Type	Serial	Serial	Serial	Serial	Serial	Serial
Specified α	0.025	0.025	0.025	0.025	0.025	0.025
No. of Endpoints	4	4	4	4	4	4
No. of Families	2	2	2	2	2	2
Rejection Region	Right-Tail	Right-Tail	Right-Tail	Right-Tail	Right-Tail	Right-Tail
MCP Results						
Conjunctive Power First Family	0.151	0.366	0.567	0.717	0.826	0.899
Conjunctive Power Last Family	0.066	0.218	0.411	0.588	0.726	0.83
Disjunctive Power Last Family	0.106	0.303	0.515	0.677	0.801	0.886
Overall FWER	0	0	0	0	0	0
Sample Size						
Maximum	100	200	300	400	500	600

- Conjunctive Power is power to reject every false null hypothesis
- Disjunctive Power is power to reject at least one false null hypothesis
- 600 patients are needed for 90% conjunctive power for first family

Next investigate by simulation: how many patients do we need for 90% conjunctive power for last family?

Maximize conjunctive power of last family by choice of test



For a 600 patient trial, Hommel's test has the most conjunctive power for last family, 0.835

	Sim11	Sim12	Sim13	Sim14	Sim15
Mnemonic	MN-2S-ME	MN-2S-ME	MN-2S-ME	MN-2S-ME	MN-2S-ME
Test Parameters					
Multiple Comparisons Procedure	Bonferroni	Holm's step down	Hochberg's step up	Hommel's step up	Fixed Sequence
Test Type	1-Sided	1-Sided	1-Sided	1-Sided	1-Sided
Design Type	Serial	Serial	Serial	Serial	Serial
Specified α	0.025	0.025	0.025	0.025	0.025
No. of Endpoints	4	4	4	4	4
No. of Families	2	2	2	2	2
Rejection Region	Right-Tail	Right-Tail	Right-Tail	Right-Tail	Right-Tail
MCP Results					
Conjunctive Power First Family	0.896	0.896	0.895	0.895	0.897
Conjunctive Power Last Family	0.78	0.83	0.827	0.835	0.829
Disjunctive Power Last Family	0.88	0.883	0.881	0.882	0.865
Overall FWER	0	0	0	0	0
Sample Size					
Maximum	600	600	600	600	600

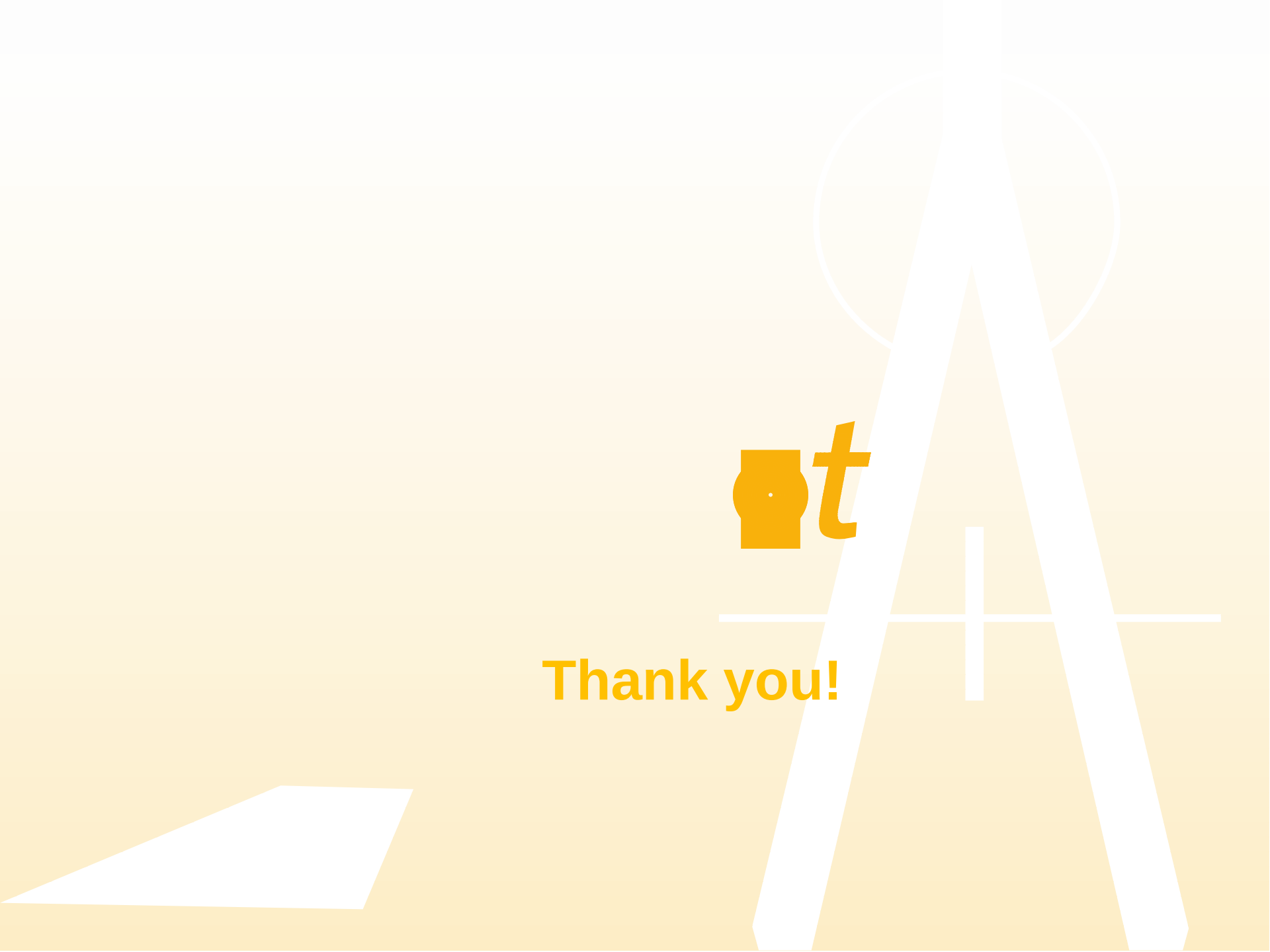
Increase Sample Size to Boost Hommel's Test to 90% Conjunctive Power for Second Family

	Sim18
Mnemonic	MN-2S-ME
Test Parameters	
Multiple Comparisons Procedure	Hommel's step up
Test Type	1-Sided
Design Type	Serial
Specified α	0.025
No. of Endpoints	4
No. of Families	2
Rejection Region	Right-Tail
MCP Results	
Conjunctive Power First Family	0.942
Conjunctive Power Last Family	0.902
Disjunctive Power Last Family	0.935
Overall FWER	0
Sample Size	
Maximum	700

- 700 patients needed to obtain 90% conjunctive power for last family
- At this sample size, conjunctive power of first family is 94%

Coming soon…in 6.3

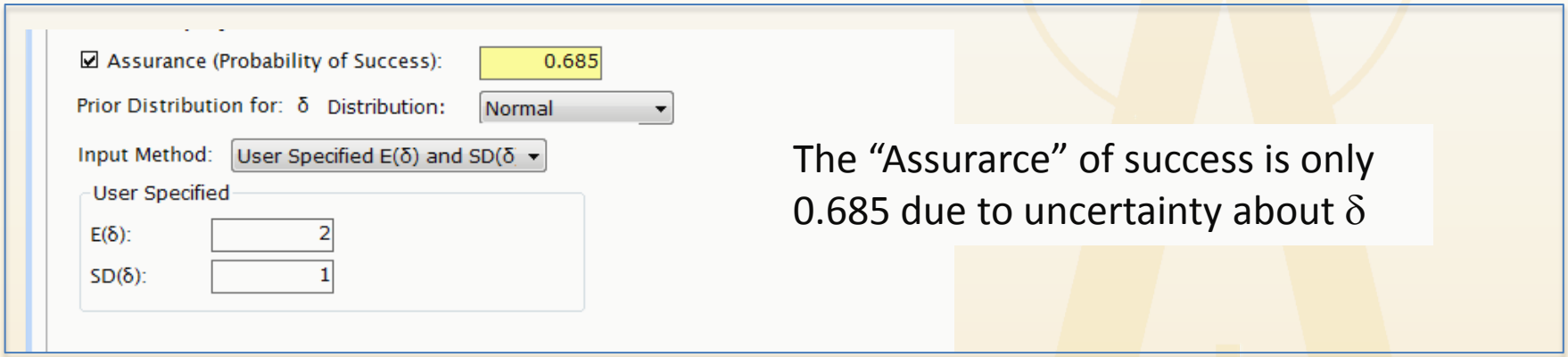
1. Adaptive dose selection with SSR
2. Interim simulations
3. More MCP options


Thank you!

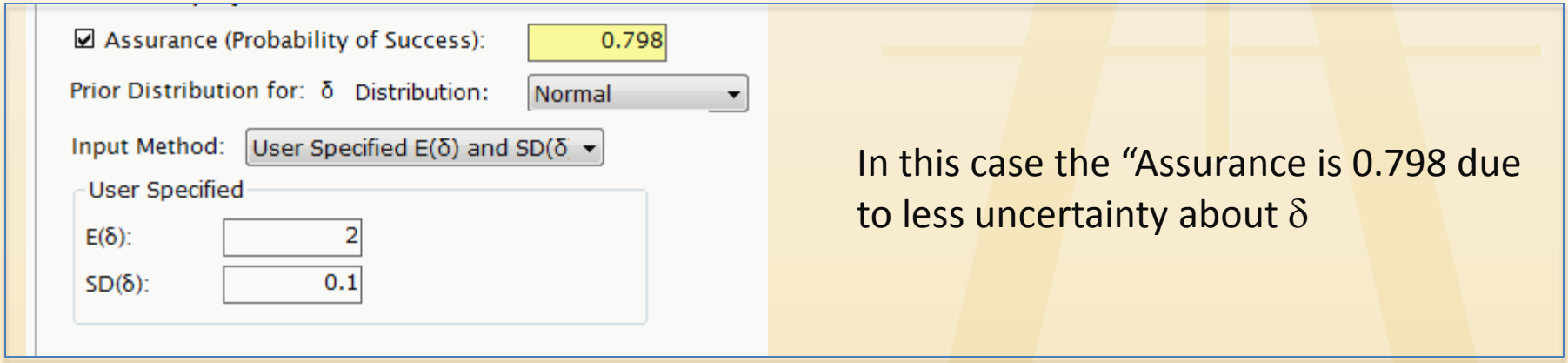
Backup slides

Bayesian Probability of Success (Assurance)

Assign a prior distribution to δ to capture the uncertainty associated with it



The "Assurarce" of success is only 0.685 due to uncertainty about δ



In this case the "Assurance is 0.798 due to less uncertainty about δ