STATISTICAL ANALYSIS PLAN

HYPERBARIC OXYGEN BRAIN INJURY TREATMENT TRIAL (HOBIT)

Byron J Gajewski, PhD Department of Biostatistics University of Kansas Medical Center	Renee H. Martin, PhD Data Coordination Unit Department of Public Health Sciences Medical University of South Carolina	Caitlyn Meinzer, PhD Data Coordination Unit Department of Public Health Sciences Medical University of South Carolina	Scott Berry, PhD Berry Consultants

General Goals of Phase II, & III Trials

- **Phase II (e.g. HOBIT)**
  - Tests safe dose(s) of a drug or treatment on a larger group of patients
  - Focus us efficacy
  - Can be one dose or multiple doses
  - Larger in size e.g. 50-200 patients
- **Phase III**
  - Confirmatory trial using dose identified vs standard of care
  - Usually 2-arms
  - Large in size e.g. 500-1000 patients

A “best” clinical trial design would be:

- Smaller: The number of patients used in the trial
- Stronger: Power of the trial
- Faster: Finish the trial fast
- Benefit more patients: While being smaller, stronger, and faster we would like the trial to somehow put more patients on better in-trial therapies
- Setup for phase III

How do we go towards smaller, stronger, faster, and more beneficial clinical trials?

- One promising solution is: Bayesian Adaptive Designs

"Given the number of times in which an unknown event has happened and failed [... Find] the chance that the probability of its happening in a single trial lies somewhere between any two degrees of probability that can be named."

Reverend Thomas Bayes (1702-1761)

Fixed Clinical Trial Designs

- Most popular trials
- Simpler to implement, statistically sound
- Fixed sample size and randomization set, using the maximum subjects allowed

Adaptive Clinical Trial Designs

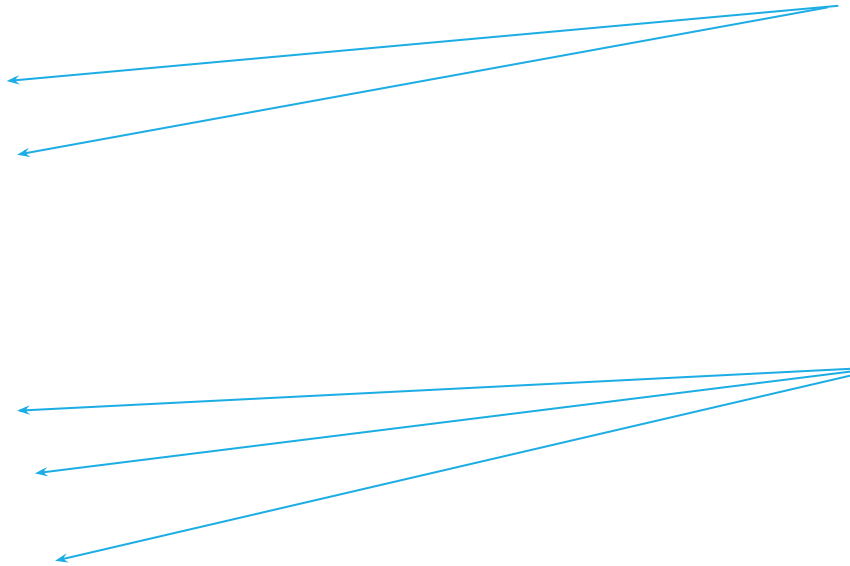
- Observations during the trial allow updates
  - Response Adaptive Randomization
  - Stopping early (interim analyses)
  - Dropping/adding doses or subgroups
  - Sample size
- More patients on better treatment
- More complex to implement, BUT also are statistically sound
- Bayesian
  - Allows prior information
  - Posterior probability interpretation

What is Response Adaptive Randomization (RAR)?

- As we get patient responses relative to the dose assigned, change the randomization plan to be proportional to the better treatment
- Why?
  - Better power for 3 treatment doses or more
  - More trial patients get better dose

How Does RAR work?

- Scenario 1:
  - Control: 40%
  - Dose 1: 40%
  - Dose 2: 45%
  - Dose 3: 50%
- Scenario 2:
  - Control: 40%
  - Dose 1: 40%
  - Dose 2: 40%
  - Dose 3: 40%



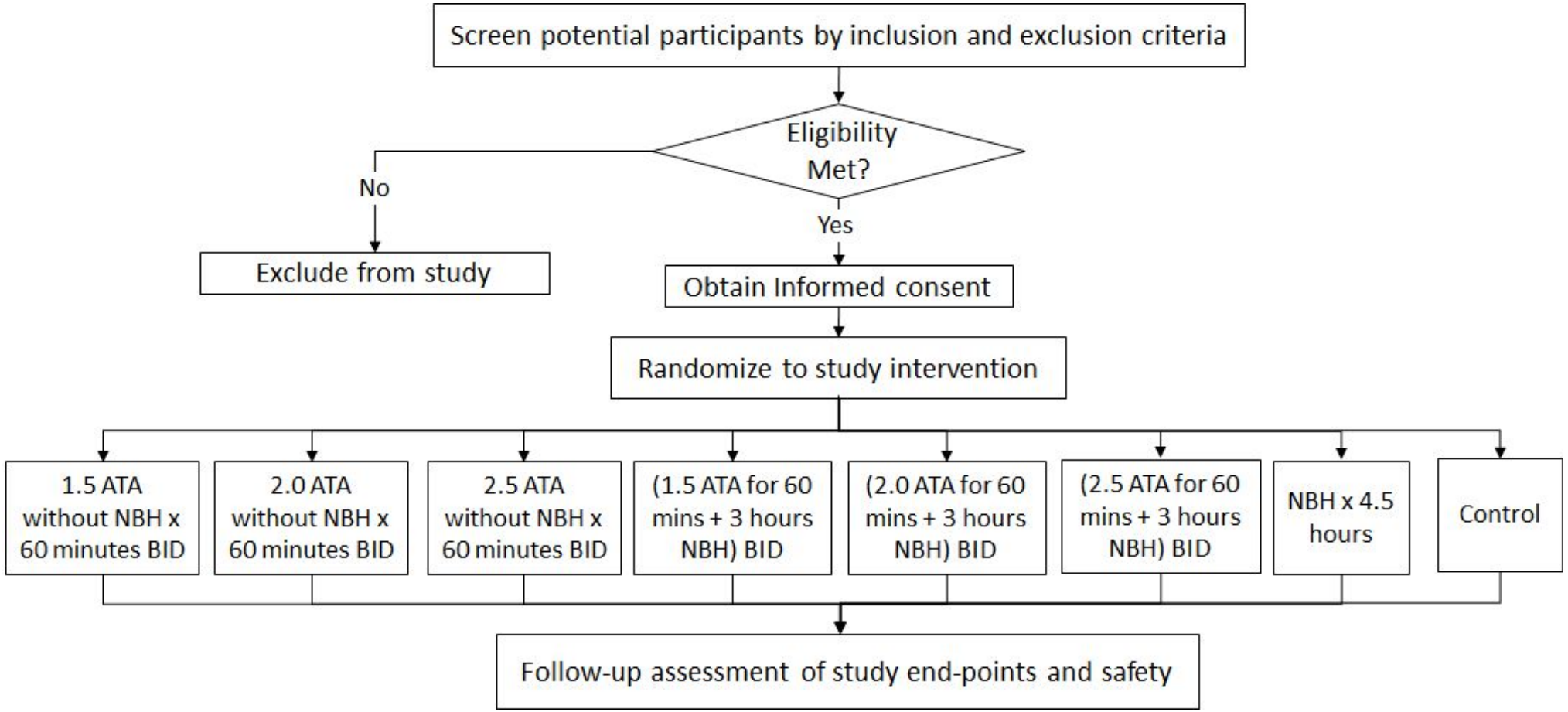
As we learn, more pts placed here! Takes less resources distinguish doses 1 and 3 than doses 2 and 3, lets use our resources right!

Allocate pts fairly equally, behaves like a fixed trial!

Aims of HOBIT

1. (Dose selection) The first aim is to select the combination of treatment parameters (pressure and intervening NBH) that is most likely to demonstrate improvement in the rate of good neurological outcome at 6 months following severe TBI injury versus standard-of-care therapy in a subsequent confirmatory trial.

2. (Signal of efficacy) The second aim is to determine whether there is a >50% probability of *hyperoxia* treatment demonstrating improvement in the rate of good neurological outcome at 6 months following severe TBI injury versus standard-of-care therapy in a subsequent confirmatory trial.


Screen potential participants by inclusion and exclusion criteria
Eligibility Met?
No
Yes
Exclude from study
Obtain Informed consent
Randomize to study intervention
1.5 ATA without NBH x 60 minutes BID
2.0 ATA without NBH x 60 minutes BID
2.5 ATA without NBH x 60 minutes BID
(1.5 ATA for 60 mins + 3 hours NBH) BID
(2.0 ATA for 60 mins + 3 hours NBH) BID
(2.5 ATA for 60 mins + 3 hours NBH) BID
NBH x 4.5 hours
Control
Follow-up assessment of study end-points and safety

Endpoint @ 26 weeks

Probability of poor Outcome on IMPACT	Glasgow Outcome Scale-Extended						
	Upper Good Recovery	Lower Good Recovery	Upper Moderate Disability	Lower Moderate Disability	Upper Severe Disability	Lower Severe Disability	Vegetative or Death
GOS -E	8	7	6	5	4	3	2/1
0 to <0.21							
0.21 to <0.41					Poor Outcome		
0.41 to <0.56		Favorable Outcome					
0.56 to ≤1.0							

Longitudinal Data

Primary Endpoint

Patient	#~1 month (30 days +/- 7)	~3 months (90 days +/- 14)	~6 months (180 days +/- 30 days)
	Sliding GOS-E (favorable or Poor outcome)	Sliding GOS-E (favorable or Poor outcome)	Sliding GOS-E (favorable or Poor outcome)
1			
2			
…			
52			Predict this until data collected
53			Predict this until data collected

* IMPACT Score (Probability of Poor Outcome) is calculated at baseline

Longitudinal Data

Primary Endpoint

Patient	#~1 month (30 days +/- 7)	~3 months (90 days +/- 14)	~6 months (180 days +/- 30 days)
	Sliding GOS-E (favorable or Poor outcome)	Sliding GOS-E (favorable or Poor outcome)	Sliding GOS-E (favorable or Poor outcome)
1			
2			
…			
52			Predict this until data collected
53			Predict this until data collected

Prediction uses all Previous Data

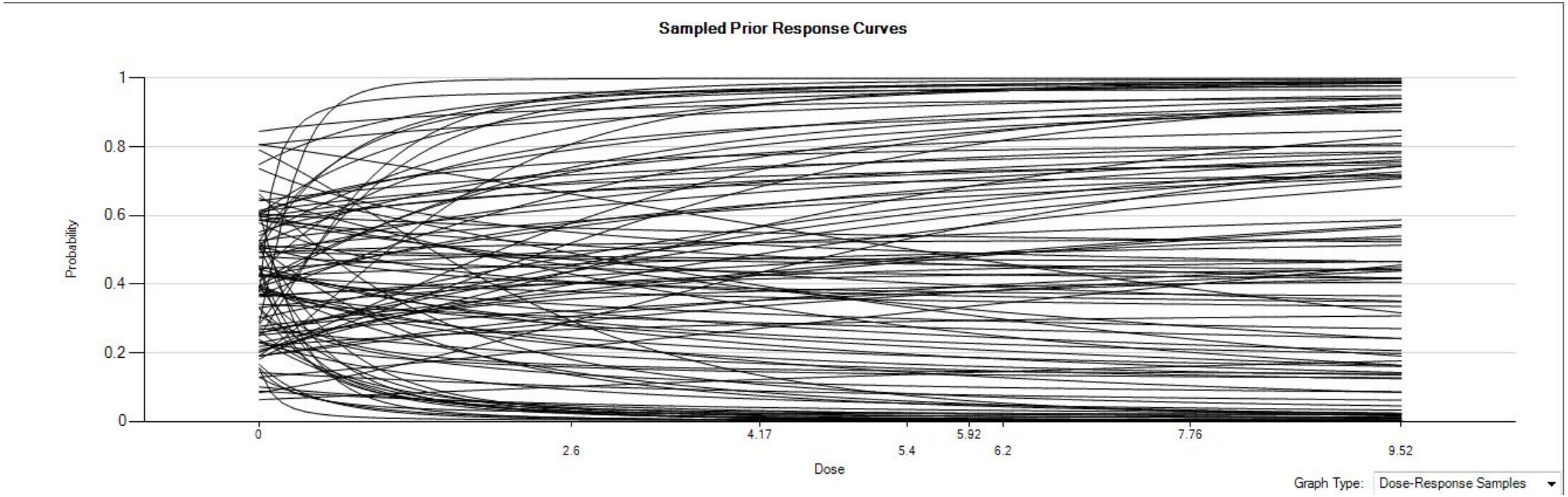
* IMPACT Score (Probability of Poor Outcome) is calculated at baseline

Dosing the Arms (key for statistical model)

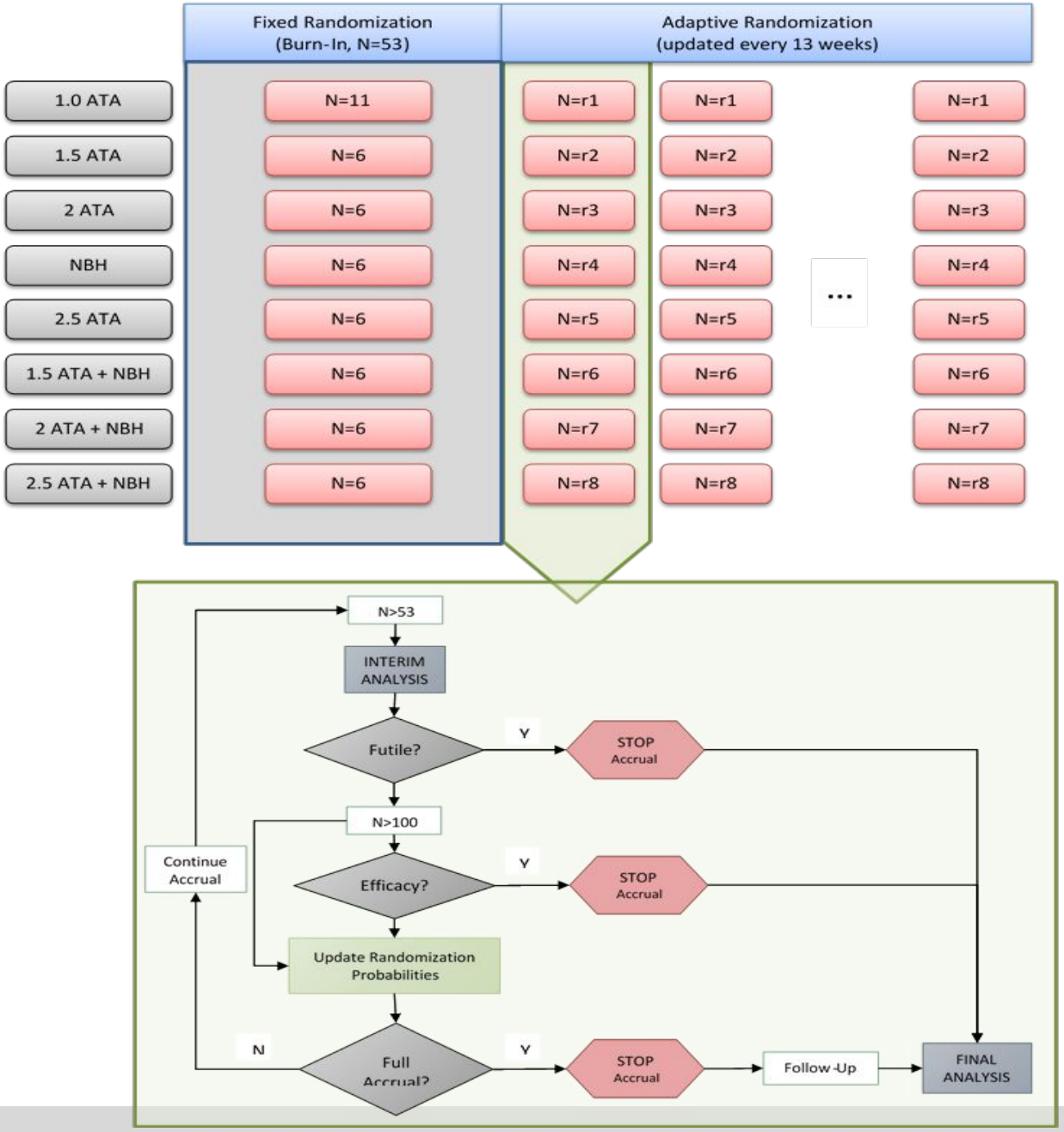
	Arm	Dose (Daily Oxygen Toxicity Units)
1	Control	0
2	1.5 ATA without NBH x 60 minutes BID	260
3	2.0 ATA without NBH x 60 minutes BID	417
4	NBH x 4.5 hours BID	540
5	2.5 ATA without NBH x 60 minutes BID	592
6	1.5 ATA x 60 minutes + 3 hours NBH BID	620
7	2.0 ATA x 60 minutes + 3 hours NBH BID	776
8	2.5 ATA x 60 minutes + 3 hours NBH BID	952

NBH is 100% FiO2 at 1.0 ATA

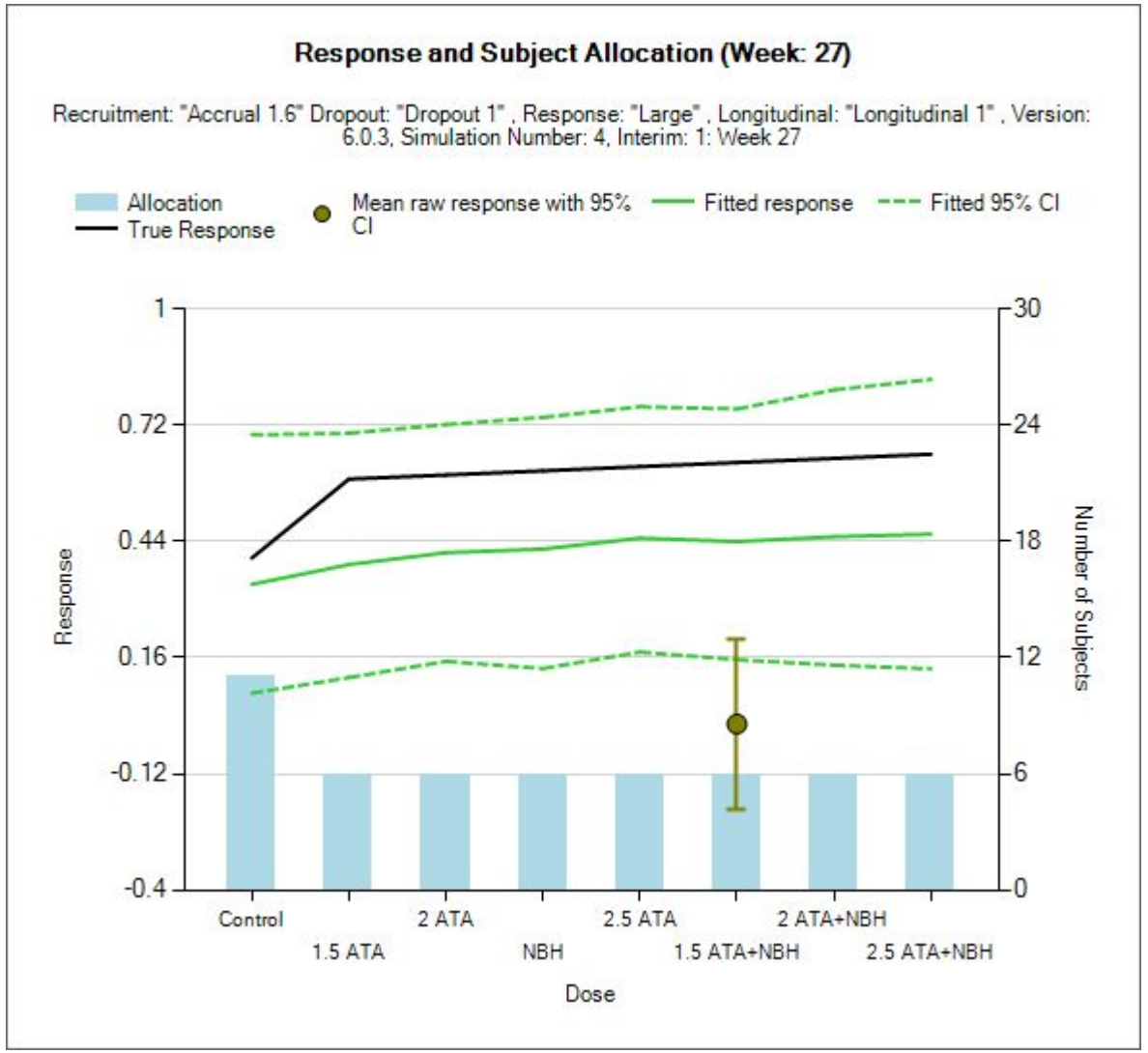
Possible Dose Response Curves*
Prior: Before Trial Begins

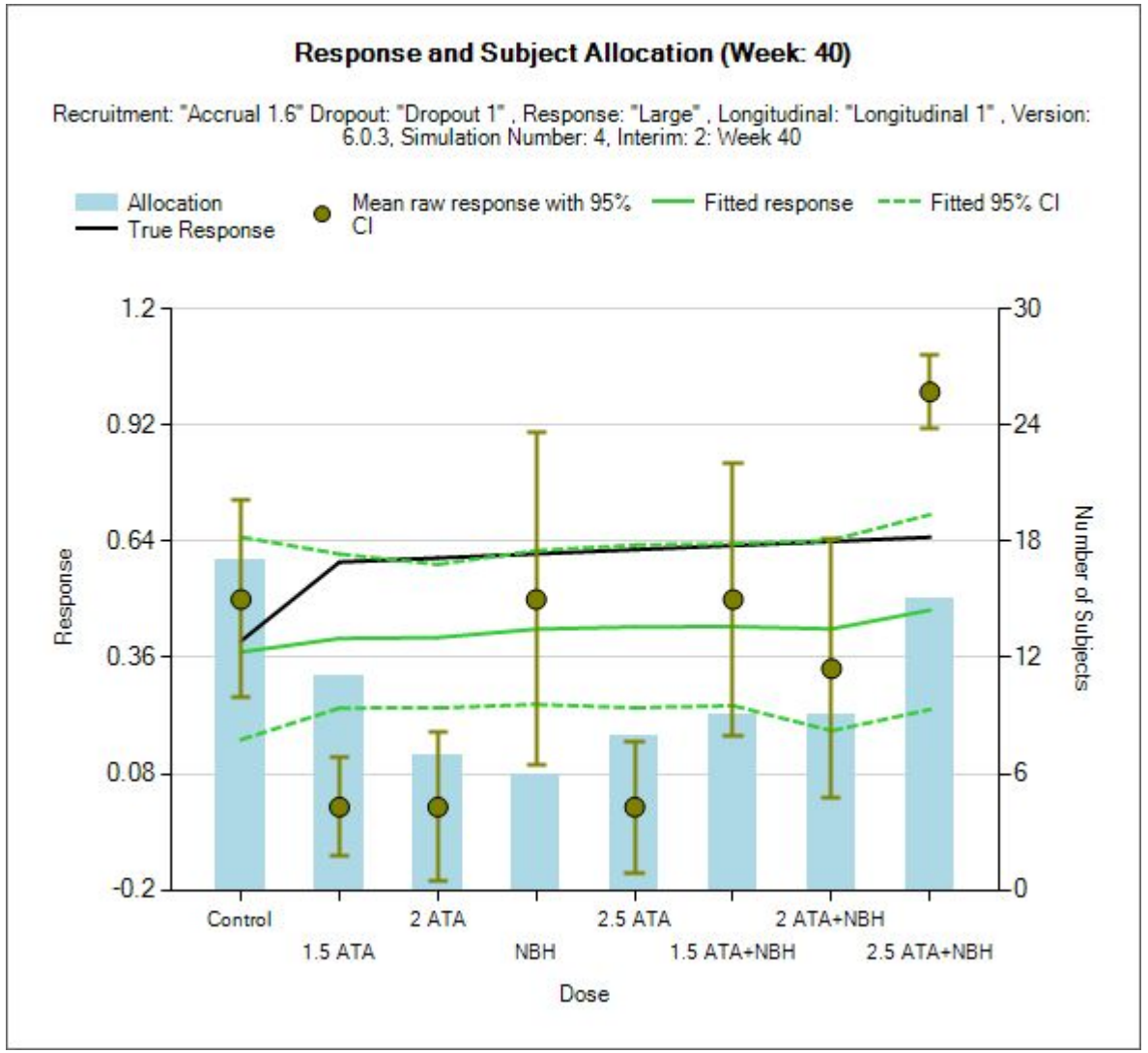


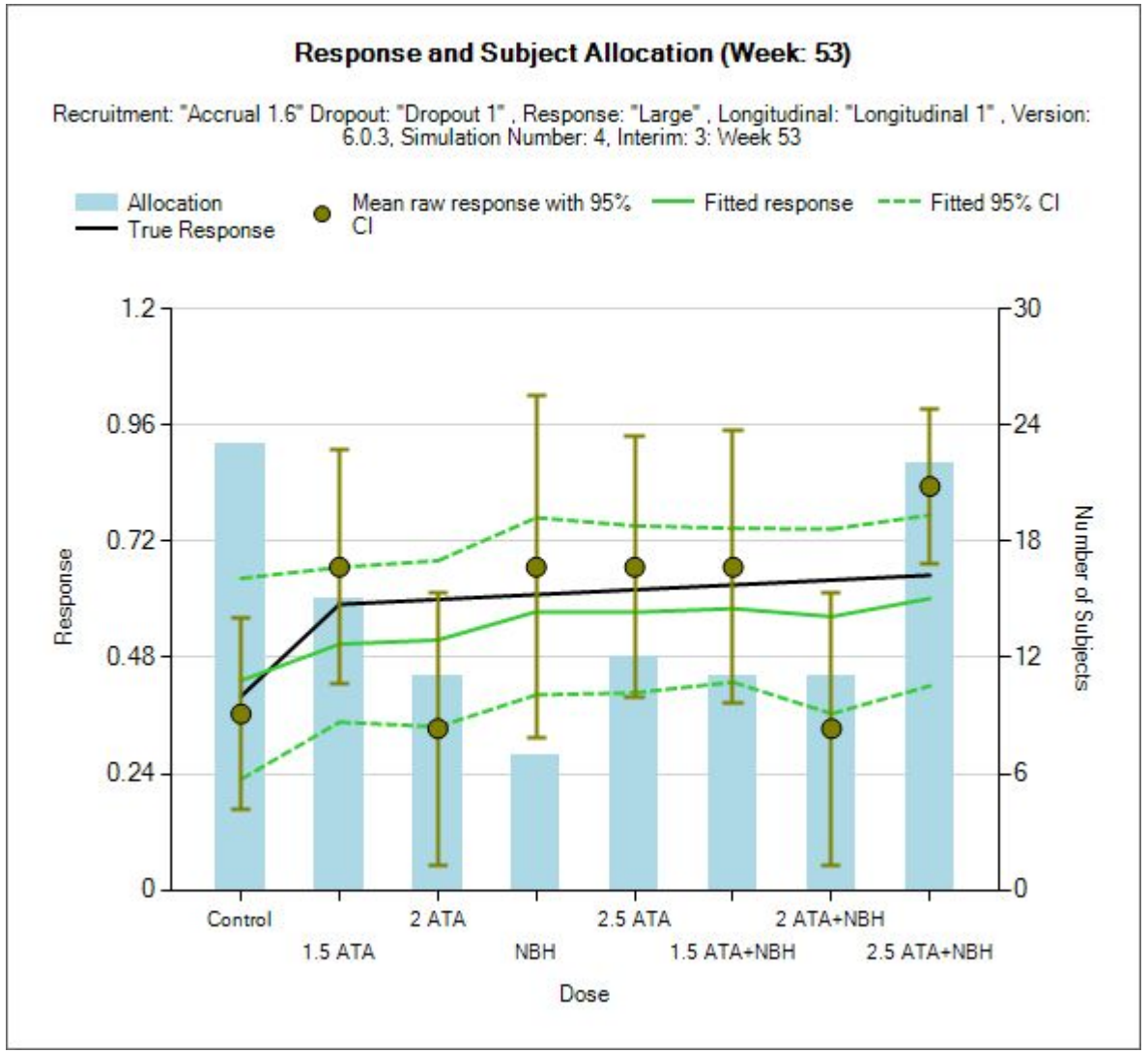
*Hierarchical Logistic Regression

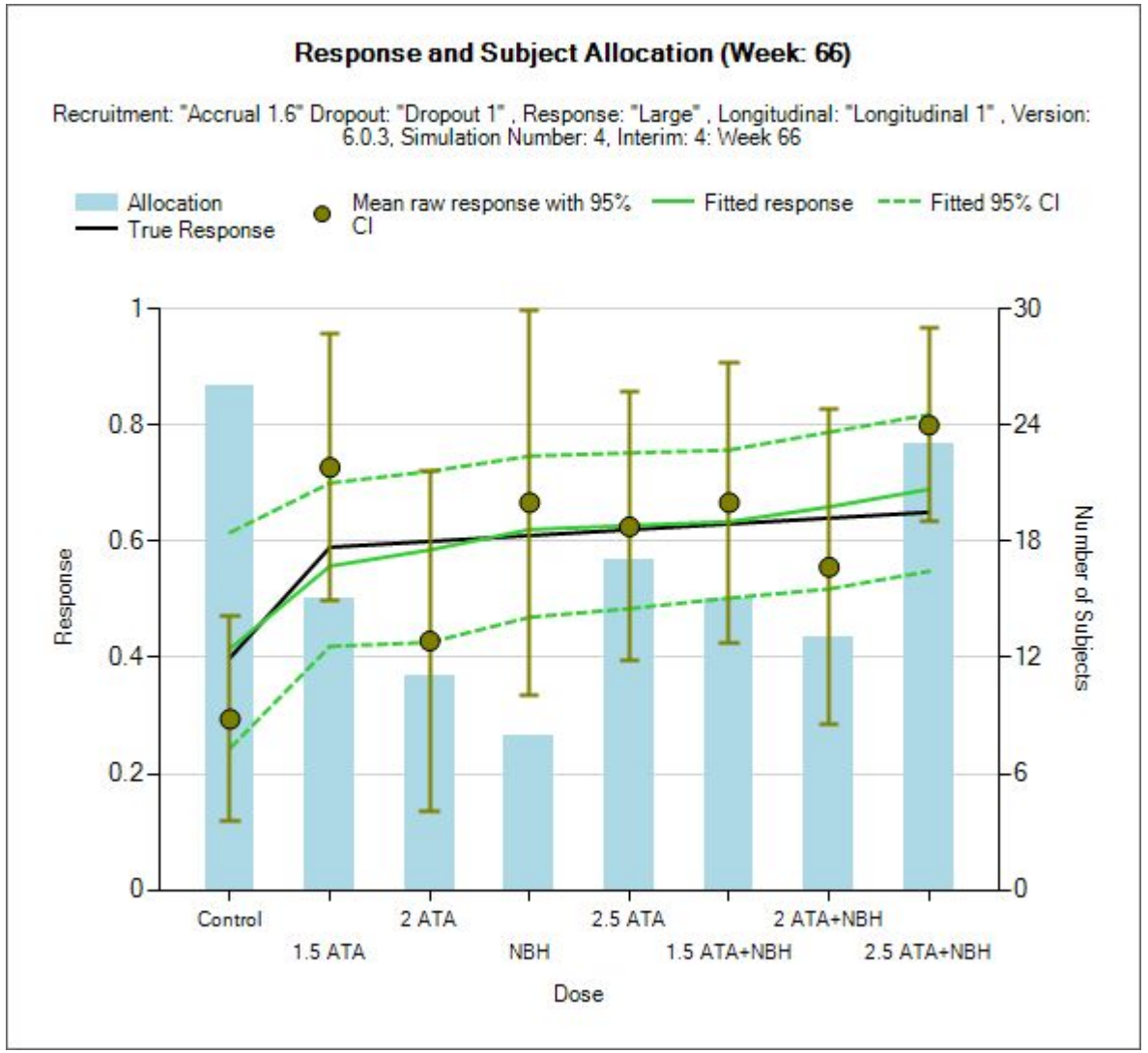

Fixed Randomization
(Burn-In, N=53)
Adaptive Randomization
(updated every 13 weeks)
1.0 ATA
1.5 ATA
2 ATA
NBH
2.5 ATA
1.5 ATA + NBH
2 ATA + NBH
2.5 ATA + NBH
N=11
N=6
N=6
N=6
N=6
N=6
N=6
N=6
N=r1
N=r2
N=r3
N=r4
N=r5
N=r6
N=r7
N=r8
...
N>53
INTERIM ANALYSIS
Futile?
Y
STOP Accrual
N>100
Efficacy?
Y
STOP Accrual
Continue Accrual
Update Randomization Probabilities
N
Full Accrual?
Y
STOP Accrual
Follow-Up
FINAL ANALYSIS

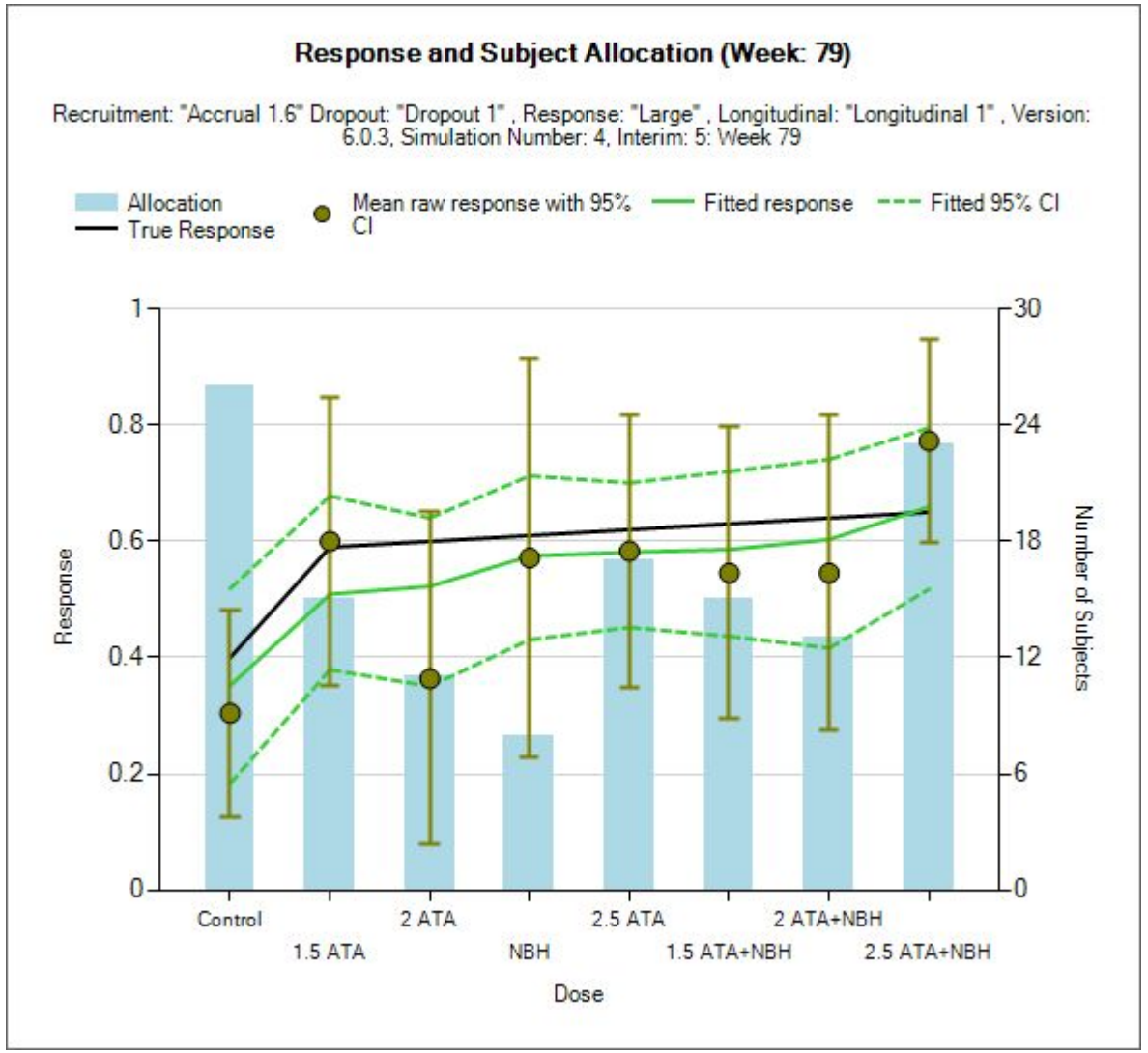
Example Trial with Simulated Data

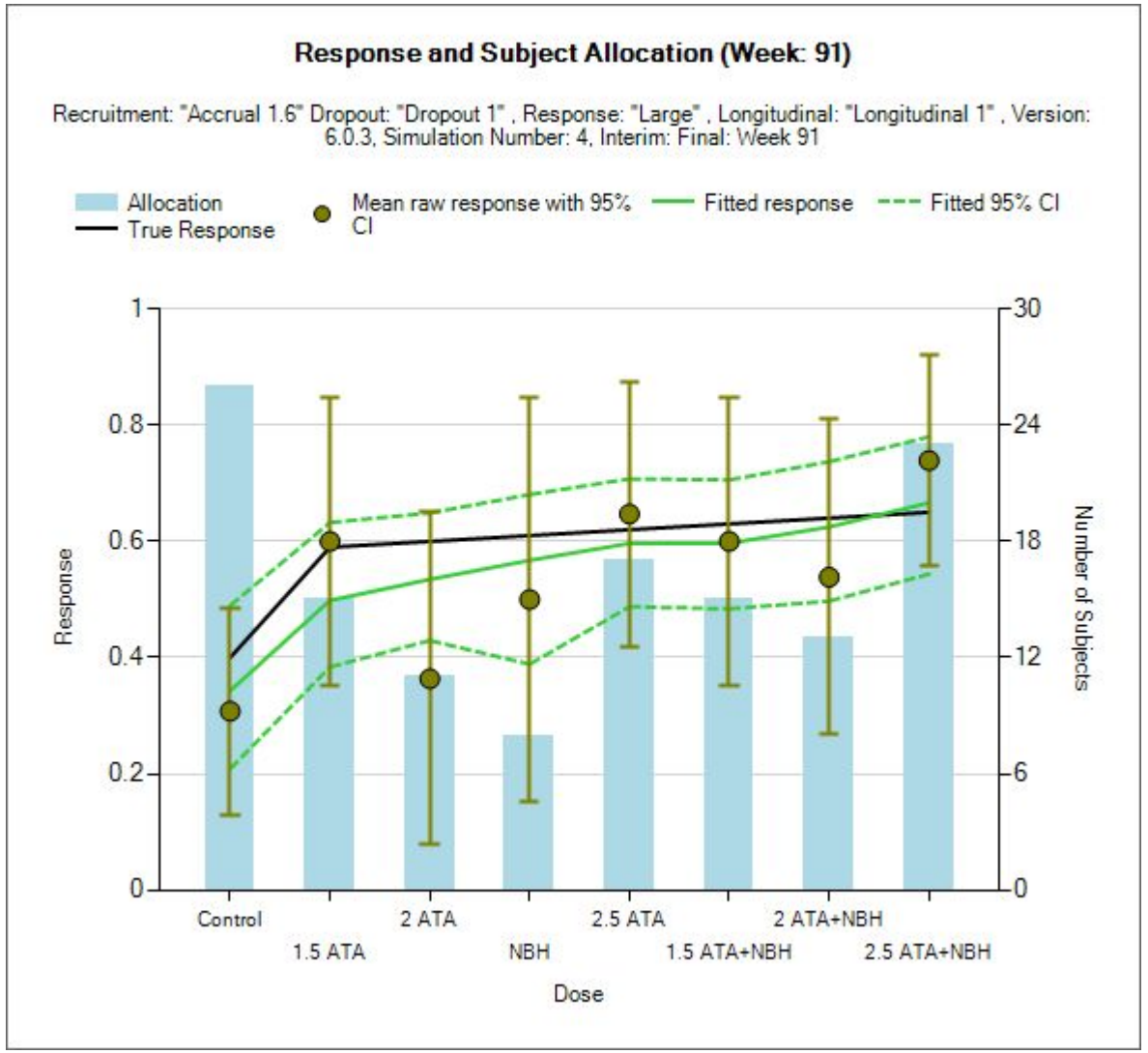

Response and Subject Allocation (Week: 27)
Recruitment: "Accrual 1.6" Dropout: "Dropout 1" , Response: "Large" , Longitudinal: "Longitudinal 1" , Version: 6.0.3, Simulation Number: 4, Interim: 1: Week 27
Allocation
True Response
Mean raw response with 95% CI
Fitted response
Fitted 95% CI
Response
1
0.72
0.44
0.16
-0.12
-0.4
Number of Subjects
30
24
18
12
6
0
Control
1.5 ATA
2 ATA
NBH
2.5 ATA
1.5 ATA+NBH
2 ATA+NBH
2.5 ATA+NBH
Dose


Response and Subject Allocation (Week: 40)
Recruitment: "Accrual 1.6" Dropout: "Dropout 1" , Response: "Large" , Longitudinal: "Longitudinal 1" , Version: 6.0.3, Simulation Number: 4, Interim: 2: Week 40
Allocation
True Response
Mean raw response with 95% CI
Fitted response
Fitted 95% CI
Response
1.2
0.92
0.64
0.36
0.08
-0.2
Number of Subjects
30
24
18
12
6
0
Control
1.5 ATA
2 ATA
NBH
2.5 ATA
1.5 ATA+NBH
2 ATA+NBH
2.5 ATA+NBH
Dose


Response and Subject Allocation (Week: 53)
Recruitment: "Accrual 1.6" Dropout: "Dropout 1" , Response: "Large" , Longitudinal: "Longitudinal 1" , Version: 6.0.3, Simulation Number: 4, Interim: 3: Week 53
Allocation
True Response
Mean raw response with 95% CI
Fitted response
Fitted 95% CI
Response
Number of Subjects
1.2
0.96
0.72
0.48
0.24
0
30
24
18
12
6
0
Control
1.5 ATA
2 ATA
NBH
2.5 ATA
1.5 ATA+NBH
2 ATA+NBH
2.5 ATA+NBH
Dose


Response and Subject Allocation (Week: 66)
Recruitment: "Accrual 1.6" Dropout: "Dropout 1" , Response: "Large" , Longitudinal: "Longitudinal 1" , Version: 6.0.3, Simulation Number: 4, Interim: 4: Week 66
Allocation
True Response
Mean raw response with 95% CI
Fitted response
Fitted 95% CI
Response
Number of Subjects
1
0.8
0.6
0.4
0.2
0
30
24
18
12
6
0
Control
1.5 ATA
2 ATA
NBH
2.5 ATA
1.5 ATA+NBH
2 ATA+NBH
2.5 ATA+NBH
Dose


Response and Subject Allocation (Week: 79)
Recruitment: "Accrual 1.6" Dropout: "Dropout 1" , Response: "Large" , Longitudinal: "Longitudinal 1" , Version: 6.0.3, Simulation Number: 4, Interim: 5: Week 79
Allocation
True Response
Mean raw response with 95% CI
Fitted response
Fitted 95% CI
Response
Number of Subjects
1
0.8
0.6
0.4
0.2
0
30
24
18
12
6
0
Control
1.5 ATA
2 ATA
NBH
2.5 ATA
1.5 ATA+NBH
2 ATA+NBH
2.5 ATA+NBH
Dose


Response and Subject Allocation (Week: 91)
Recruitment: "Accrual 1.6" Dropout: "Dropout 1" , Response: "Large" , Longitudinal: "Longitudinal 1" , Version: 6.0.3, Simulation Number: 4, Interim: Final: Week 91
Allocation
True Response
Mean raw response with 95% CI
Fitted response
Fitted 95% CI
Response
Number of Subjects
1
0.8
0.6
0.4
0.2
0
30
24
18
12
6
0
Control
1.5 ATA
2 ATA
NBH
2.5 ATA
1.5 ATA+NBH
2 ATA+NBH
2.5 ATA+NBH
Dose

Results (Large effect, simulation 4)

- This trial stopped early for success:
  - The dose identified as being the best is dose 8*
  - The posterior probability dose 8 is better than control is 0.9996
  - The proportion of patients with favorable outcome in dose 8 is 0.66 (95% CrI 0.53-0.78)
  - The proportion of patients with favorable outcome in control is 0.34 (95% CrI 0.21-0.49)
  - The posterior predictive probability of phase III success for dose 8 vs Control is 0.99479

*2.5 ATA x 60 minutes + 3 hours NBH BID

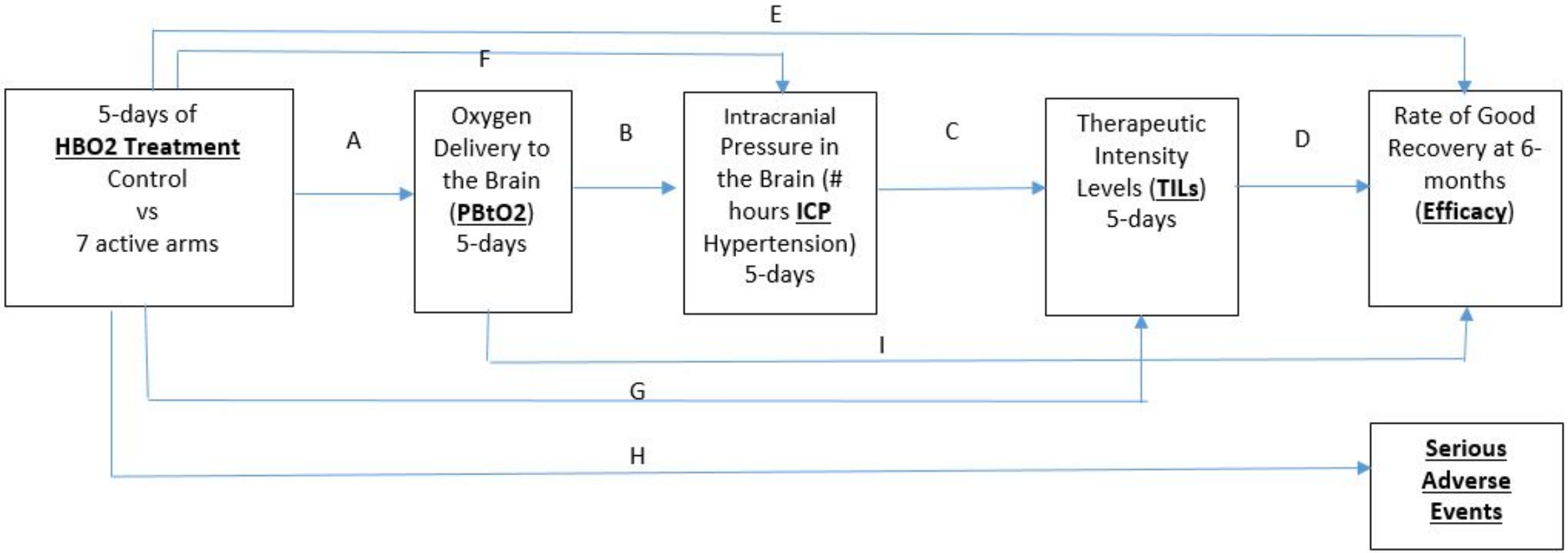
Comparison of Fixed and Bayesian adaptive designs: Large Effect

Trial Design*	Size	Strength	Speed	Benefit (%best arm among active arms)
Fixed & Separate Dose Model (no longitudinal)	200	85%	151 weeks	14%
Bayesian Adaptive	161	97%	126 weeks	22%

Compared to fixed trial Bayesian adaptive design is:
much smaller,
much stronger,
much faster, and
benefits more %patients.

*Both designs have Type I error rates of 20% (aggressive but a comparison of apples to apples)

Secondary Analyses



Thank you!