



# A new approach to trial design and probabilistic risk assessment for trials with dual survival endpoints

Pantelis Vlachos  
Cytel, Inc



# Agenda

- Motivation
- Dual TTE Endpoints: Logistics
- Case Study
- Q&A

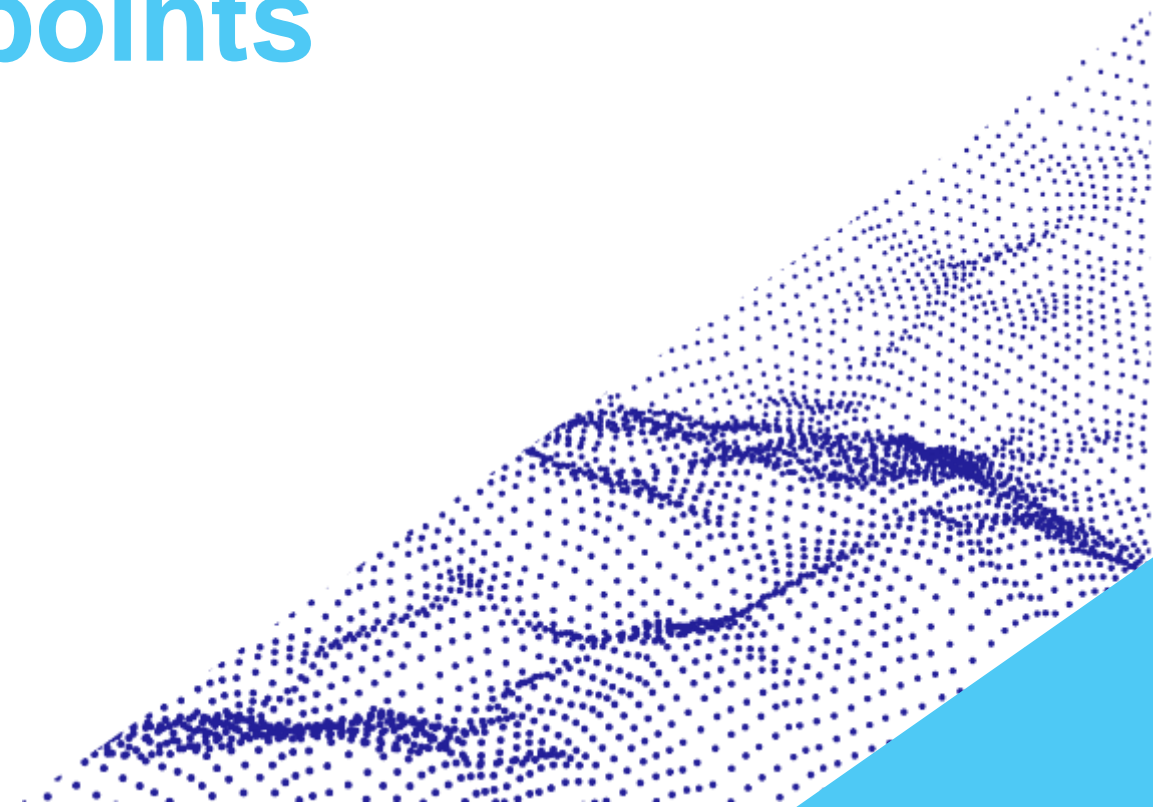


# Motivation

- Interim futility analysis with intermediate endpoints (Goldman, Leblanc and Crowley, 2008)
  - Futility testing commonly performed at very low levels (e.g., one-sided alpha 0.0025) at one or two times before final analysis.
- In TTE studies, problems might arise for adaptations, when using info on patients which are under risk at the interim. (Joergens, Wassmer et al, 2019)
  - Use Surrogate endpoints as basis for adaptations
- Systematic literature review to identify surrogate endpoints validated in oncology (Savina, PhD Thesis, 2018)



# Designing for Two Endpoints



# Key questions when designing trials with two endpoints



How do we define success?



How do we specify correlation, or other relationships between endpoints?



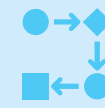
In the absence of early stopping, when do we end the trial?



How do we adjust for multiple testing of hypotheses?



How is the timing of interim analyses determined?



How do we define early stopping rules?

# Designing a study with two endpoints (at least one TTE)

## Primary + Secondary

## Primary + Primary

### Plans

**Plan 1**

Study Objective: Two Arm Confirmatory | Phase (Optional): 3

Target Population: All Comers | Control Arm: Standard of Care

	Priority	Endpoint Name	Endpoint Type
EP1	Primary	PFS	Time to Event
EP2	Secondary	OS	Time to Event

Winning Condition:  At least EP1,  Both endpoints

### Plans

**Plan 1**

Study Objective: Two Arm Confirmatory | Phase (Optional): 3

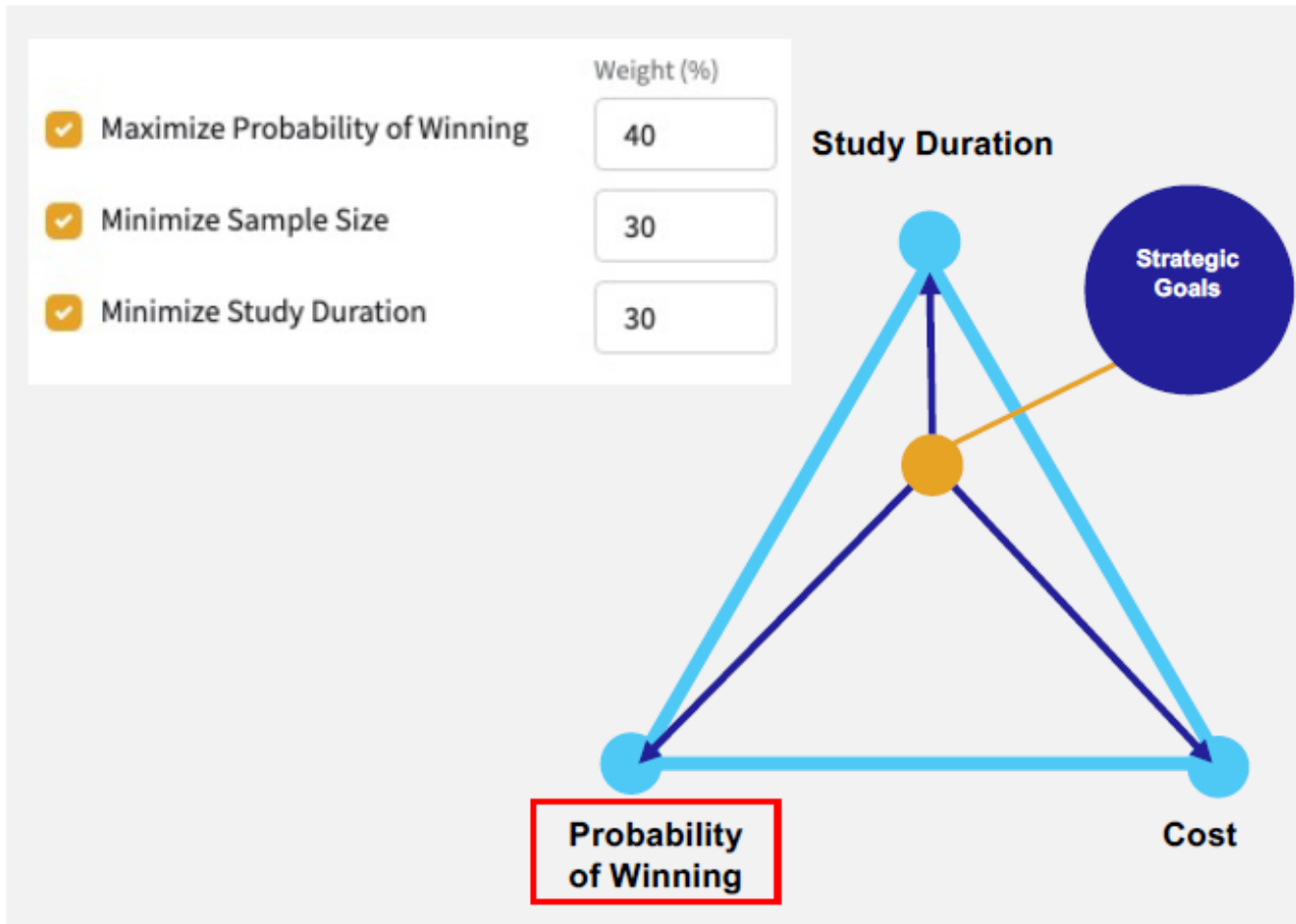
Target Population: All Comers | Control Arm: Standard of Care

	Priority	Endpoint Name	Endpoint Type
EP1	Primary	PFS	Time to Event
EP2	Primary	OS	Time to Event

Winning Condition:  At least EP1,  At least EP2,  At least one endpoint,  Both endpoints

Power is no longer sufficient to define success when considering two endpoints. Instead, Probability of Winning defines success based on the user-specified Winning Condition.

# Scoring system uses Probability of Winning rather than Power



Models can be scored on performance criteria that reflect strategic goals

The score is a weighted scaled function of performance criteria

- Probability of Winning
- Study Duration
- Study Cost (or Sample Size)

Selecting general design-agnostic criteria enable broad strategic comparisons

Scoring is meant to surface areas of interest in the design map that merit further exploration

# Generating Response Data

## Correlated Endpoints

**Response Set 1**

Plan: Plan 1

Target Population: All Comers  
Control Arm: Standard of Care  
Treatment Arm: Drug X

EP1 (Primary - Time to Event)  
EP2 (Secondary - Time to Event)  
Winning Condition

**General** | Dropout Rate

Input Method: Median Survival Time

Correlation: Uncorrelated

Endpoint Rules: None

EP1 Distribution: Exponential

Control (Month): 20

**POSITIVE CORRELATION**

- Very Weak Positive
- Weak Positive
- Moderate Positive
- Strong Positive
- Very Strong Positive

**NEGATIVE CORRELATION**

- Very Weak Negative
- Weak Negative
- Moderate Negative
- Strong Negative
- Very Strong Negative

## Generating Correlated Outcomes

Randomly generate two correlated random variables from a standard Bivariate Normal distribution with correlation qualitatively specified by the user:

- **0**: Uncorrelated
- **+/- 0.15**: Very Weak
- **+/- 0.3**: Weak
- **+/- 0.5**: Moderate
- **+/- 0.7**: Strong
- **+/- 0.85**: Very Strong

Using Cumulative Distribution Functions, transform – if necessary – the data to the desired distributions, exponential, piecewise exponential, binomial, etc.

Solara will report the actual observed correlation between the endpoints from simulation



# Generating Response Data

## Endpoint Rules

Response Set 1

Plan	Target Population	All Comers	EP1 (Primary - Time to Event)	PFS
Plan 1	Control Arm	Standard of Care	EP2 (Secondary - Time to Event)	OS
	Treatment Arm	Drug X	Winning Condition	Both endpoints

General Dropout Rate

Input Method: Median Survival Time

Correlation: Uncorrelated

Endpoint Rules: EP1 cannot occur after EP2

✓ If EP1>EP2, then EP1=EP2  
If EP1>EP2 then EP1 is censored

## Dictating Logical Relationships

Some endpoints have logical relationships that dictate the order in which they will be observed for a subject, eg. PFS cannot occur after OS.

Other pairs of endpoints do not display this kind of relationship.

Solara needs to know this to generate the data appropriately.

It also needs to know what to do when such a violation occurs in the data generation process. There are two options:

- Count an event: Set the value of one endpoint to be the value of the other, eg. if a value of PFS time was generated that was large than the value of OS time, then PFS time = OS time
- Count a censoring event: The one endpoint is censored by the other, eg. Time to Progression if longer than Survival time would be censored by the time of death

# Defining a Fixed Design

## Primary + Secondary

Design Set 1

Plan: Plan 1

Target Population: All Comers  
Control Arm: Standard of Care  
Treatment Arm: Drug X

EP1 (Primary - Time to Event): PFS  
EP2 (Secondary - Time to Event): OS  
Winning Condition: Both endpoints

Arms: 2

Statistical Design: Fixed Sample

Planned End of Trial: Full info for both endpoints

General

Sample Size: 450

Allocation Ratio: 1

1-Sided Type 1 Error: 0.025

Multiplicity Adjustment: Hierarchical

Testing Order: Start with EP1

EP1

Hypothesis: Superiority

Test Statistic: Logrank

Target Number of Events: 162

Critical Point: -1.959964

EP2

Hypothesis: Superiority

Test Statistic: Logrank

Target Number of Events: 331

Critical Point: -1.959964

### Planned End of Trial

- Full info for EP1 (Primary)
- Full info for both endpoints

### Multiplicity Adjustment options:

- Hierarchical:
  - Testing order is start with primary endpoint EP1

# Defining a Fixed Design

## Primary + Primary

**Design Set 2**

Plan: Plan 2

Target Population: All Comers

Control Arm: Standard of Care

Treatment Arm: Drug X

EP1 (Primary - Time to Event): PFS

EP2 (Primary - Time to Event): OS

Winning Condition: At least EP2

Arms: 2

Statistical Design: Fixed Sample

Planned End of Trial: Full Info for EP2

**General**

Sample Size: 182

Allocation Ratio: 1

1-Sided Type I Error: 0.025

Multiplicity Adjustment options:  
Hierarchical  
✓ Split  
None

**EP 1**

Hypothesis: Superiority

Test Statistic: Logrank

Target Number of Events: 88

Type-1 Error Allocation (%): 80

Type-1 Error Allocated: 0.020000

Critical Point: -2.053749

**EP 2**

Hypothesis:

Test Statistic:

Target Number of Events:

Type-1 Error Allocation (%):

### Planned End of Trial

- Full info for EP1
- Full info for EP2
- Full info for both endpoints

### Multiplicity Adjustment options:

- Hierarchical
  - Specify testing order
- Split : weighted Bonferroni
  - Specify allocation
- None

# Defining a Group Sequential Design

## Full info for both endpoints

Arms: 2 | Statistical Design: Group Sequential | Planned End of Trial: Full Info for both endpoints

General | **Early Stopping**

Synchronize Interims: Based on EP1

EP1				EP2			
Analysis	Analysis Spacing (%)	Efficacy	Futility	Analysis	Analysis Spacing (%)	Efficacy	Futility
IA1	50	<input checked="" type="checkbox"/>	<input type="checkbox"/>	IA1	Determined	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Final	100	<input checked="" type="checkbox"/>	<input type="checkbox"/>	IA2	Determined	<input checked="" type="checkbox"/>	<input type="checkbox"/>
				IA3	80	<input checked="" type="checkbox"/>	<input type="checkbox"/>
				Final	100	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

+ Add Interim to Both | + Add Interim to EP2

**EP1**

**EFFICACY**

Efficacy Boundary Family: Spending Functions | Spending Function: Lan-DeMets

**FUTILITY**

Futility Boundary Family: None

Synchronize interim analyses based on one of the endpoints

Information fraction for the other endpoint is determined during simulation with respect to the target number of events the user has specified

# Defining a Group Sequential Design

## Full info for driving endpoint

Arms: 2 | Statistical Design: Group Sequential | Planned End of Trial: Full info for EP1

General | **Early Stopping**

Synchronize Interims: Based on EP1  Info. fraction at interim analyses for EP2 same as EP1

EP1				EP2			
Analysis	Analysis Spacing (%)	Efficacy	Futility	Analysis	Analysis Spacing (%)	Efficacy	Futility
IA1	50	<input checked="" type="checkbox"/>	<input type="checkbox"/>	IA1	Determined	<input type="checkbox"/>	<input checked="" type="checkbox"/>
IA2	75	<input checked="" type="checkbox"/>	<input type="checkbox"/>	IA2	Determined	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Final	100	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Final	Determined	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

+ Add Interim to Both | + Add Interim to EP2

**EP1**

**EFFICACY**

Efficacy Boundary Family: Spending Functions | Spending Function: Lan-DeMets

Parameter: O'Brien-Fleming | Type 1 Error: 0.025

**FUTILITY**

Futility Boundary Family: None

Synchronize interim analyses based on one of the endpoints

Information fraction for the other endpoint is determined during simulation with respect to the target number of events the user has specified OR equated to the information fraction for the synchronizing endpoint

# Stopping Logic

- Win on Both Endpoints / Hierarchical Testing

EP1				EP2			
Analysis	Analysis Spacing (%)	Efficacy	Futility	Analysis	Analysis Spacing (%)	Efficacy	Futility
IA1	60	<input checked="" type="checkbox"/>	<input type="checkbox"/>	IA1	Determined	<input checked="" type="checkbox"/>	<input type="checkbox"/>
IA2	80	<input checked="" type="checkbox"/>	<input type="checkbox"/>	IA2	Determined	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Final	100	<input checked="" type="checkbox"/>	<input type="checkbox"/>	IA3	Determined	<input checked="" type="checkbox"/>	<input type="checkbox"/>
				IA4	85	<input checked="" type="checkbox"/>	<input type="checkbox"/>
				Final	100	<input checked="" type="checkbox"/>	<input type="checkbox"/>

IA	PFS	OS	Decision
IA1	✓ Eff	✓ Eff	Win
IA1	✓ Eff	✗ Eff	Continue
IA1	✗ Eff		Continue

IA	PFS	OS	Decision
IA1	✗ Eff		Continue
IA2	✓ Eff	✗ Eff	Continue
FA/IA3		✓ Eff	Win
FA/IA3		✗ Eff	Continue

IA	PFS	OS	Decision
IA1	✗ Eff		Continue
IA2	✗ Eff		Continue
FA/IA3	✓ Eff	✓ Eff	Win
FA/IA3		✗ Eff	Continue

IA	PFS	OS	Decision
IA1	✗ Eff		Continue
IA2	✗ Eff		Continue
FA/IA3	✗ Eff		Lose

# Case Study -

**A Randomized, Multi-Center, Double-blind, Placebo Controlled, Phase 3 Study to Investigate Safety and Efficacy of Treatment X in combination with Agent A compared with Placebo in combination with Agent B in Participants with Previously Untreated Locally Advanced, Unresectable or Metastatic PD-L1 Selected Non-Small Cell Lung Cancer (NSCLC)**

# Case Study 1 – ONCOLOGY/Lung

Parameter	Initial Inputs
Planned sample size	504
Number of events (if applicable)	328 for OS and 278 for PFS
Treatment/control effect	HR = 0.58 for PFS and HR=0.73 for OS
Standard deviation (if applicable)	
Follow-up time (if applicable)	
Allocation ratio	1:1
Type-1 error (1-sided)	0.1% for PFS and 2.4% for OS
Target average power	PFS: 92.7%, OS: 80%
Number of interim analyses (if applicable)	1 IA for PFS and 3 IA's for OS
Timing of interim analyses (if applicable)	PFS: 75% information OS: 32%, 50%, 75% information
Alpha spending function (if applicable)	Lan DeMets OBF boundaries
Promising zone minimum/maximum (if applicable)	
Target conditional power (if applicable)	
Beta spending function (if applicable)	

## Exploration Goals

### Primary Outcome –

Overall Survival

Progressions –free survival

### Optimization Aim:

Other multiplicity approaches, Varying HR, varying information fraction, different data maturity for PFS/Final analysis, Probability of observing median in the active and control arms, critical values for HR (0.7 for PFS and 0.8 for OS)

### Additional Information:

Time for primary analysis - recruitment +median of control arm

Accrual: 9 pts/m for first 6 m, 25 pts/m for 6-12 and 38 pts/m thereafter

mPFS = 6.9 m (curve plateaus – **piecewise exponential** with 50% at 6.9 m and 30% at 18 m), mOS = 22.2 (exponential curve)



# Case Study 1 – ONCOLOGY/Lung – Simulation Plan

Parameter	Initial Inputs
Planned sample size	450, 504, 550
Number of events (OS)	255, 328, 390
Number of events (PFS)	230, 278, 310
Treatment/control effect (OS)	0.7, 0.73, 0.75
Treatment/control effect (PFS)	0.55, 0.58, 0.6
mPFS	6.9 months
mOS	22.2 months
Enrollment	9 pts/m for first 6 m, 25 pts/m for 6-12 and 38 pts/m thereafter
Allocation ratio	1:1
Type-1 error (1-sided)	0.1% for PFS and 2.4% for OS
Target average power	PFS: 92.7%, OS: 80%
Number of interim analyses	1 IA for PFS and 3 IA's for OS
Timing of interim analyses (PFS)	60%, 70%, 75%
Timing of interim analyses (OS)	(* , * , 75%), (* , * , 60%) *:determined
Alpha spending function (if applicable)	Gamma (-2, -3, -4)
Beta spending function (if applicable)	Gamma (-40, -4)
Total Models	48,440

## Exploration Goals

### Primary Outcome –

Overall Survival

Progressions –free survival

### Optimization Aim:

Other multiplicity approaches, Varying HR, varying information fraction, different data maturity for PFS/Final analysis, Probability of observing median in the active and control arms, critical values for HR (0.7 for PFS and 0.8 for OS)

### Additional Information:

Time for primary analysis - recruitment +median of control arm

Accrual: 9 pts/m for first 6 m, 25 pts/m for 6-12 and 38 pts/m thereafter

mPFS = 6.9 m (curve plateaus – **piecewise exponential** with 50% at 6.9 m and 30% at 18 m), mOS = 22.2 (exponential curve)

# Exploring the Output

# 48440 Models = 865 Designs x 56 Scenarios

PFS ev: 310  
OS ev: 390

Gain 15%  
Gain 15%

Favorite

Designs

Output

0.8

0.6

0.4

0.2

Sample Size  
Number of Events (EP2)  
Number of Events (EP1)  
Futility Parameters (EP1)  
Efficacy Parameters (EP1)  
Interim Analysis Spacing (EP2)  
Interim Analysis Spacing (EP1)  
Robustness Score

Favorite Scenario

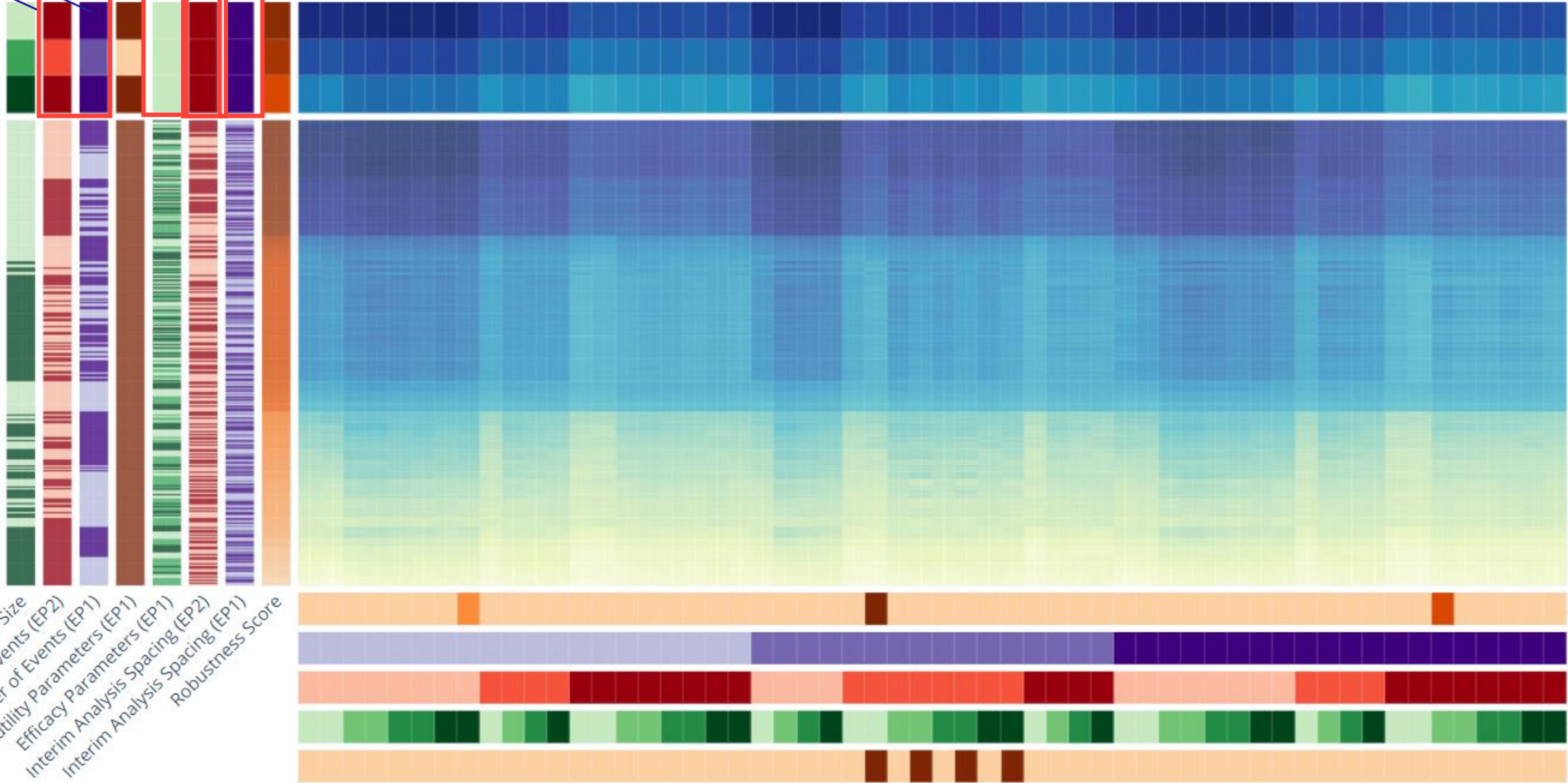
HR (EP1)

HR (EP2)

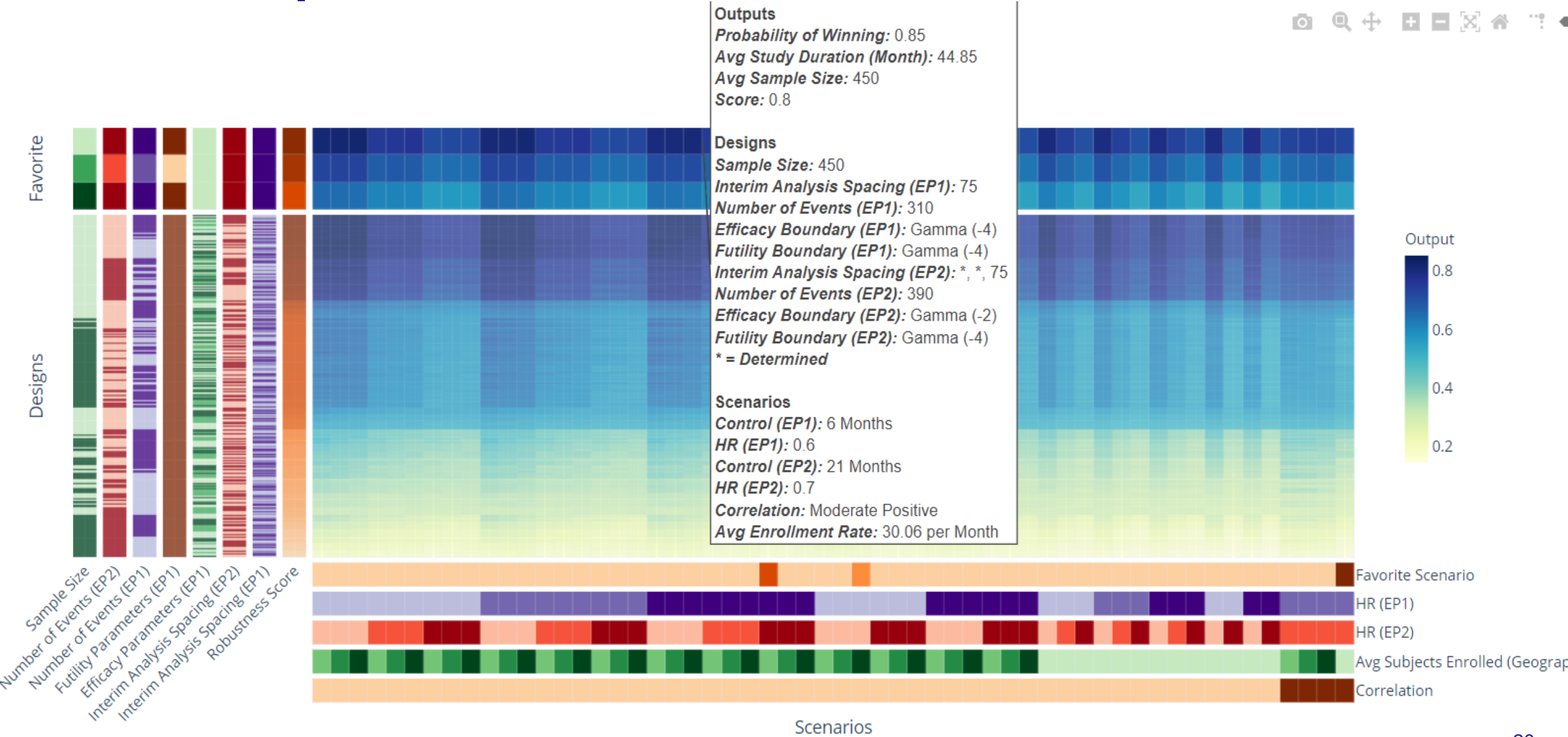
Avg Subjects Enrolled

Correlation

Scenarios



# Quick output



# Tabular summary- Favorite Designs

Designs	AtLeast80PowerShort (Target Number of Events (EP1)=310;Target Number of Events (EP2)=390;Sample Size=550;Multiplicity Adjustment=Split;Analysis Spacing (%) Info (EP1)=[75];Analysis Spacing (%) Info (EP2)=[ "determined","determined",75];Efficacy Parameter (EP1)=-4;Futility Parameter (EP1)=-4)			AtLeast80PowerSmallN (Target Number of Events (EP1)=310;Target Number of Events (EP2)=390;Sample Size=450;Multiplicity Adjustment=Split;Analysis Spacing (%) Info (EP1)=[75];Analysis Spacing (%) Info (EP2)=[ "determined","determined",75];Efficacy Parameter (EP1)=-4;Futility Parameter (EP1)=-4)			Reference Design (Target Number of Events (EP1)=278;Target Number of Events (EP2)=328;Sample Size=504;Multiplicity Adjustment=Split;Analysis Spacing (%) Info (EP1)=[75];Analysis Spacing (%) Info (EP2)=[ "determined","determined",75];Efficacy Parameter (EP1)=-4;Futility Parameter (EP1)=N/A)		
Scenarios	Optimistic Scenario (HR (EP1)=0.55;HR (EP2)=0.7;Avg Subjects Enrolled=30)	Pessimistic Scenario (HR (EP1)=0.6;HR (EP2)=0.75;Avg Subjects Enrolled=22)	Reference Scenario (HR (EP1)=0.58;HR (EP2)=0.73;Avg Subjects Enrolled=9,25,38)	Optimistic Scenario (HR (EP1)=0.55;HR (EP2)=0.7;Avg Subjects Enrolled=30)	Pessimistic Scenario (HR (EP1)=0.6;HR (EP2)=0.75;Avg Subjects Enrolled=22)	Reference Scenario (HR (EP1)=0.58;HR (EP2)=0.73;Avg Subjects Enrolled=9,25,38)	Optimistic Scenario (HR (EP1)=0.55;HR (EP2)=0.7;Avg Subjects Enrolled=30)	Pessimistic Scenario (HR (EP1)=0.6;HR (EP2)=0.75;Avg Subjects Enrolled=22)	Reference Scenario (HR (EP1)=0.58;HR (EP2)=0.73;Avg Subjects Enrolled=9,25,38)
Outputs									
Average Probability Of Winning	0.933	0.744	0.822	0.909	0.716	0.8	0.863	0.679	0.741
Average Sample Size (Overall)	546.359	535.415	549.578	450	448.034	450	503.066	497.349	503.968
Average Number of Events (EP1) (Overall)	241.547	254.483	249.093	241.47	253.79	248.477	219.626	233.564	227.975
Average Number of Events (EP2) (Overall)	274.517	286.12	283.238	276.462	287.754	285.52	257.078	261.488	262.569
Average Study Duration (Month)	37.638	40.937	42.763	49.452	52.546	55.64	37.102	39.093	41.889
Average Accrual Duration (Month)	18.178	24.291	21.07	14.968	20.318	18.448	16.736	22.555	19.868



# Optimizing Further

Filters
Test Scenarios

New Filter Set ▾ Save As

Add Filter... ▾

**PROBABILITY OF WINNING (%)** 🗑️

Reference Scenario ▾

○

**PROBABILITY OF WINNING (%)** 🗑️

Pessimistic Scenario ▾

○○

**TEAM PRIORITIES (%)**

Power

Sample Size




Duration

## 69 Results of Reference Scenario



Sort by: Avg. Duration (Shortest) ↓↑

	Avg. Sample Size	Probability of Winning	Avg. Duration (Months)	
<input checked="" type="checkbox"/>	550 (547 - 550)	82.2%	42.8 (21.3 - 59.4)	<span style="color: blue;">♥</span> <span style="margin-left: 10px;">☰</span>
<input type="checkbox"/>	550 (547 - 550)	82.1%	42.8 (21.3 - 59.4)	
<input type="checkbox"/>	550 (547 - 550)	82.1%	42.8 (21.3 - 59.4)	
<input type="checkbox"/>	549 (531 - 550)	82.3%	43.3 (20.6 - 59.4)	
<input type="checkbox"/>	549 (531 - 550)	81.9%	43.3 (20.6 - 59.4)	
<input type="checkbox"/>	549 (531 - 550)	82.1%	43.3 (20.6 - 59.4)	
<input type="checkbox"/>	547 (482 - 550)	82.1%	43.3 (19.3 - 59.4)	

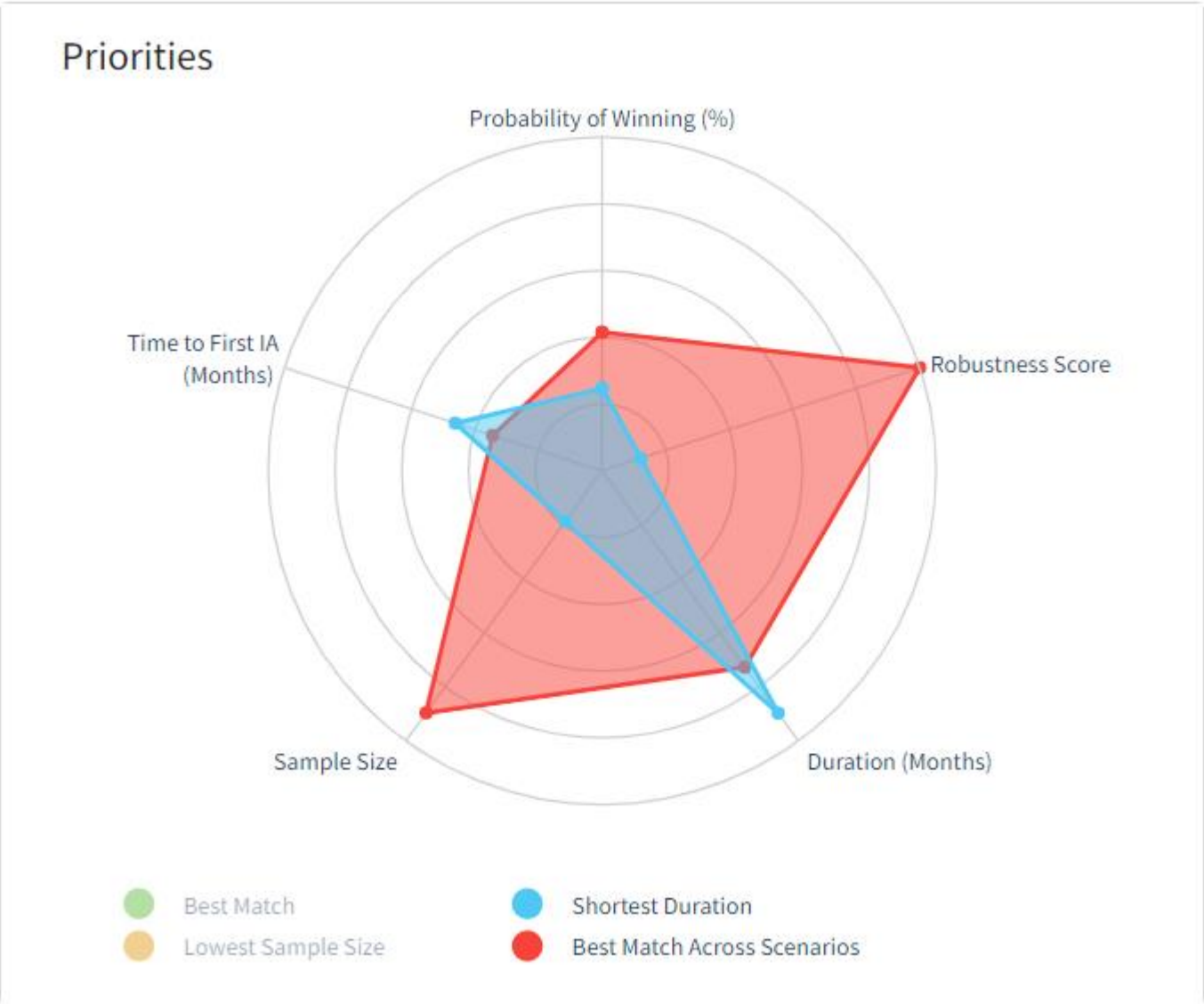
# Under Pessimistic Scenario

Filters		Test Scenarios	
 Reference Scenario			
 Optimistic Scenario			
 Pessimistic Scenario			

69 Results of Pessimistic Scenario			Sort by: Avg. Duration (Shortest) ↓↑	
<input checked="" type="checkbox"/>	Avg. Sample Size 536 (438 - 550)	Probability of Winning 74.4%	Avg. Duration (Months) 40.9 (19.9 - 56.4)	 
<input type="checkbox"/>	Avg. Sample Size 535 (438 - 550)	Probability of Winning 74.3%	Avg. Duration (Months) 40.9 (19.9 - 56.4)	
<input type="checkbox"/>	Avg. Sample Size 536 (438 - 550)	Probability of Winning 74.2%	Avg. Duration (Months) 40.9 (19.9 - 56.4)	
<input type="checkbox"/>	Avg. Sample Size 535 (419 - 550)	Probability of Winning 74.2%	Avg. Duration (Months) 41 (19 - 56.5)	
<input type="checkbox"/>	Avg. Sample Size 535 (419 - 550)	Probability of Winning 74.3%	Avg. Duration (Months) 41 (19 - 56.4)	
<input type="checkbox"/>	Avg. Sample Size 536 (419 - 550)	Probability of Winning 74.5%	Avg. Duration (Months) 41.1 (19 - 56.5)	
<input type="checkbox"/>	Avg. Sample Size 538 (380 - 550)	Probability of Winning 74%	Avg. Duration (Months) 41.2 (17.2 - 56.4)	

# Graphical Summaries: Radar Plot



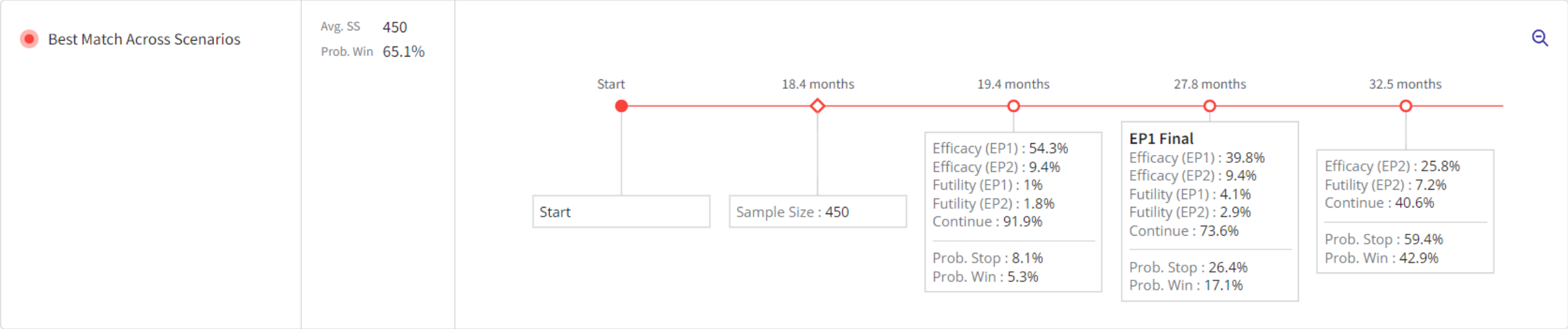


# Timelines Comparison

● Start/End ○ Interim ◇ Initial Enroll End



# Detailed Timelines per Design

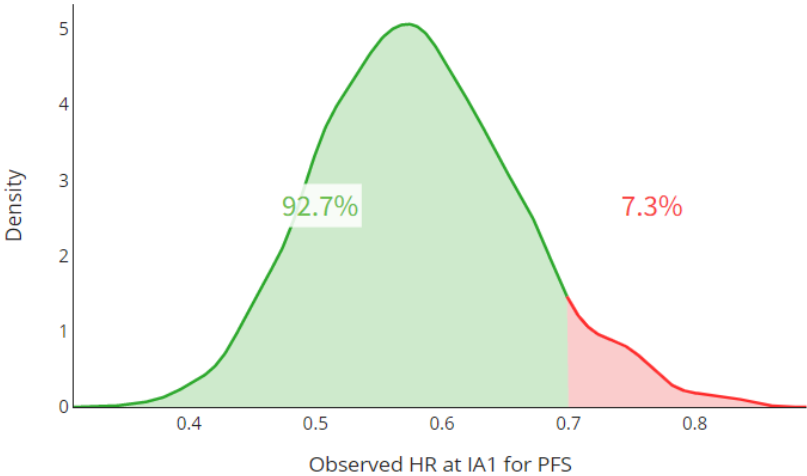


# Calculate chance of winning at different looks and different endpoints

Group Sequential ♥ AtLeast80PowerSmallIN 🚩 Reference Scenario

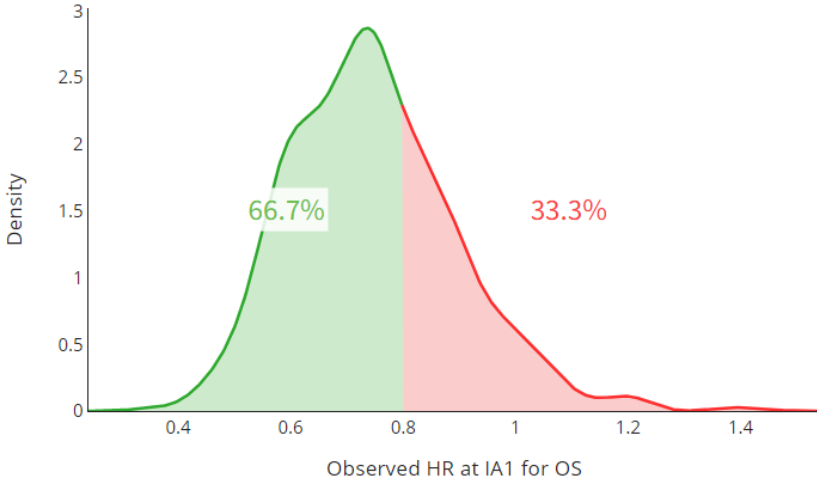
Score **0.727** Robustness Score **0.742** Weighted Probability of Winning **80.6%** Likelihood **0.167** Synchronize Interims **Based on EP1** EP1 (PFS): Superiority 2-Arms, Until End of Study EP2 (OS): Superiority 2-Arms, Until End of Study

Percentage of Observed HR <= 0.7 at IA1 for PFS



Save Thresholds

Percentage of Observed HR <= 0.8 at IA1 for OS



# Detailed look for each design

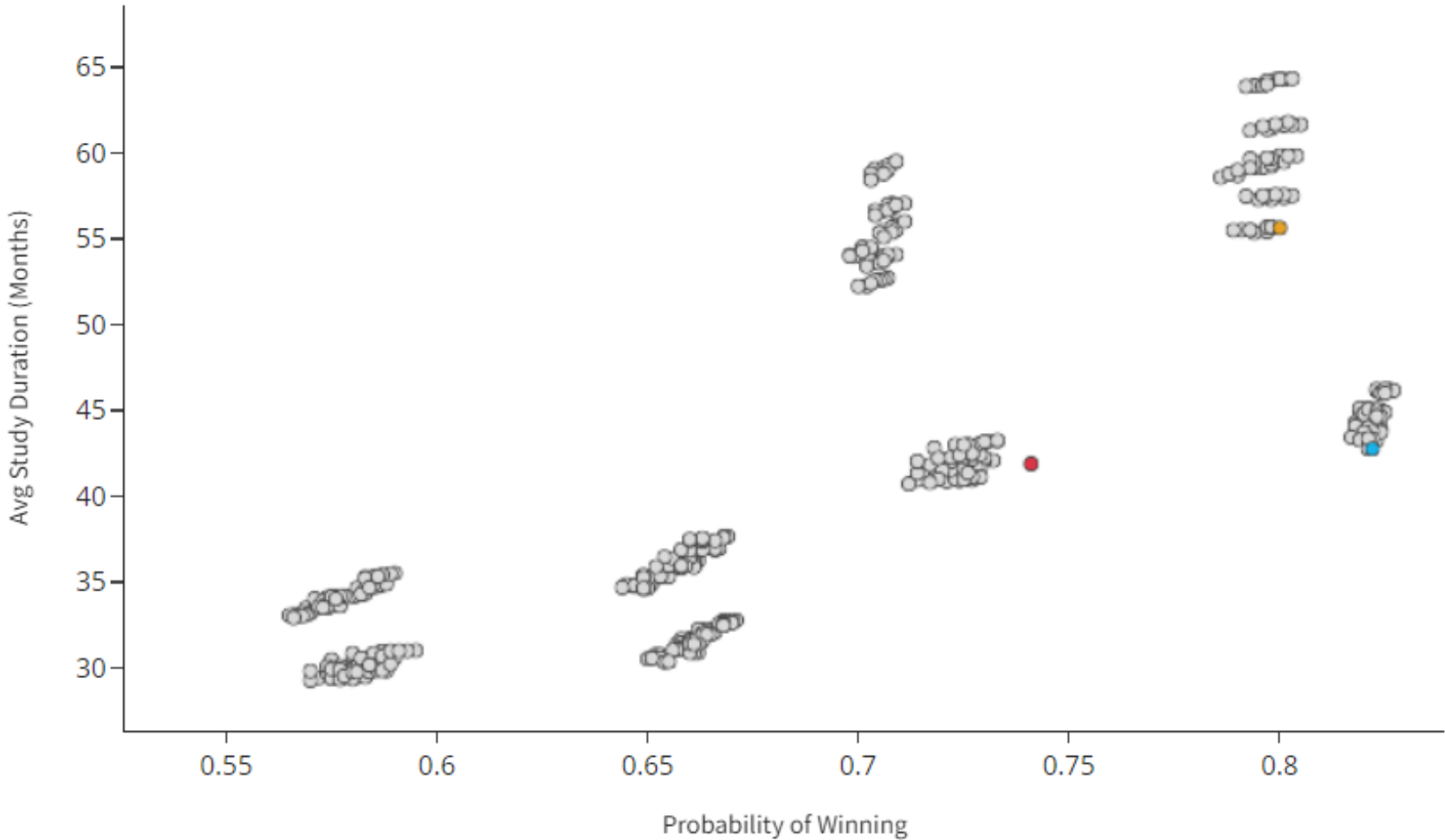
## Simulation Boundaries and Incremental Boundary Crossing Probabilities



Analysis #	Events for EP1	Events for EP2	EP1 Boundary Crossing		EP2 Boundary Crossing		Decisions		Total Simulations	
			Crossing For		Crossing For		Stopping Trial For			
			Efficacy	Futility	Efficacy	Futility	Efficacy (win)	Futility	Count	%
1	233	105.830	781	17	113	8	88	25	113	11.300
2	310	155.676	168	33	32	3	46	36	82	8.200
3	0	293			501	40	501	40	541	54.100
4	0	390			165	99	165	99	264	26.400
Total			949	50	811	150	800	200	1000	100
%			94.900	5	81.100	15	80	20		

# Duration vs Power Plot

Avg Study Duration vs Probability of Winning



# Changing Objectives

## Reference Response scenario

Design Name	Planned Sample Size	Planned Number of events	Avg. Sample Size	Avg. Events	Power	Avg. Study Duration (Months)	Marginal power (PFS)	Marginal Power (OS)	Alpha Spending Function	Beta Spending Function	Interim Analyses (*: Determined)
<b>Reference Design</b>	504	PFS: 278 OS: 328	504	PFS: 275 OS: 252	74.4%	40	89.9%	74.8%	PFS: Gamma (-4) OS: Gamma (-4)	None	PFS: 70% OS: (28, 40, 75)%
<b>Power Optimized Design</b>	504	PFS: 310 OS: 390	504	PFS: 305 OS: 304	83.8%	52	94.5%	83.9%	PFS: Gamma (-3) OS: Gamma (-4)	None	PFS: 75% OS: (27, 38, 60)%
<b>Sample Size Optimized Design</b>	450	PFS: 278 OS: 390	450	PFS: 273 OS: 288	80.8%	57	92%	80.9%	PFS: Gamma (-3) OS: Gamma (-3)	PFS: Gamma (-4) OS: Gamma (-40)	PFS: 60% OS: (18, 34, 75)%
<b>Duration Optimized Design</b>	550	PFS: 310 OS: 390	550	PFS: 302 OS: 282	80.9%	43	94.7%	81.3%	PFS: Gamma (-4) OS: Gamma (-2)	PFS: Gamma (-4) OS: Gamma (-4)	PFS: 75% OS: (26, 37, 75)%
<b>Balanced Design</b>	450	PFS: 310 OS: 390	450	PFS: 303 OS: 273	79.7%	54	93.6%	80.1%	PFS: Gamma (-3) OS: Gamma (-2)	PFS: Gamma (-4) OS: Gamma (-4)	PFS: 60% OS: (21, 40, 75)%



Thank you



**Pantelis Vlachos**

[Pantelis.Vlachos@cytel.com](mailto:Pantelis.Vlachos@cytel.com)

VP Customer Success

Cytel Inc. | Geneva, CH

<https://calendly.com/pantelis-vlachos/30min>

# Cytel Events for **Statistical Innovators**

## **Spring 2023**

Hear from your peers & industry experts  
and take part in interactive workshops

**REGISTER NOW**

[www.cytel.com/software-event-booking](http://www.cytel.com/software-event-booking)



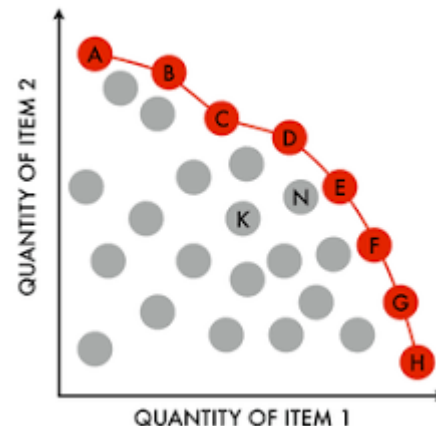
**Cytel**



# Appendix

# Items to Note

- Alpha-spending Functions: O'Brien-Fleming (OF) & Gamma
  - Gamma= -4 approximates OF
  - Gamma= -1 approximates Pocock
- Target Value needs to be specified for futility interim analyses
  - It's the value of the Alternative Hypothesis that is intended to be rejected in favor of the Null Hypothesis
- The Pareto Set of simulated designs is identified by Solara
  - no individual score criterion (e.g., power, sample size, and duration) can be better off without making at least one other criterion worse off or without any loss thereof.



# Target Value of HR for Futility IA

- In order to compute the futility boundary using the beta-spending function we have to solve the equation below:

Keeping this value of  $\eta$  and the previously obtained efficacy boundary values  $\{u_1, u_2, \dots, u_K\}$  fixed, compute the futility boundary  $\{l_1, l_2, \dots, l_K\}$  as follows:

$$P_\eta(W(t_1 \leq l_1) = \beta(t_1)) \quad (\text{B.68})$$

This means we need to know how much type-2 error to spend. However, we never specify power in Solara nor the alternative hypothesis (that give us eta), so how do we do this?

We know what the number of events is and that determines max information under proportional hazard. We could get  $\eta_1$  if only we knew  $\delta_1$ , the target HR:

$$I_{\max} = \left[ \frac{\eta_1}{\delta_1 - \delta_0} \right]^2$$

Once we solve for  $\eta$  we can build the boundary, because we also know what the power is for  $D_{\max}$  events when trying to detect a difference of  $\delta_1$ .

The target value of HR for futility is  $\delta_1$