Bayesian Adaptive Clinical Trials in the 21st Century

Ben Saville, Ph.D.

Berry Consultants

About me

- Ben Saville, Ph.D.
  - Statistical Scientist at Berry Consultants
  - Adjunct assistant professor at Vanderbilt University School of Medicine
  - Assistant Professor of Biostatistics at Vanderbilt 2008-2014
  - Ph.D. in Biostatistics from University of North Carolina at Chapel Hill in 2008
  - Primary expertise
    - Phase 2/3 Bayesian adaptive clinical trials
    - Consulting with medical device & pharmaceutical companies, academic investigators
    - Interactions with FDA

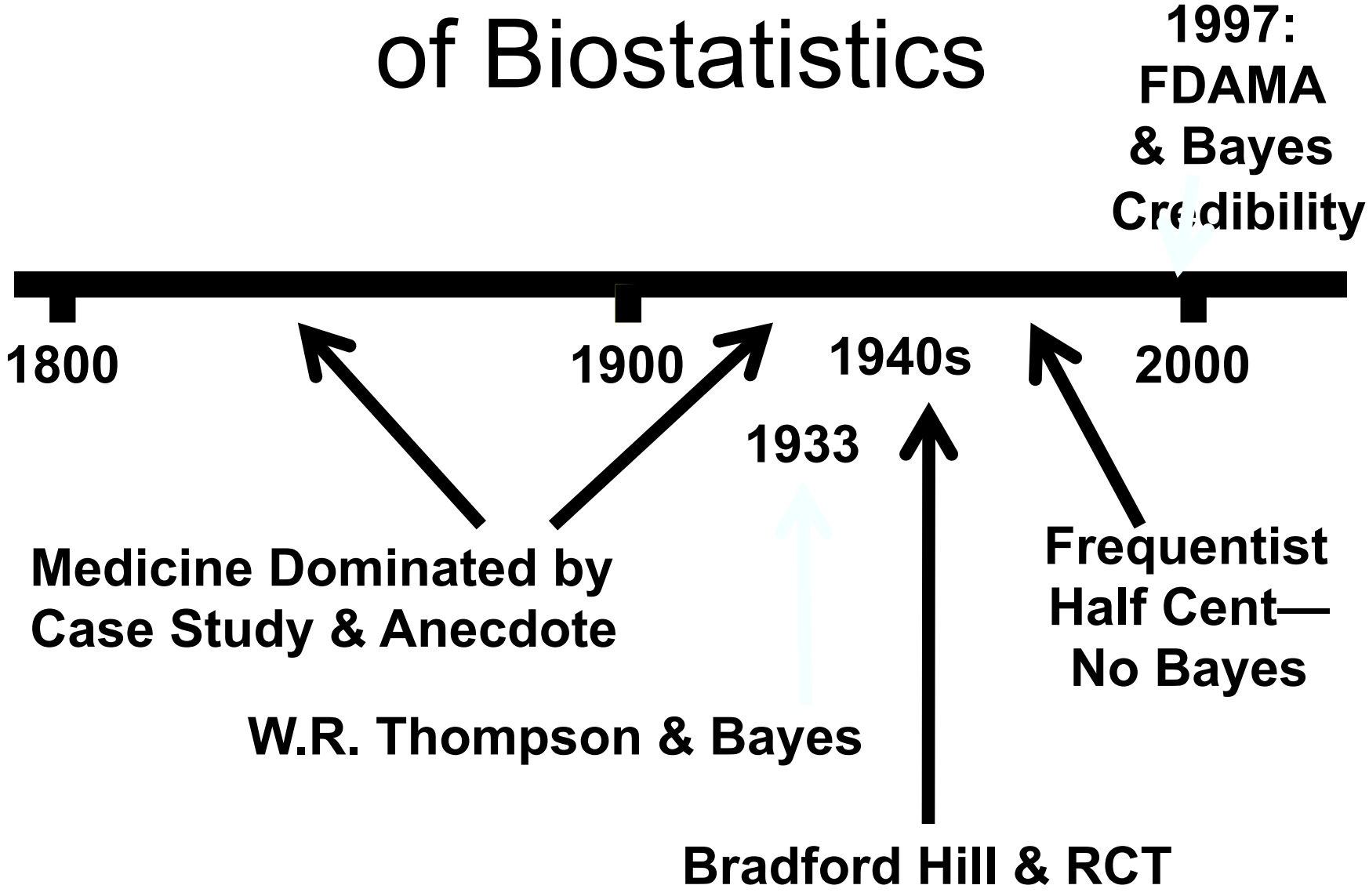
About Berry Consultants

- Founded by Don and Scott Berry in 2000
- Don Berry, Ph.D.
  - Professor of Biostatistics, MD Anderson Cancer Center
  - World renowned Bayesian leader
  - Over 300 published articles and several books
- Scott Berry, Ph.D., President
  - Former assistant professor at Texas A&M University
  - Involved in hundreds of Bayesian adaptive trials
  - Primary designer of FACTS software
- 15 Biostatisticians, 4 Admin/IT, 1 physician

Acknowledgements

- Many individuals contributed to the material in this presentation:
  - Don Berry
  - Scott Berry
  - Jason Connor
  - Roger Lewis
  - Anna McGlothlin

Rough History of Biostatistics



Top 5 Reasons for Bayes

1. On-line learning
2. Predictive probabilities
3. Hierarchical modeling
4. Modeling generally
5. Decision analysis

Clinical trials are the final links in the chains of knowledge and for determining the roles of therapeutic advances. Unfortunately, in an important sense they are the weakest links. … the rocketships of modern biology culminate their final stage of delivery in a wagon train.

ARTICLE INFO

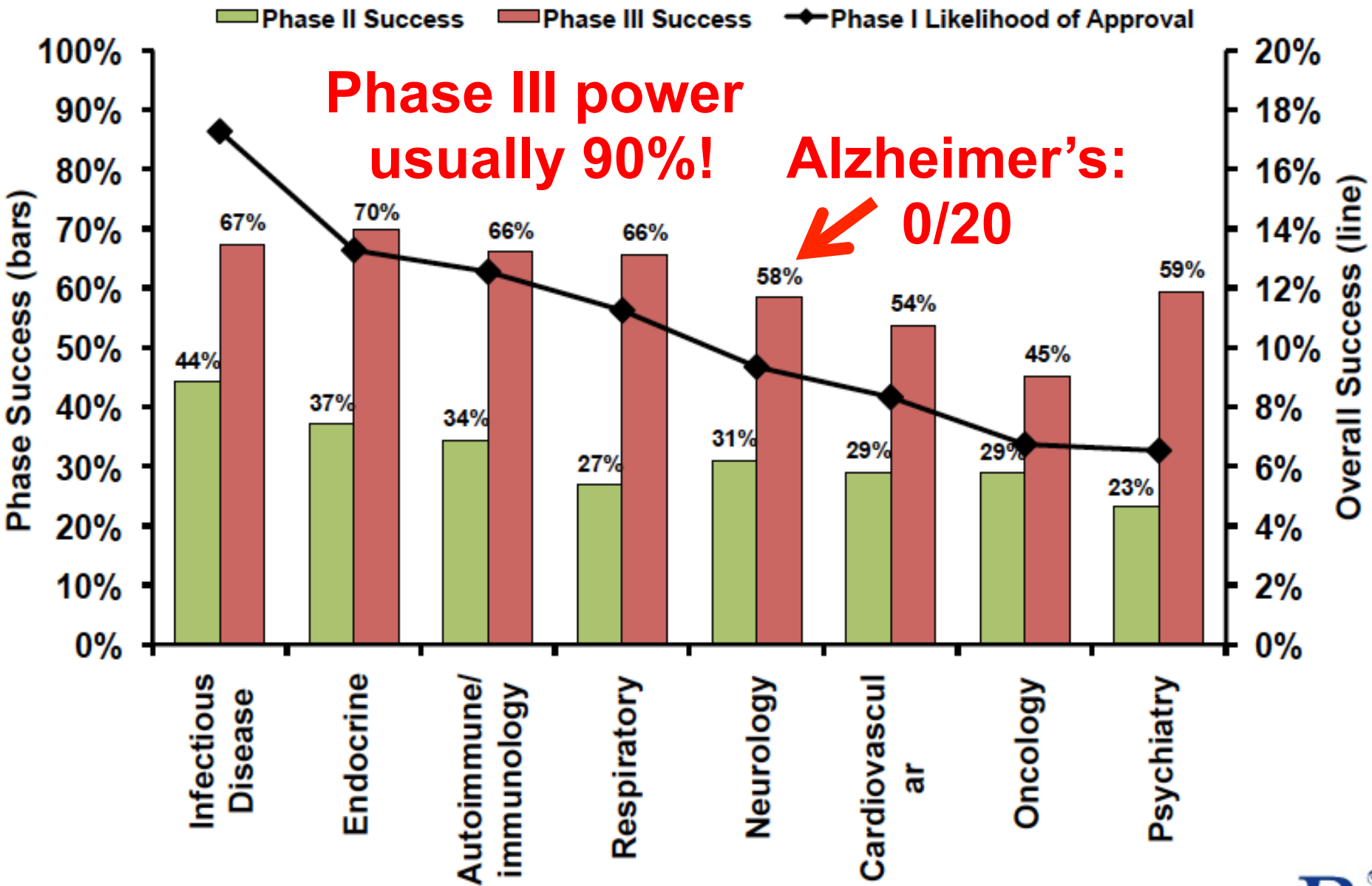
Article history:
Received 24 February 2015
Accepted 24 February 2015
Available online 11 March 2015

Keywords:
Biomarker-driven clinical trials
Platform trials
Basket trials
Bayesian adaptive clinical trials

ABSTRACT

Clinical trials are the final links in the chains of knowledge and for de
therapeutic advances. Unfortunately, in an important sense they a
This article describes two designs that are being explored today: plat
trials. Both are attempting to merge clinical research and clinical pr

SUCCESS AT PHASE II AND III



biomed tracker

Jan 2003 through Aug 2012

Bio BIOTECHNOLOGY INDUSTRY ORGANIZATION

Why Phase III Failures?

- Estimated cost per successful drug: $1.8 Billion
- Ineffective drug
  - Wrong endpoint in phase II
  - No randomization in phase II
  - Lottery
  - Regression to the mean
  - Silly subsetting
- Effective drug, lousy strategy
  - Underpowered
  - Wrong dose/schedule/concomitant Rx
  - Wrong population

Janet Woodcock (2006)
Dir CDER FDA

"Improved utilization of adaptive and Bayesian methods" could help resolve low success rate of and expense of phase III clinical trials

FDA's Critical Path Opportunities Report (2006)

"uncovered a consensus
that the two most important areas
for improving medical product
development are
biomarker development and
streamlining clinical trials."

http://www.fda.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/default.htm

Why are Study Designs (Usually) Fixed

- It's easiest to calculate type I error rates if the design parameters of the trial are all constant
- Results obtained using "Standard approaches" are generally considered valid
- Logistically simpler to execute
- Fixed designs are less sensitive to drift in the characteristics of subjects over time
  - Fears worse than reality
- We could do the math 40 years ago
  - We still can but we can also do more sophisticated things now too

Why Adapt?
The Prospective Postmortem

- Consider whether any adaptations might be added to *prospectively* address *potential* regrets

Why Adapt?
The Prospective Postmortem

- Consider whether any adaptations might be added to *prospectively* address *potential* regrets

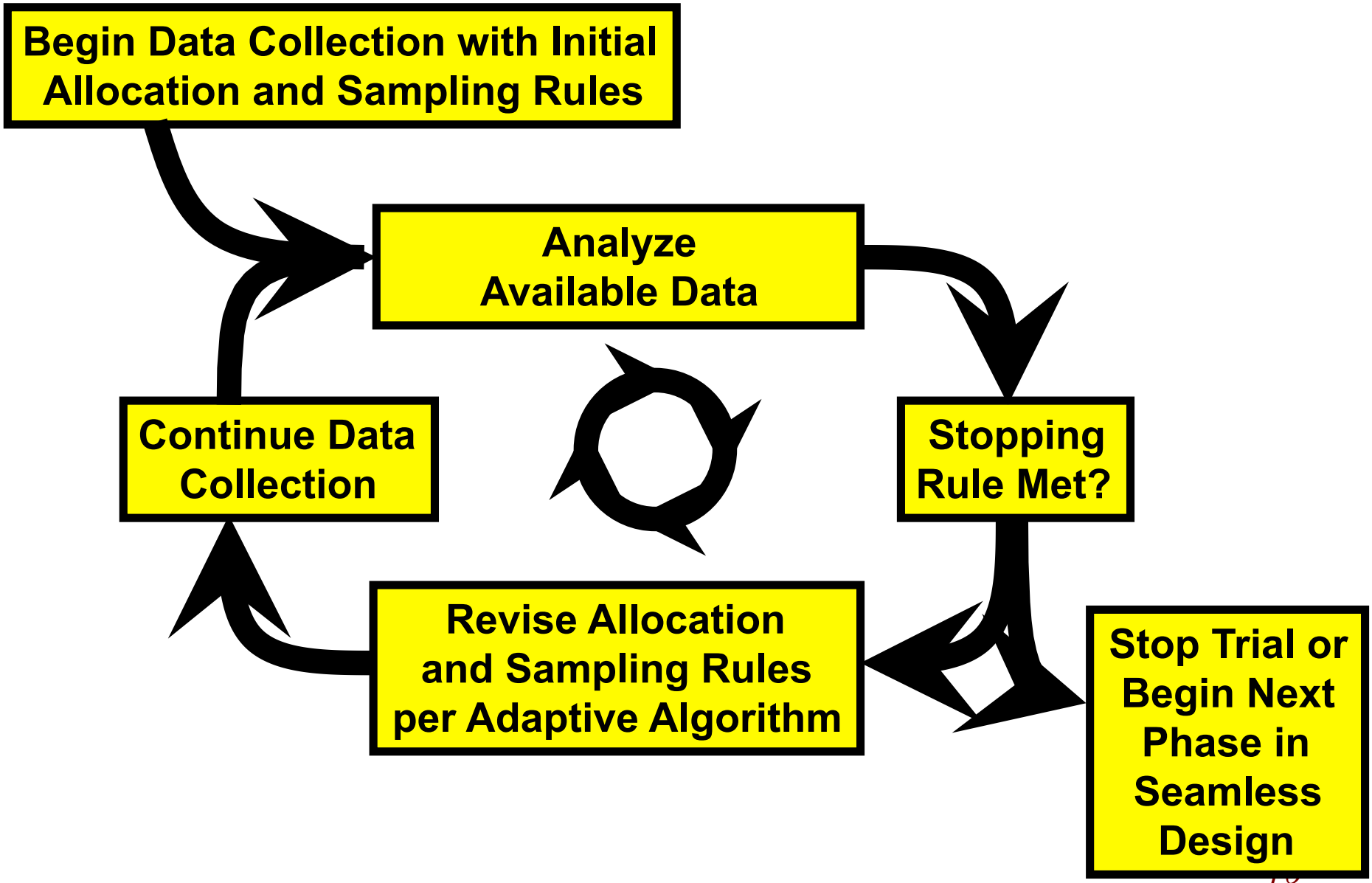
- Be honest with yourself in design Phase
  - We overestimate treatment effects
  - We underestimate variability
  - Because we need to justify a doable trial
  - Because we can't be honest in grant proposals

What are Adaptive Trials?

Trials in which key design parameters change during trial execution based upon *a priori* predefined rules and accumulating data from the trial to achieve goals of validity, scientific efficiency, and safety

- Planned: All possible adaptations defined *a priori*
- Well-defined: Criteria for adapting clearly explained
- Key parameters: *Not* minor inclusion or exclusion criteria, routine amendments, etc.
- Validity: Reliable statistical inference

Typical Prospective Adaptive Design



JAMA 2006;296:1955-1957.

Adaptive Features

- Response-adaptive randomization
- Dose-response modeling
- Adaptive sample size
- Population enrichment
- Explicit longitudinal modeling of the accumulating data based upon interim outcomes
- Extensive simulation of trial performance
- Frequent interim analyses
- Repeatedly ask when are primary questions answered

Traditional Drug Development

- Phase I
  - tens of subjects
  - first use in humans (with or without target illness)
  - generates initial dosing and toxicity information
- Phase II
  - 100 to few hundreds of subjects with target illness
  - gain initial information on dose-response relationship (i.e., "proof of concept"), side effects
- Phase III
  - confirm superiority of new treatment
  - Typically large and expensive

Challenges in Traditional Approach

- Phase II
  - a wide range of doses are possibly the "best" choice
  - consider combinations of treatments?
  - different durations, schedules of treatment?
  - different combinations may work best on patients with different histologies or biomarkers
  - *can not do a fixed trial over all possibilities*

- Currently we pick 2 or 3 (of many possible) doses or combinations, hope we're right, & run a trial

Challenges in Traditional Approach

- Phase II solution: Adaptively randomize
  - start looking across many doses / durations / combos
  - stop enrolling patients unlikely to benefit
  - drop arms / lower randomization probabilities on poorly performing strategies
  - increase randomization probabilities on promising strategies
  - by the end only looking where the effect might be
    - learning about strategies that matter
    - assigning patients to strategies most likely to help them

Challenges in Traditional Approach

- Phase III
  - often still don't really know the right dose
  - don't really know what to expect in the control arm
  - don't know anything about rarer side effects
  - *Yet traditional statistical approaches require that the trial characteristics be completely defined prior to enrolling the first phase III patient*

Challenges in Traditional Approach

- Phase III Solution: adaptive sample size
  start with 2+ arms & drop all but one & control
  - measure treatment effect as trial progresses
  - measure variability & control event rate as we go
  - ask "If we stop enrolling now & track patients will we have sufficient evidence in one year?"
    If so stop accrual, wait, perform the final analysis
  - ask "If we enroll to the max will we have high chance of achieving goal?"
    If not stop for futility
  - use predictive probabilities based on in-trial data to guide sample size

When is Adaptation Most Valuable

- Outcomes or biomarkers available rapidly relative to time required for entire trial
- Substantial morbidity, risks, costs
- Large uncertainty regarding relative efficacy, adverse event rates, variability, patient population in trial, etc.
- Logistically practical
- Able to secure buy-in of stakeholders

Some Current Areas of Application

- Alzheimer's Disease
- Aneurysm
- Asthma
- Atrial Fibrillation
- Cancer Diagnostics
- Cancer Screening
- Cancer Therapeutics
- Crohn's Disease
- Diabetes
- DVT
- Ebola
- Heart Valves
- Ebola
- Emphysema
- HIV
- Libido
- Lymphoma
- Lung Cancer
- Lupus
- Migraines
- Multiple Sclerosis
- Obesity
- Pain
- Parkinson's
- Pandemic Flu
- Pre-term Labor
- Rheumatoid Arthritis
- Sepsis
- Smoking Cessation
- Spinal Cord Injury
- Spinal Implants
- Stroke
- Tinnitus
- Uterine Cancer
- Vaccines

Example
Adaptive Dose Finding Trial

- Treatment of Post-Operative Ileus
  - Major abdominal surgery
  - IV infusion after surgery
  - No approved drug
- Primary endpoint is *recovery of bowel function*
- Intrinsically a time-to-event endpoint
- Placebo median of ≈100 hours
- Clinically significant difference ≈ 10-15 hours
- Censor & offer rescue meds at 168 hours

Details

- 7 active doses available:
  - 20, 40, 80, 160, 320, 480, 600 *m*g/kg
- Placebo comparison
- What do we want to learn?
  - ED_{90} dose?
  - MED dose?
  - Achieve clinically significant difference?
  - Can we run phase III with a reasonable size?

Goals of Adaptive Design

- Find the ED_{90}
  - 90% of max effect
  - > 15 hour Δ vs. placebo
- Learn about ED_{90}
  - Pr(Beat Placebo in Phase III)
  - Use predictive probabilities
- Find MED
  - Smallest dose with 15 hour Δ vs. placebo

Use NDLM to model dose-response curve

$$t_{i,d} \sim F_d(t) = 1 - \frac{\theta_d^4}{\theta_d^4 + t_d^4}$$

$$\theta_1 \sim N(100, 1000^2)$$

$$\theta_d \sim N(\theta_{d-1}, \tau^2), \quad d \in \{2, \ldots 8\}$$

$$\tau^2 \sim \Gamma^{-1}(2, 0.0078)$$

Prior on tau^2 has mean s=8, n=4 observations worth of information a = n/2 b = 2/(ns^2)

Adaptive Randomization

- Randomize 5 patients per dose, then adapt
- Update randomization probabilities (r_d) every week
- Focus randomization on areas of interest
  - ED_{90}
  - MED
  - Phase III power of ED_{90}
- Once we "know" ED90 or MED focus randomization on the other area
  - Pr(*d* is ED_{90}) > 0.60
  - Pr(*d'* is MED) > 0.60
- If $r_d < 0.05$, drop dose & rescale probabilities
  - Dropped doses may re-enter

Early Stopping Rules

- Look at the data every week
- Stop for Success if

  Pr(d is the ED90) $\geq$ 0.60 for some d in {2…8}

  Pr(d' is the MED) $\geq$ 0.60 for some d' in {2…8}
- Stop for Futility if

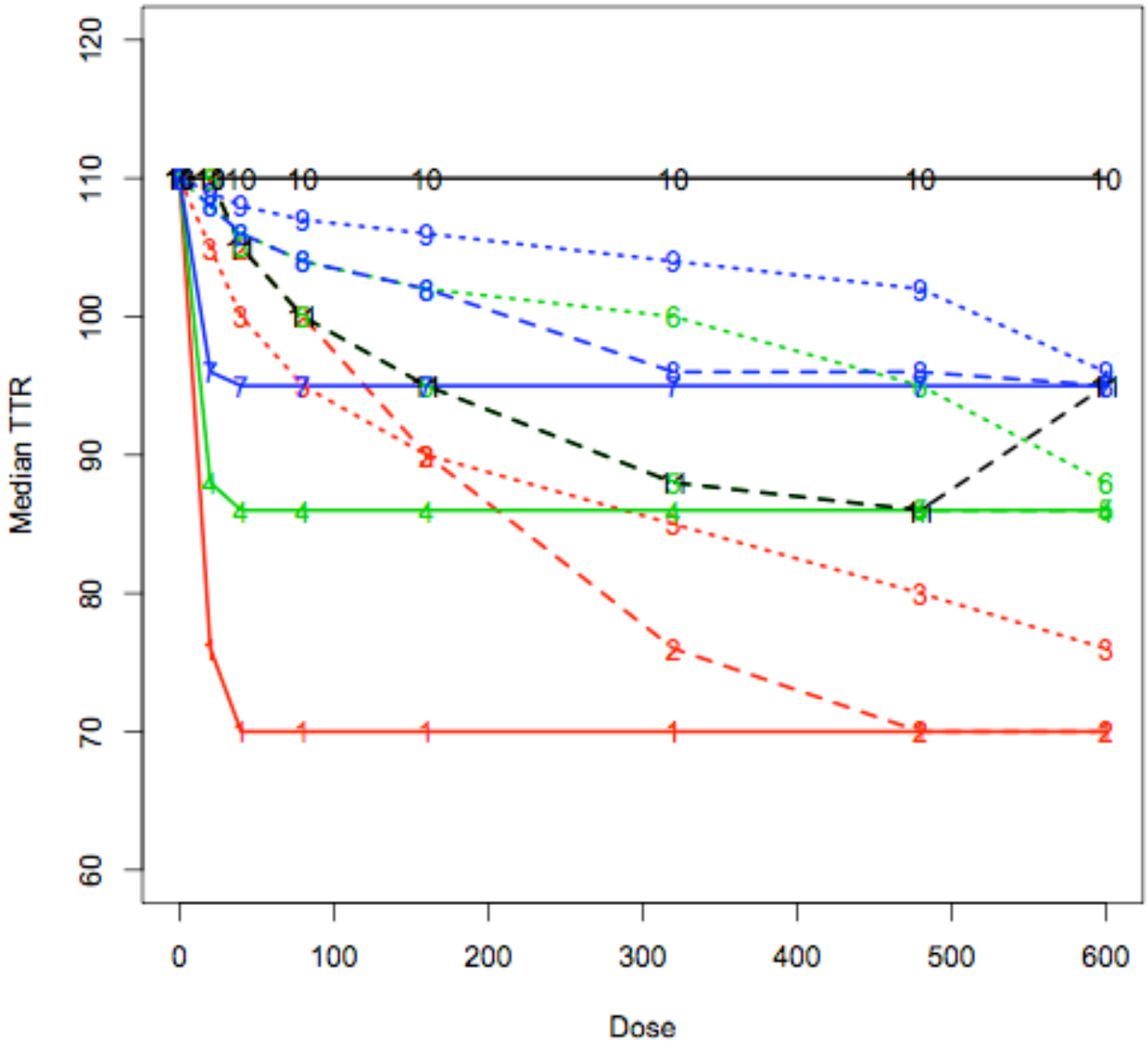
  $\geq$ 100 patients enrolled

  Pr(most likely ED90 wins Phase 3 > 0.80) < 0.20
- Otherwise update randomization probabilities

  & keep enrolling & repeat analysis in one week
- Maximum sample size = 250

Scenarios for Simulation

Scenario	0	20	40	80	160	320	480	600	MED	ED90
#1 (Large Effect, Low ED90)	110	76	70	70	70	70	70	70	20	40
#2 (Large Effect, Medium ED90)	110	110	105	100	90	76	70	70	160	480
#3 (Large Effect, High ED90)	110	105	100	95	90	85	80	76	80	600
#4 (Medium Effect, Low ED90)	110	88	86	86	86	86	86	86	20	20
#5 (Medium Effect, Medium ED90)	110	110	105	100	95	88	86	86	160	320
#6 (Medium Effect, High ED90)	110	108	106	104	102	100	95	88	480	600
#7 (Small Effect, Low ED90)	110	96	95	95	95	95	95	95	40	20
#8 (Small Effect, Medium ED90)	110	108	106	104	102	96	96	95	600	320
#9 (Small Effect, High ED90)	110	109	108	107	106	104	102	96	None	600
#10 (No Effect)	110	110	110	110	110	110	110	110	None	None
#11 (Medium Effect, Non-Monotone)	110	110	105	100	95	88	86	95	160	320


Median TTR
120
110
100
90
80
70
60
0
100
200
300
400
500
600
Dose

Scenario	0	20	40	80	160	320	480	600	MED	ED90
#1 (Large Effect, Low ED90)	110	76	70	70	70	70	70	70	20	40
#2 (Large Effect, Medium ED90)	110	110	105	100	90	76	70	70	160	480
#3 (Large Effect, High ED90)	110	105	100	95	90	85	80	76	80	600

#	P(S)	Early	Cap	Fut	SS		Sample Size Probability Selected Phase III Dose Probability of MED						
					Mean	SD	0	20	40	80	160	320	480	600
1	1.000	0.450	0.550	0.000	194	63	35.9	23.5 0.01 **0.78**	30.9 **0.15** 0.19	26.1 0.11 0.02	22.5 0.07 0.01	19.4 0.04 0.00	17.7 0.03 0.00	18.0 0.04 0.00
2	1.000	0.860	0.140	0.000	129	68	23.0	7.8 0.00 0.07	10.8 0.00 0.12	12.7 0.00 0.19	15.1 0.00 **0.26**	20.0 0.14 0.28	20.7 **0.41** 0.07	18.6 0.30 0.01
3	0.986	0.638	0.360	0.002	169	70	30.9	9.7 0.00 0.05	15.1 0.00 0.19	17.5 0.00 **0.18**	20.2 0.01 0.21	22.9 0.03 0.17	25.2 0.15 0.12	27.7 **0.44** 0.07

Scenario	0	20	40	80	160	320	480	600	MED	ED90
#4 (Medium Effect, Low ED90)	110	88	86	86	86	86	86	86	20	20
#5 (Medium Effect, Medium ED90)	110	110	105	100	95	88	86	86	160	320
#6 (Medium Effect, High ED90)	110	108	106	104	102	100	95	88	480	600

#	P(S)	Early	Cap	Fut	SS		Sample Size Probability Selected Phase III Dose Probability of MED							
					Mean	SD	0	20	40	80	160	320	480	600
4	0.836	0.181	0.810	0.009	221	47	41.2	24.5 **0.01** **0.32**	32.7 0.05 0.30	28.5 0.04 0.10	25.6 0.02 0.06	23.2 0.01 0.02	21.5 0.02 0.02	23.5 0.04 0.01
5	0.894	0.393	0.597	0.010	199	63	37.0	8.9 0.00 0.02	15.2 0.00 0.08	19.8 0.00 0.15	24.3 0.00 **0.22**	31.3 **0.10** 0.24	31.4 0.12 0.13	30.8 0.16 0.06
6	0.672	0.370	0.566	0.064	196	62	36.3	10.1 0.00 0.01	15.8 0.00 0.06	18.1 0.00 0.06	20.2 0.00 0.07	23.7 0.01 0.07	30.3 0.02 **0.14**	41.6 **0.34** 0.25

Scenario	0	20	40	80	160	320	480	600	MED	ED90
#7 (Small Effect, Low ED90)	110	96	95	95	95	95	95	95	40	20
#8 (Small Effect, Medium ED90)	110	108	106	104	102	96	96	95	600	320
#9 (Small Effect, High ED90)	110	109	108	107	106	104	102	96	None	600

#	P(S)	Early	Cap	Fut	SS		Sample Size Probability Selected Phase III Dose Probability of MED							
					Mean	SD	0	20	40	80	160	320	480	600
7	0.383	0.069	0.827	0.104	222	45	41.3	20.6 **0.00** 0.08	28.8 0.01 **0.13**	28.3 0.01 0.07	26.5 0.01 0.04	23.9 0.01 0.02	23.9 0.01 0.02	28.5 0.02 0.01
8	0.407	0.134	0.761	0.105	215	51	39.9	11.4 0.00 0.01	17.7 0.00 0.04	21.1 0.00 0.05	25.2 0.00 0.07	31.9 0.02 0.13	31.7 **0.03** 0.06	35.9 0.08 **0.04**
9	0.254	0.123	0.664	0.213	203	58	37.5	11.6 0.00 0.01	17.5 0.00 0.02	19.8 0.00 0.03	21.8 0.00 0.03	24.1 0.00 0.03	28.6 0.01 0.04	42.2 **0.11** 0.08

Scenario	0	20	40	80	160	320	480	600	MED	ED90
#10 (No Effect)	110	110	110	110	110	110	110	110	None	None
#11 (Medium Effect, Non-Monotone)	110	110	105	100	95	88	86	95	160	320

#	P(S)	Early	Cap	Fut	SS		Sample Size Probability Selected Phase III Dose Probability of MED						
					Mean	SD	0	20	40	80	160	320	480	600
10	0.016	0.006	0.384	0.610	174	61	31.5	14.3 0.00 0.00	20.0 0.0 0.00	21.1 0.00 0.00	20.2 0.00 0.00	20.0 0.00 0.00	20.4 0.00 0.00	26.2 0.01 0.01
11	0.774	0.275	0.691	0.034	213	50	39.8	10.1 0.00 0.01	17.6 0.00 0.09	23.0 0.00 0.13	28.6 0.02 **0.22**	35.7 **0.12** 0.25	33.3 0.12 0.07	25.2 0.02 0.00

Reported to DMC Each Week

Posterior Summaries

TRT	N	Mean TH	SD TH	P(ed90)	P(Med)	P(III)	Rand
1	4	82.5	10.4	0.000	0.000	0.025	0.200
2	5	79.9	8.8	0.018	0.072	0.295	0.087
3	4	74.7	9.4	0.171	0.174	0.508	0.484
4	4	78.5	8.6	0.061	0.046	0.397	0.144
5	4	83.9	9.1	0.019	0.014	0.259	0.000
6	4	83.2	9.9	0.035	0.017	0.288	0.085
7	4	85.9	10.1	0.019	0.012	0.233	0.000
8	5	87.5	11.3	0.021	0.010	0.215	0.000

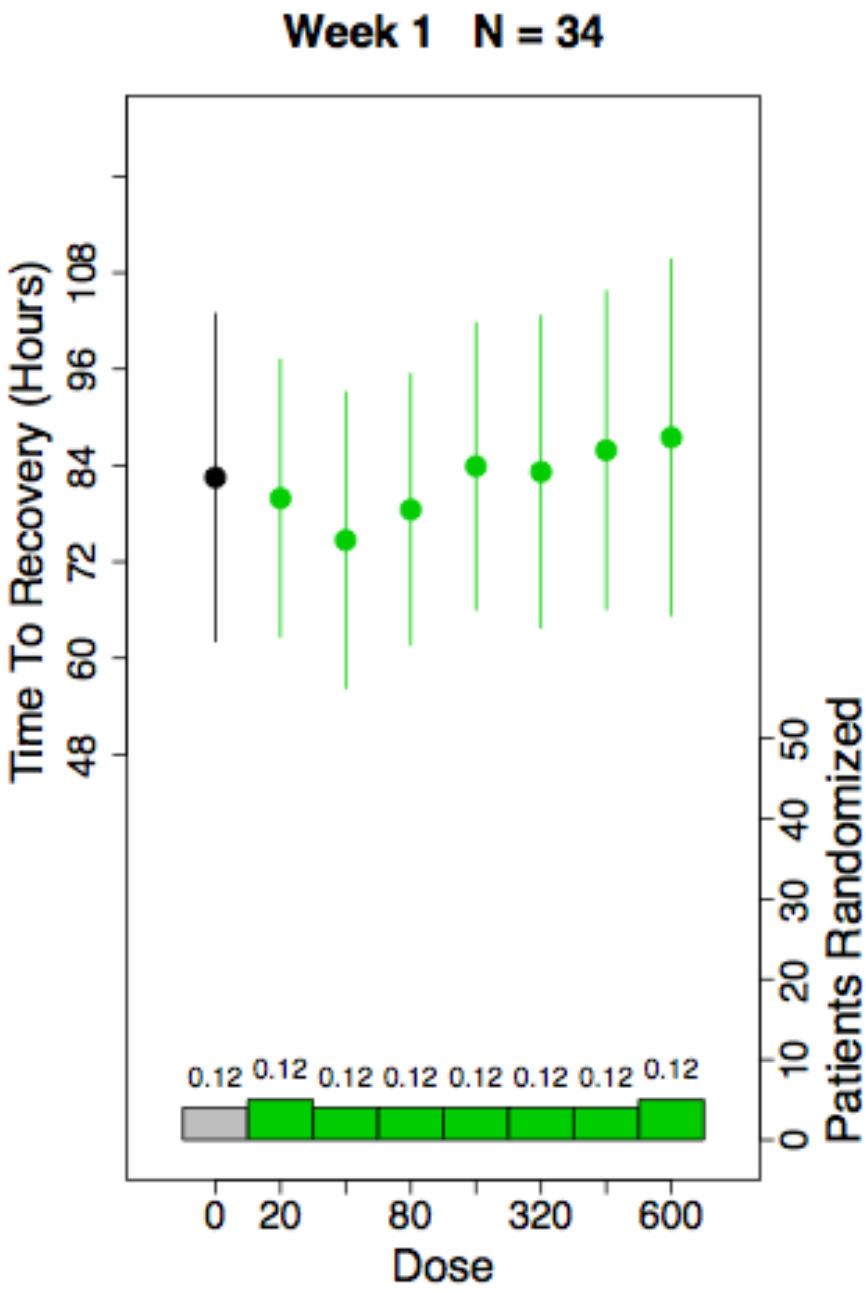
Decisions

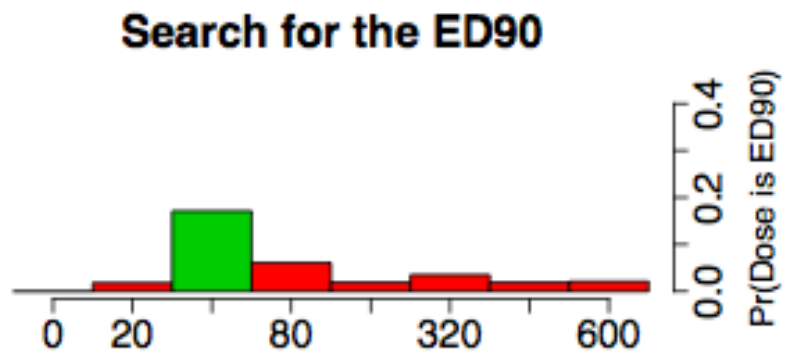
Max P(ed90) = 0.171
Is the maximum P(ed90) > 0.60? NO

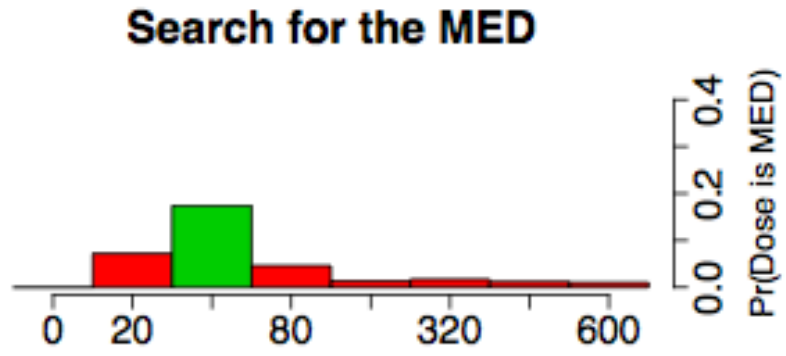
P(III | Max)= 0.508
Is the P(III) > 0.80? NO

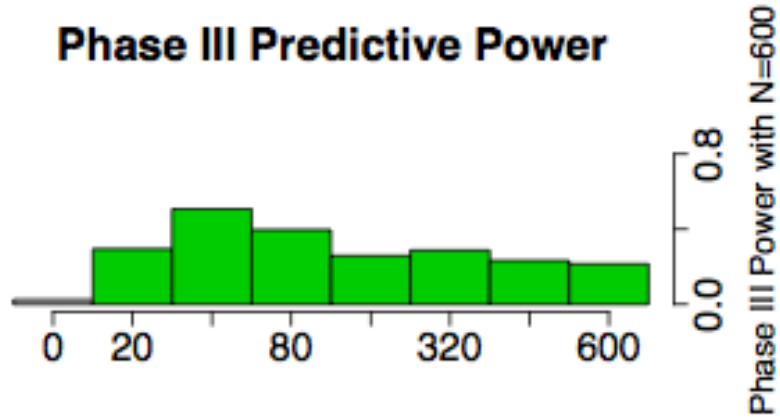
Max P(MED) = 0.174
Is the maximum P(MED) > 0.60? NO

Decision = Continue
Keep sampling? YES


Week 1 N = 34
Time To Recovery (Hours)
48
60
72
84
96
108
0.12 0.12 0.12 0.12 0.12 0.12 0.12 0.12
0
20
80
320
600
Dose
0
10
20
30
40
50
Patients Randomized


Search for the ED90
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is ED90)


Search for the MED
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is MED)


Phase III Predictive Power
0
20
80
320
600
0.0
0.8
Phase III Power with N=600

DMC Report

Posterior Summaries

TRT	N	Mean TH	SD TH	P(ed90)	P(Med)	P(III)	Rand
1	6	81.9	10.5	0.000	0.000	0.025	0.200
2	6	77.3	8.5	0.032	0.130	0.389	0.113
3	5	70.0	9.7	0.303	0.244	0.636	0.603
4	5	77.3	9.0	0.051	0.026	0.413	0.084
5	4	89.8	10.7	0.003	0.003	0.133	0.000
6	5	87.2	10.5	0.013	0.005	0.187	0.000
7	5	92.1	11.4	0.003	0.002	0.117	0.000
8	5	92.5	12.5	0.008	0.004	0.132	0.000

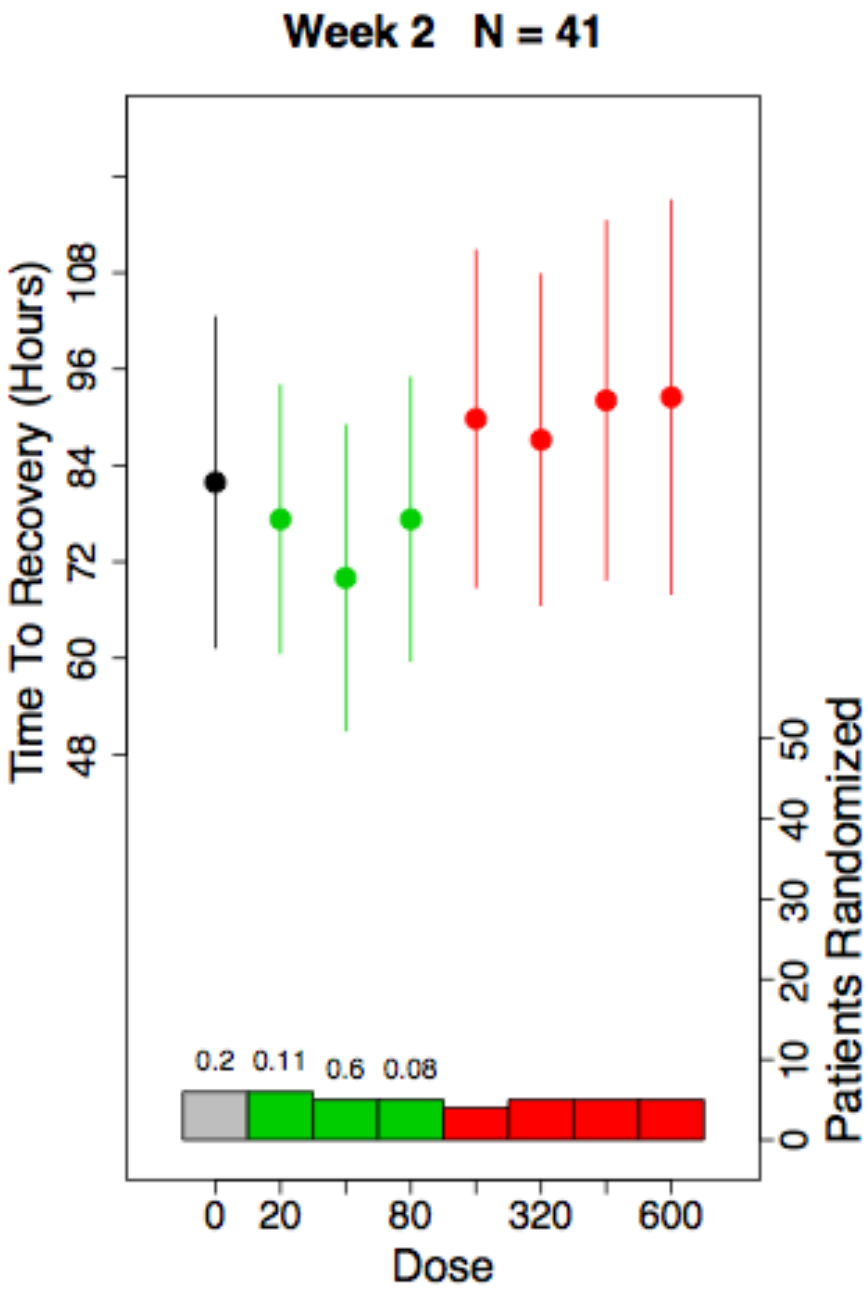
Decisions

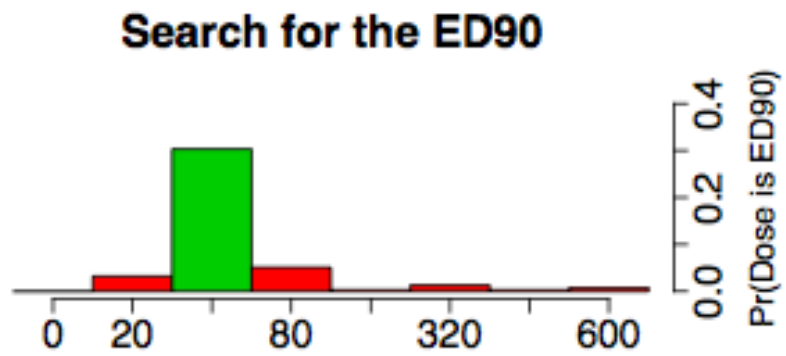
Max P(ed90) = 0.303
Is the maximum P(ed90) > 0.60? NO

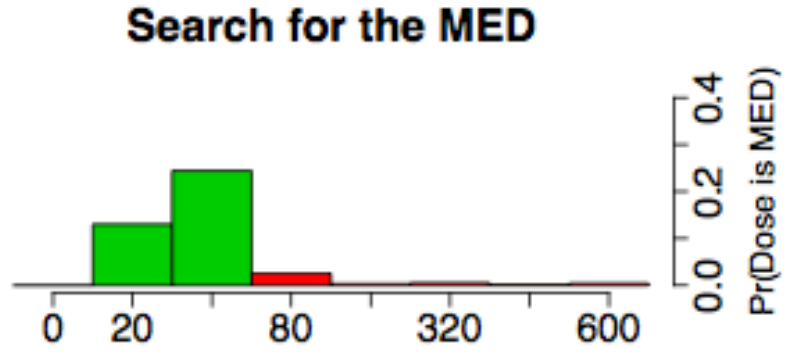
P(III | Max)= 0.636
Is the P(III) > 0.80? NO

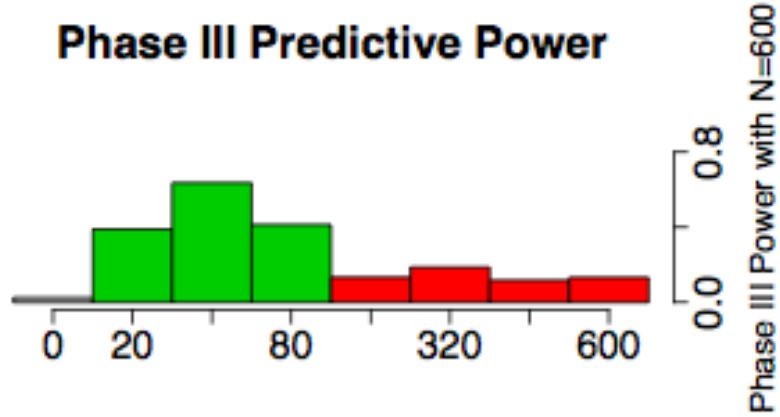
Max P(MED) = 0.244
Is the maximum P(MED) > 0.60? NO

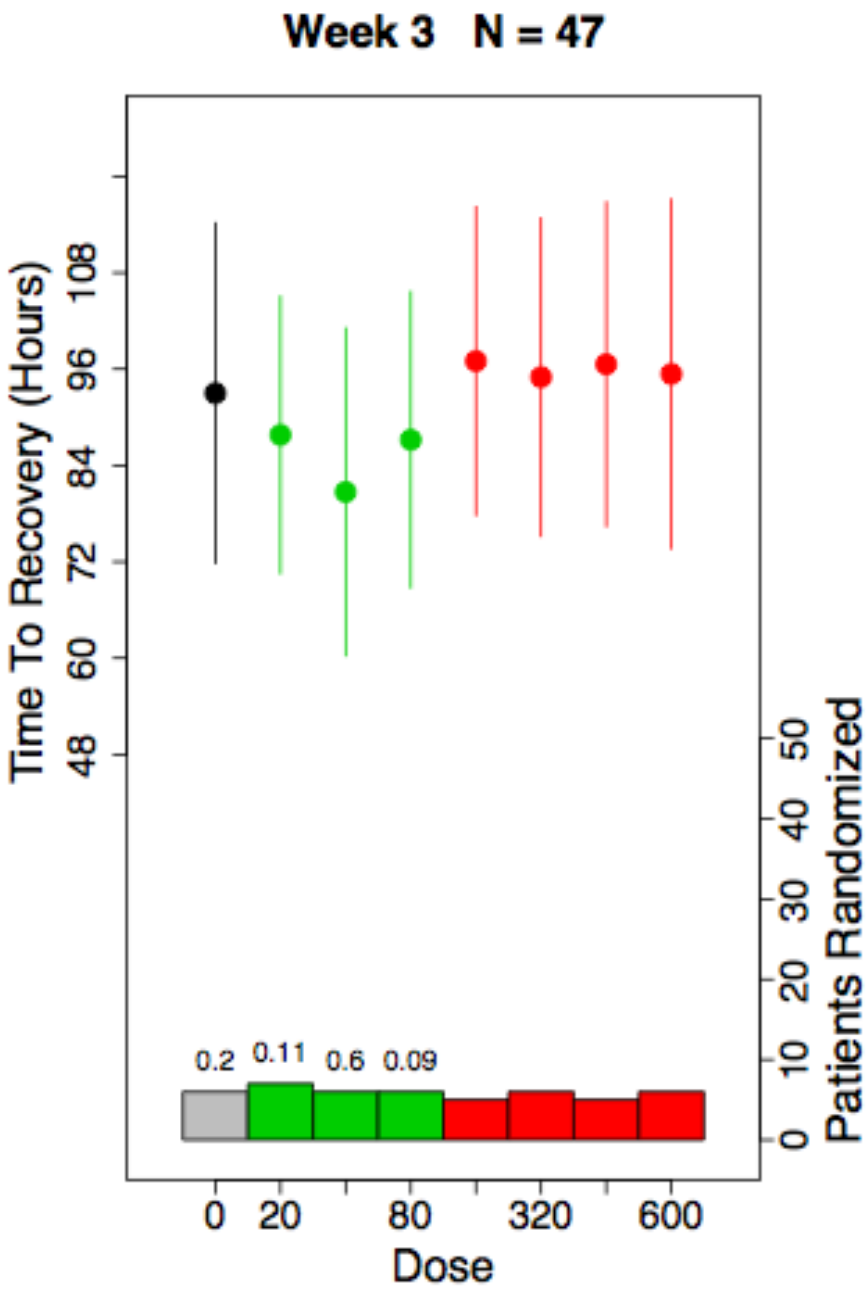
Decision = Continue
Keep sampling? YES

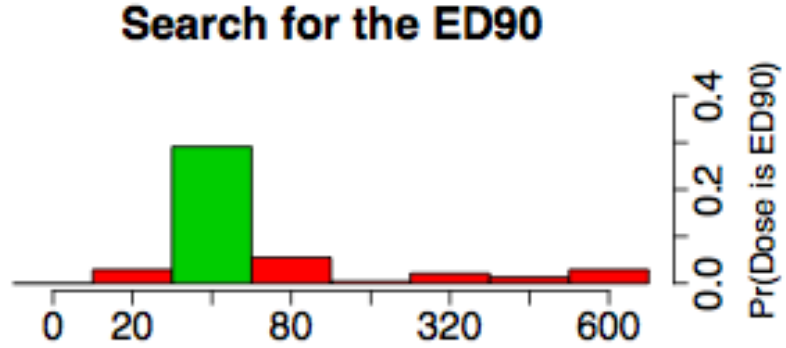

Week 2 N = 41
Time To Recovery (Hours)
48
60
72
84
96
108
0.2
0.11
0.6
0.08
0
20
80
320
600
Dose
0
10
20
30
40
50
Patients Randomized

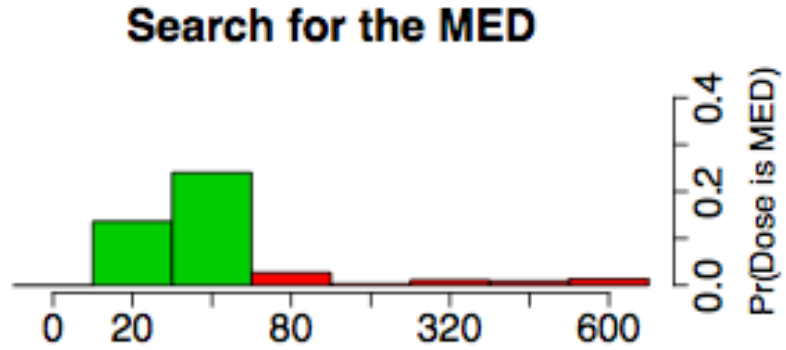

Search for the ED90
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is ED90)

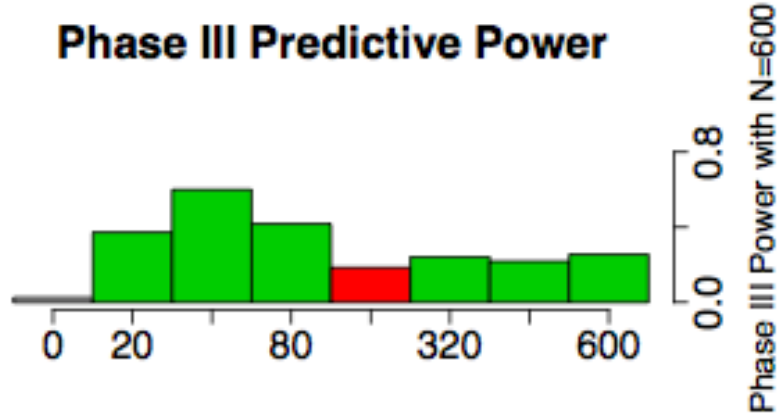

Search for the MED
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is MED)

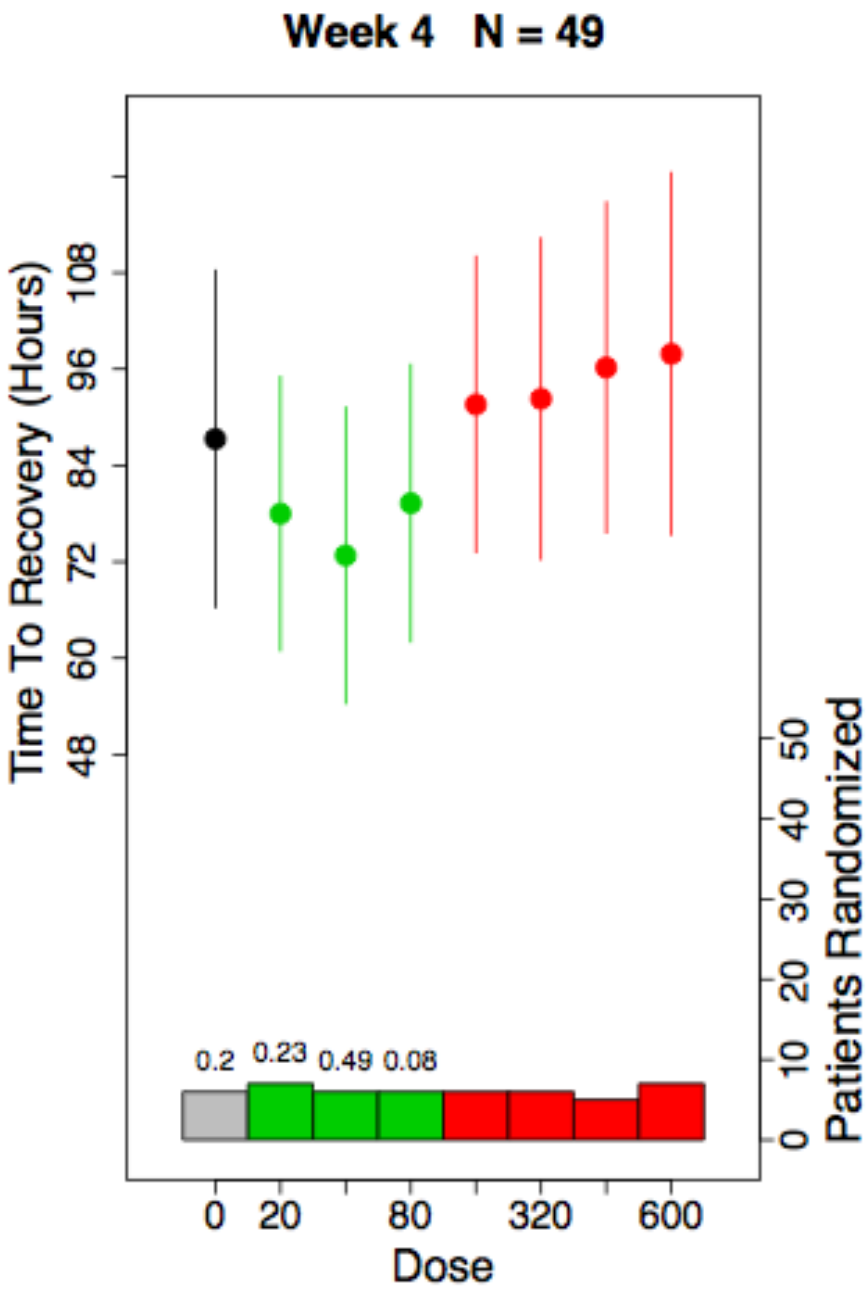

Phase III Predictive Power
0
20
80
320
600
0.0
0.8
Phase III Power with N=600

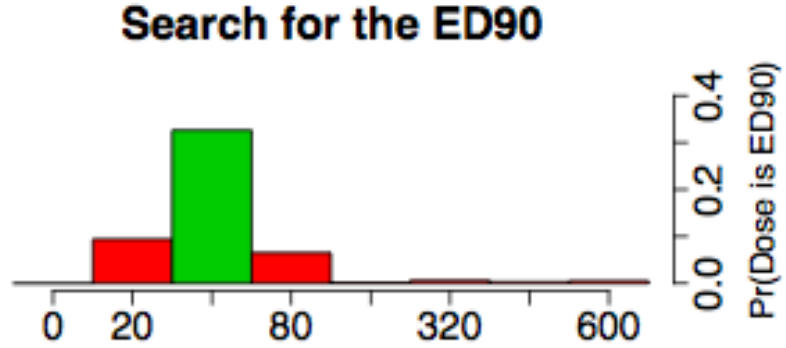

Week 3 N = 47
Time To Recovery (Hours)
48
60
72
84
96
108
0.2
0.11
0.6
0.09
0
20
80
320
600
Dose
0
10
20
30
40
50
Patients Randomized

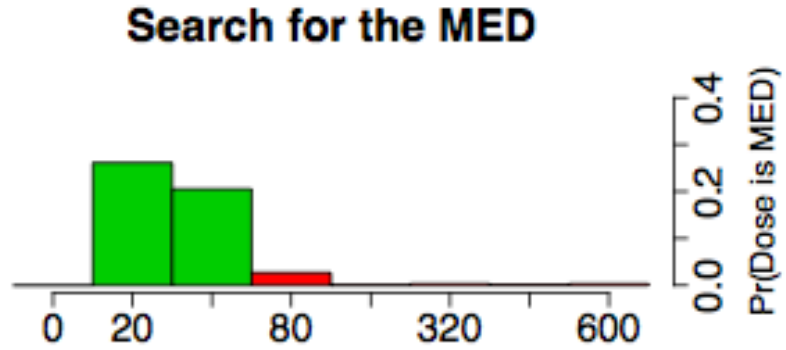

Search for the ED90
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is ED90)

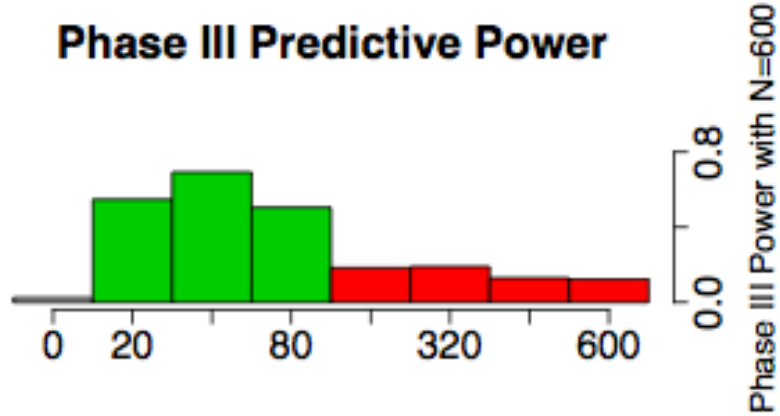

Search for the MED
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is MED)

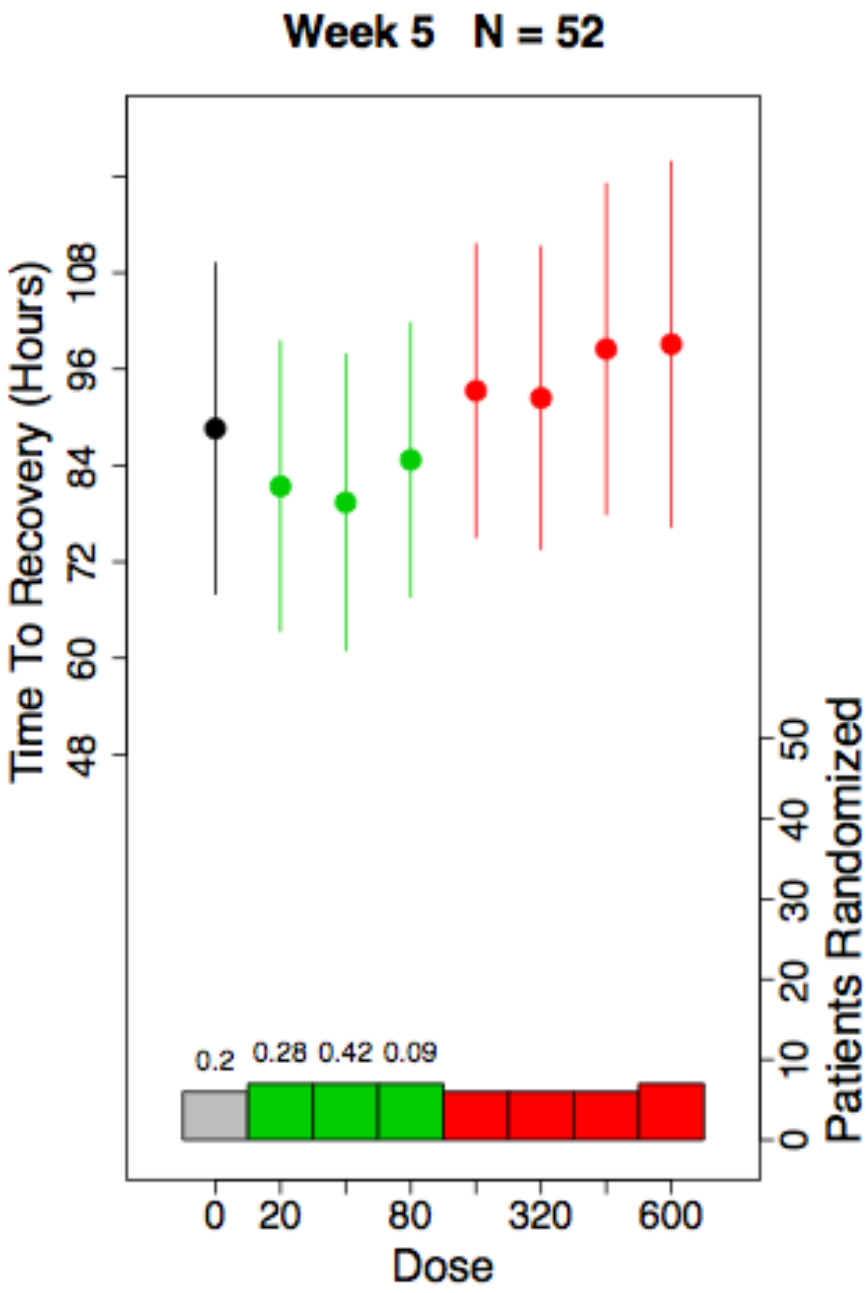

Phase III Predictive Power
0
20
80
320
600
0.0
0.8
Phase III Power with N=600

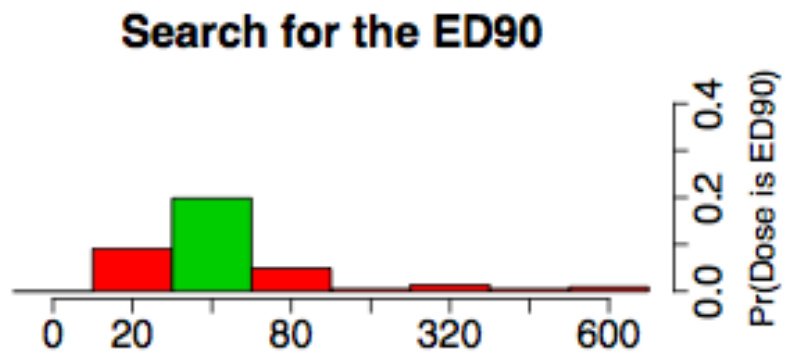

Week 4 N = 49
Time To Recovery (Hours)
48
60
72
84
96
108
0.2
0.23
0.49
0.08
0
20
80
320
600
Dose
0
10
20
30
40
50
Patients Randomized

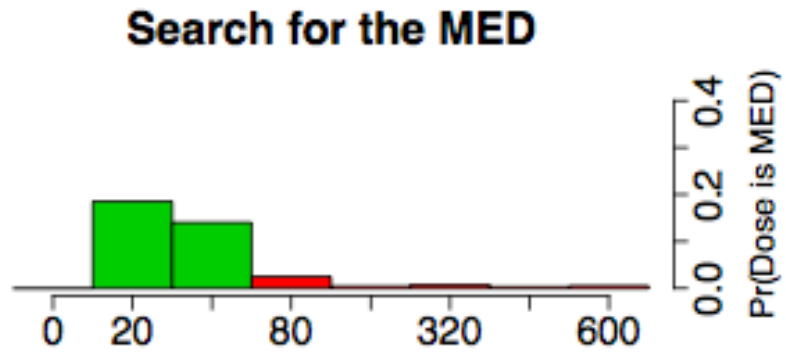

Search for the ED90
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is ED90)

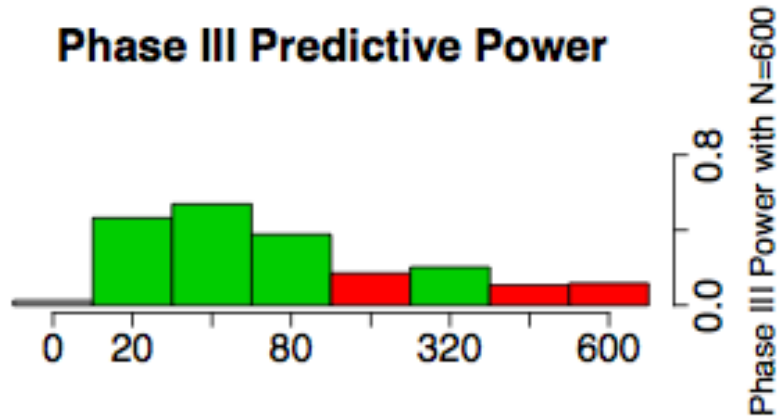

Search for the MED
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is MED)

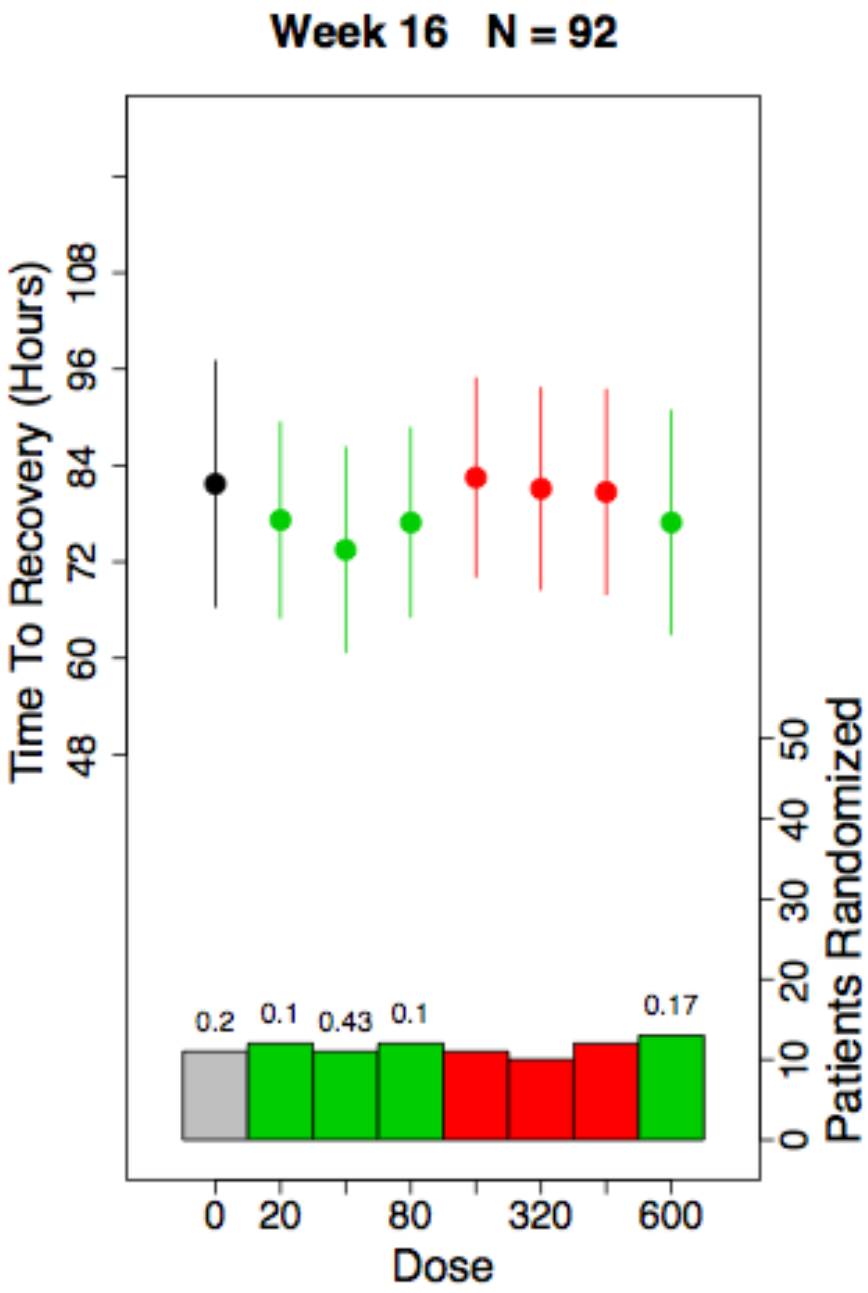

Phase III Predictive Power
0
20
80
320
600
0.0
0.8
Phase III Power with N=600

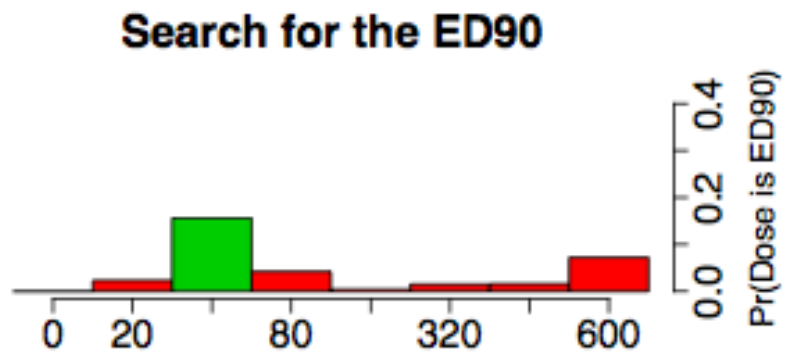

Week 5 N = 52
Time To Recovery (Hours)
48
60
72
84
96
108
0.2
0.28
0.42
0.09
0
20
80
320
600
Dose
0
10
20
30
40
50
Patients Randomized

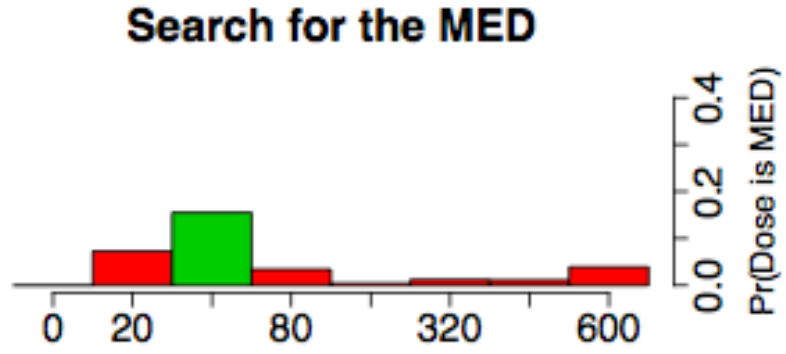

Search for the ED90
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is ED90)

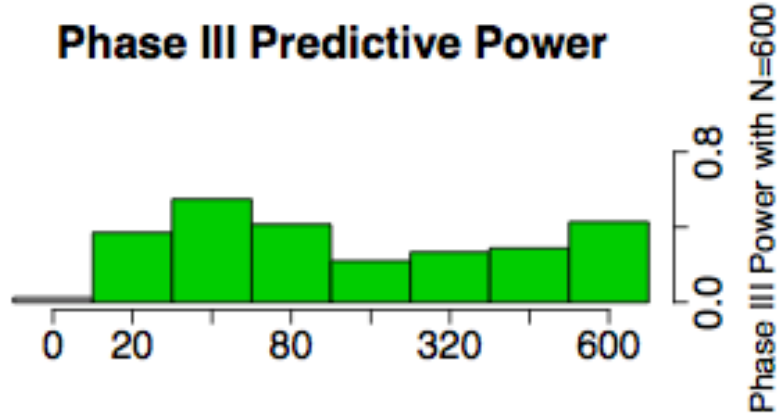

Search for the MED
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is MED)


Phase III Predictive Power
0
20
80
320
600
0.0
0.8
Phase III Power with N=600

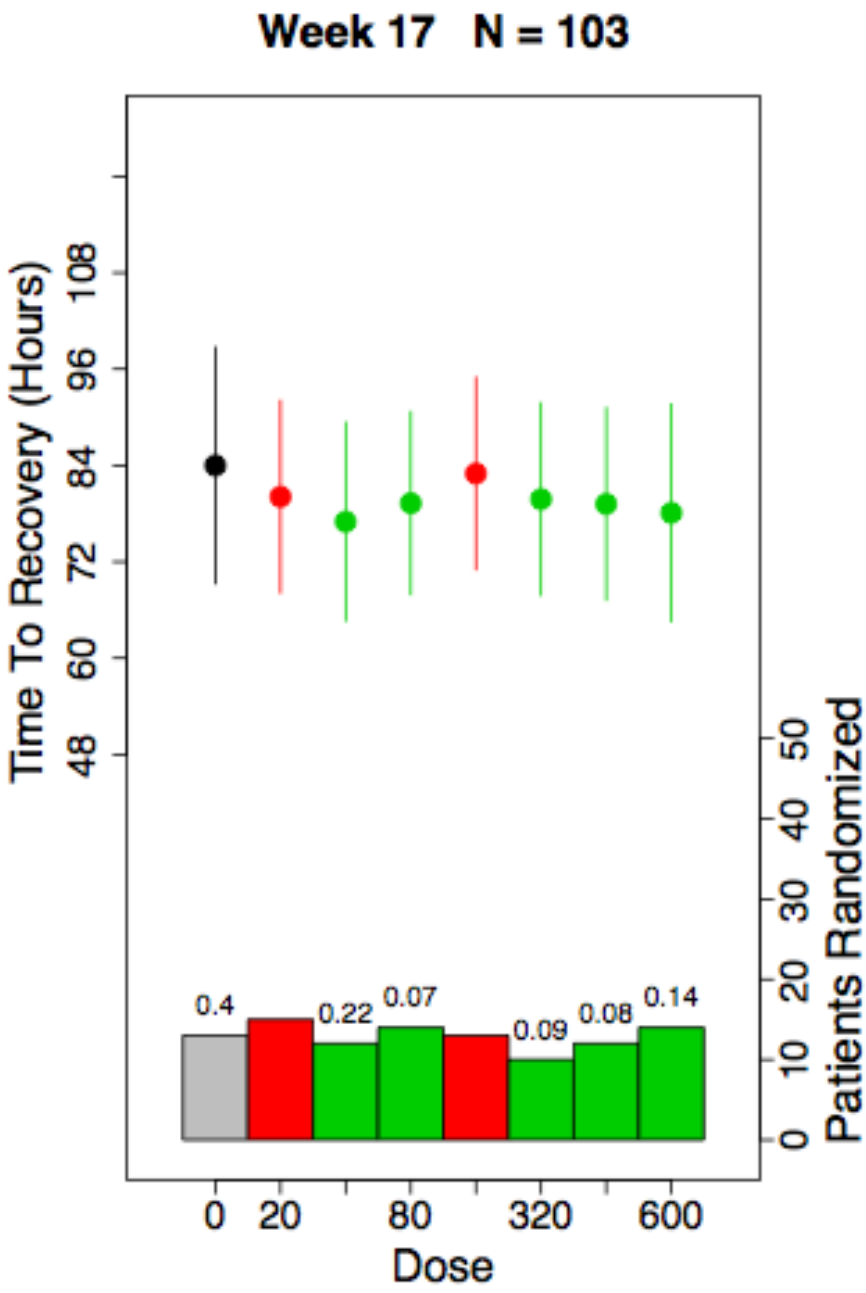

Week 16 N = 92
Time To Recovery (Hours)
48
60
72
84
96
108
0.2
0.1
0.43
0.1
0.17
0
20
80
320
600
Dose
0
10
20
30
40
50
Patients Randomized

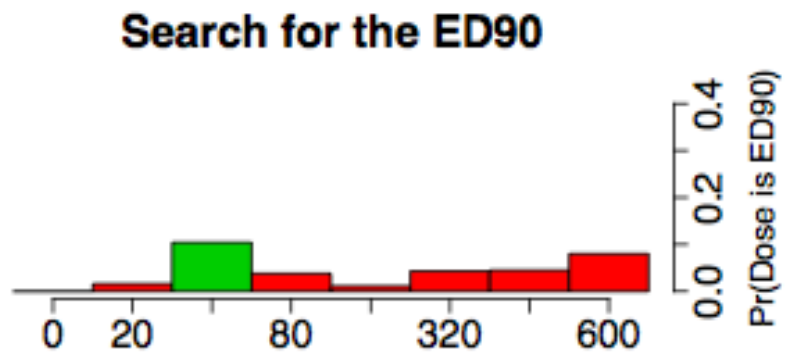

Search for the ED90
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is ED90)

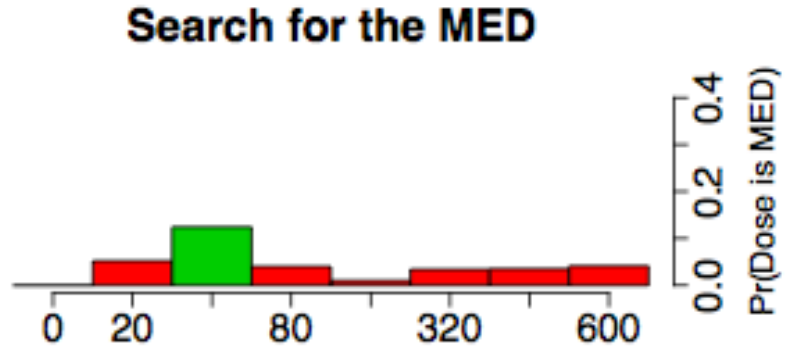

Search for the MED
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is MED)

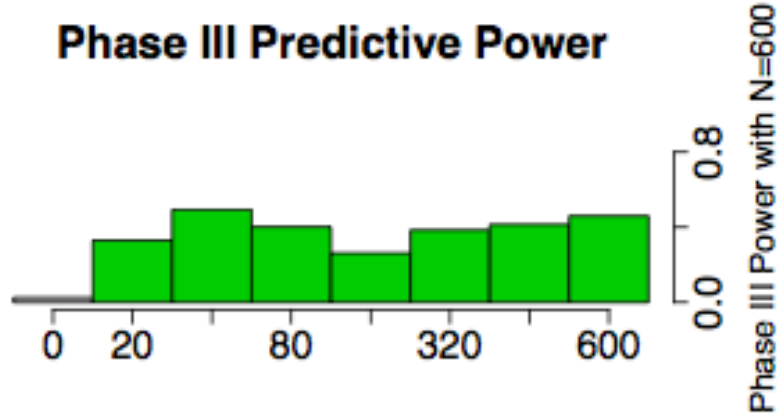

Phase III Predictive Power
0
20
80
320
600
0.0
0.8
Phase III Power with N=600

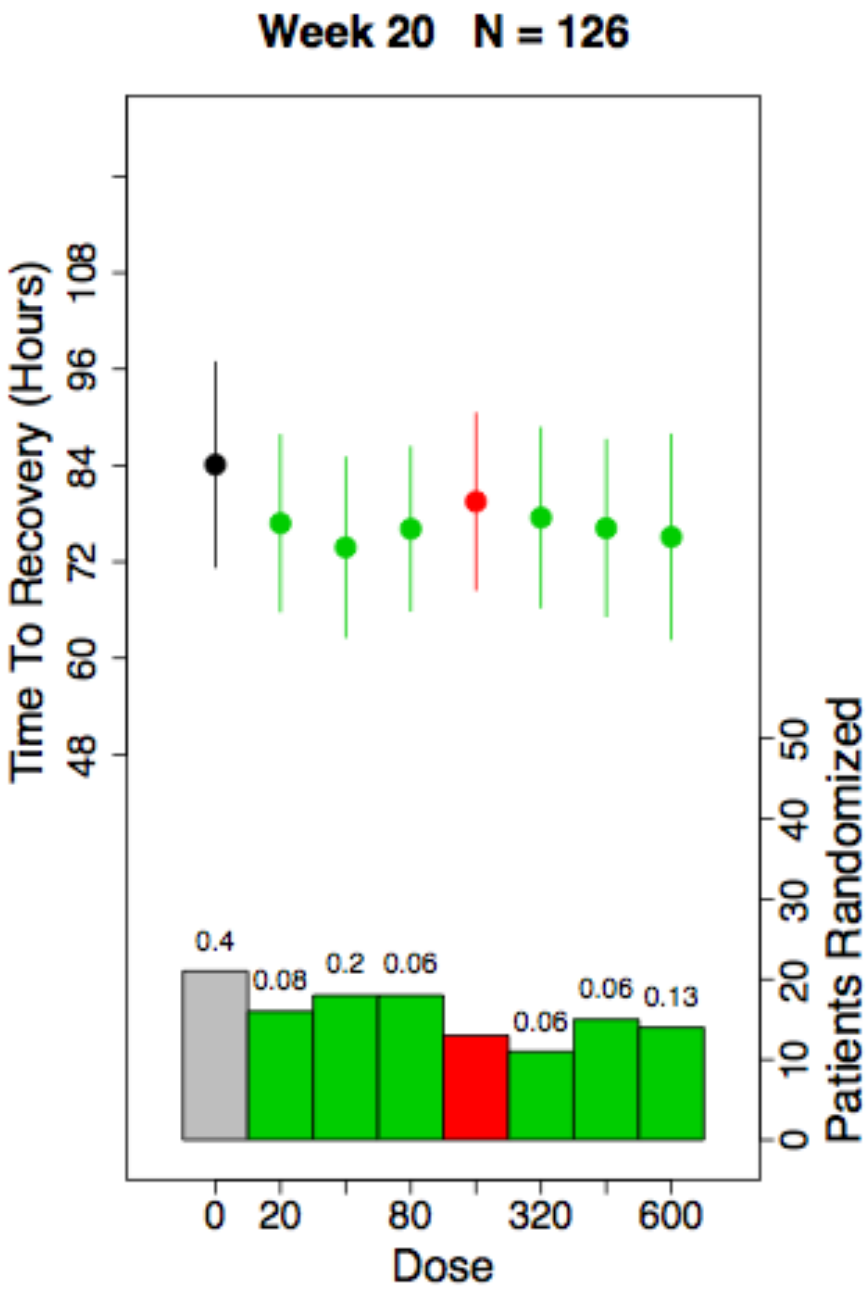
“Hmm, That’s odd”

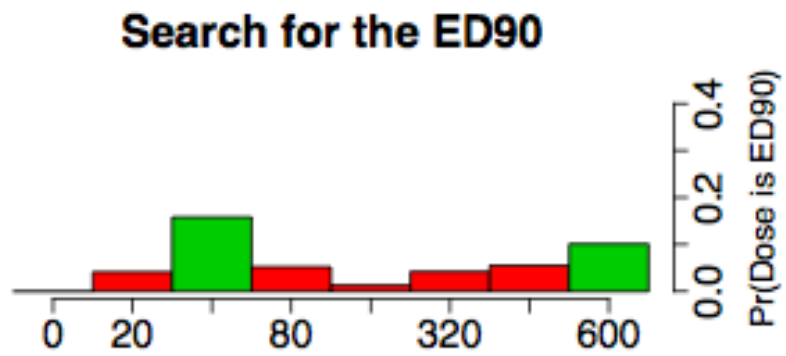

Week 17 N = 103
Time To Recovery (Hours)
48
60
72
84
96
108
0.4
0.22
0.07
0.09
0.08
0.14
0
20
80
320
600
Dose
0
10
20
30
40
50
Patients Randomized

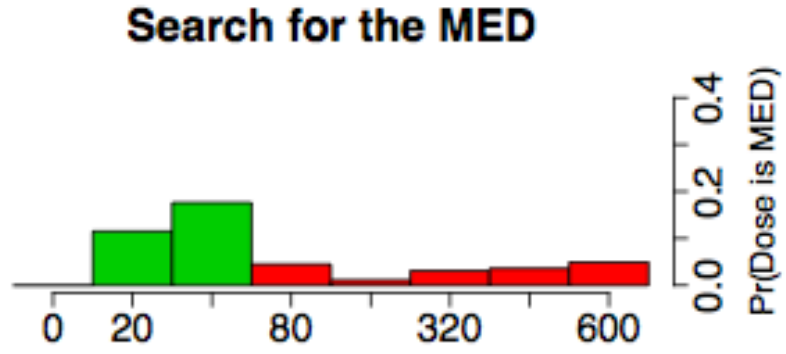

Search for the ED90
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is ED90)

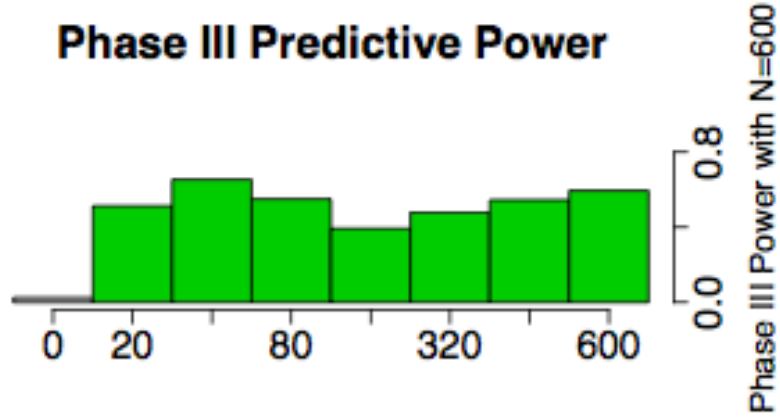

Search for the MED
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is MED)

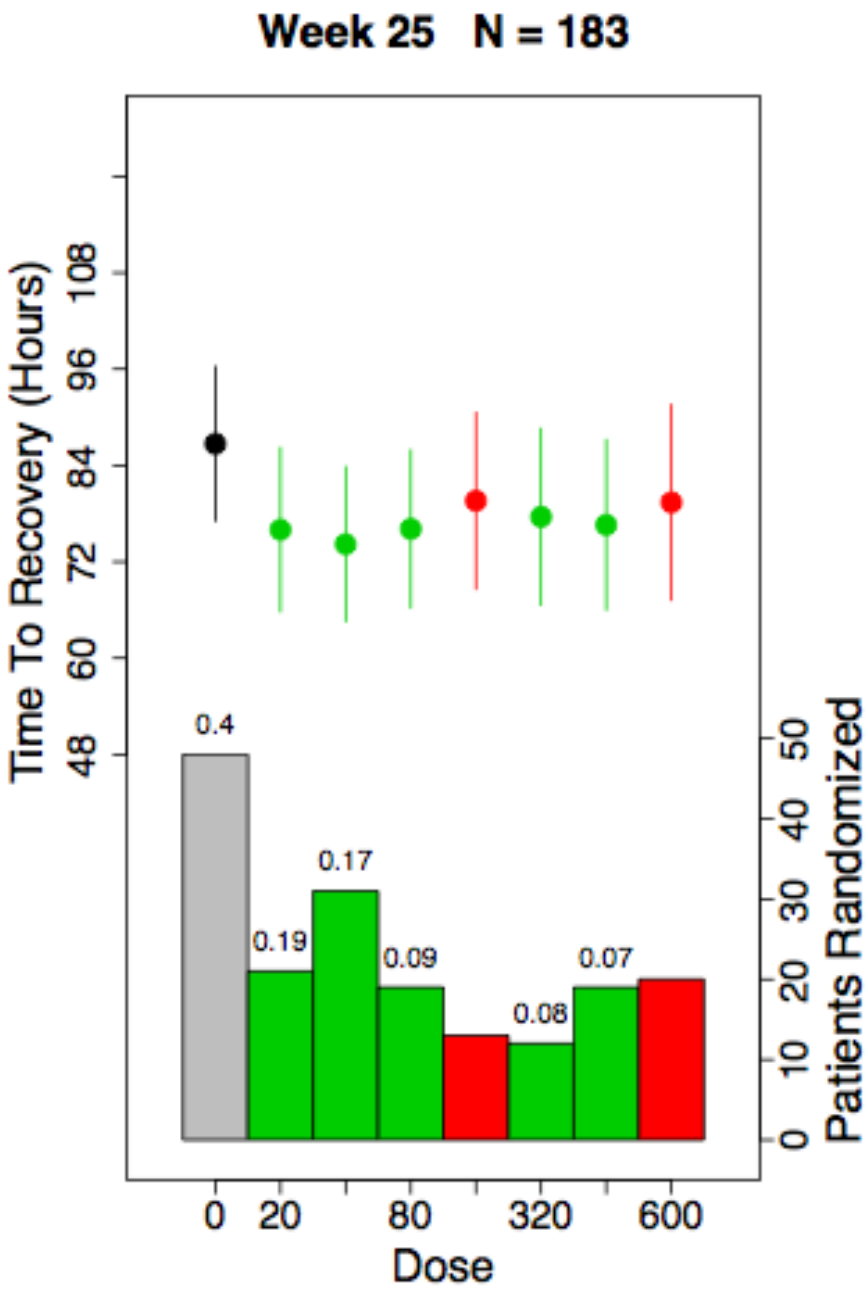

Phase III Predictive Power
0
20
80
320
600
0.0
0.8
Phase III Power with N=600

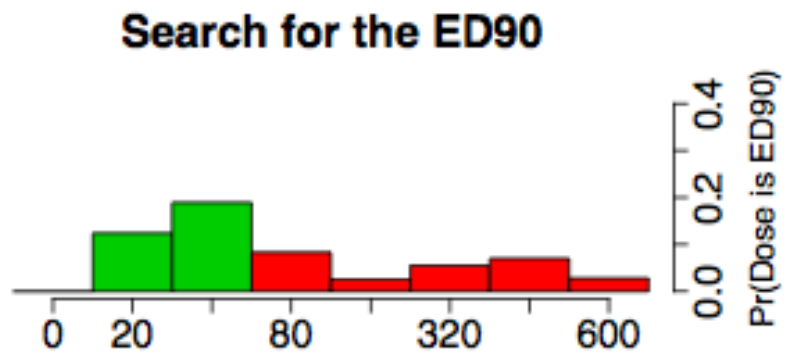

Week 20 N = 126
Time To Recovery (Hours)
48
60
72
84
96
108
0.4
0.08
0.2
0.06
0.06
0.06
0.13
0
20
80
320
600
Dose
0
10
20
30
40
50
Patients Randomized

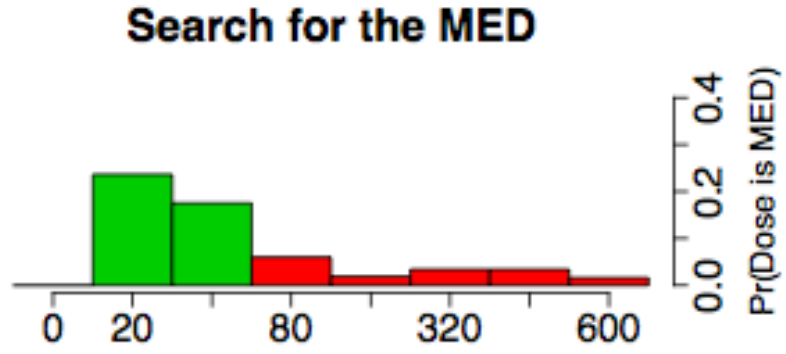

Search for the ED90
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is ED90)

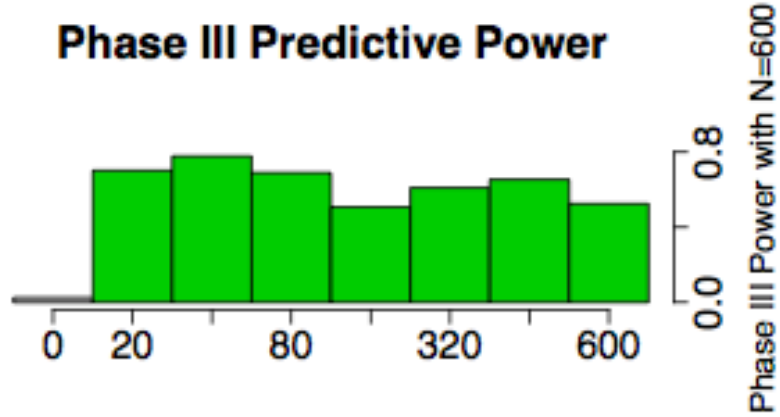

Search for the MED
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is MED)

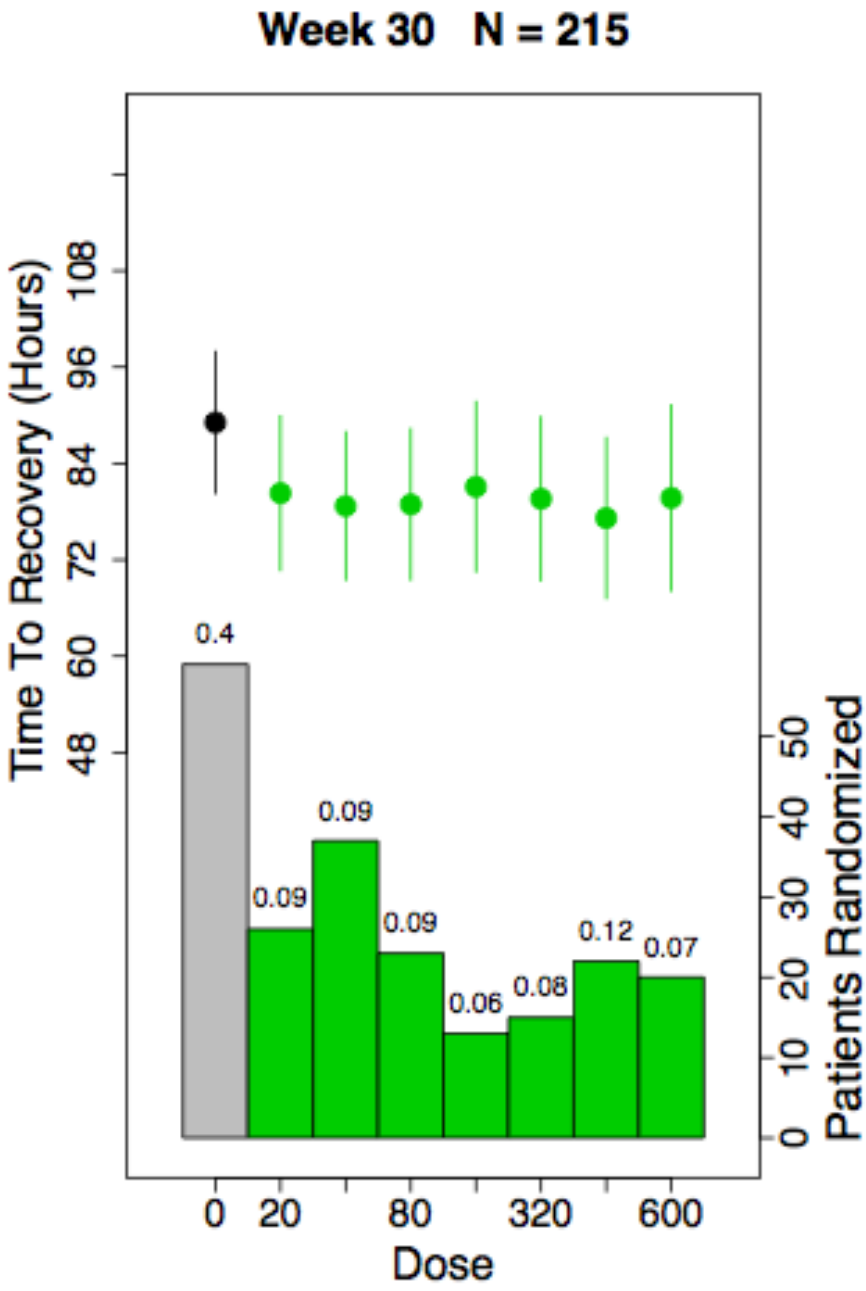

Phase III Predictive Power
0
20
80
320
600
0.0
0.8
Phase III Power with N=600

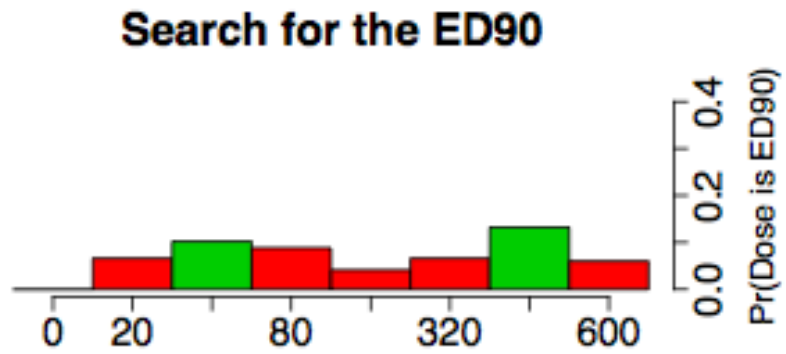

Week 25 N = 183
Time To Recovery (Hours)
48
60
72
84
96
108
0.4
0.19
0.17
0.09
0.08
0.07
0
20
80
320
600
Dose
0
10
20
30
40
50
Patients Randomized

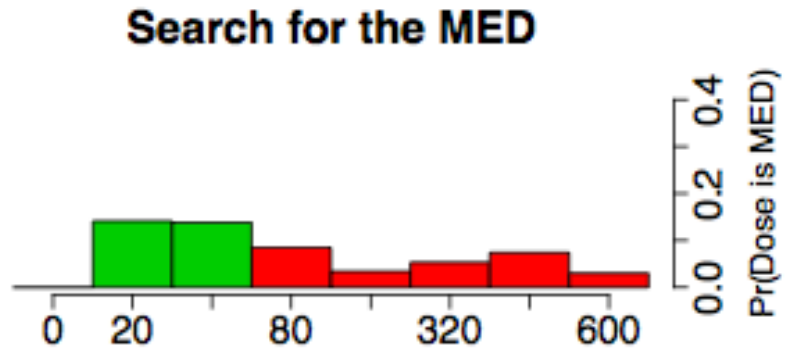

Search for the ED90
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is ED90)

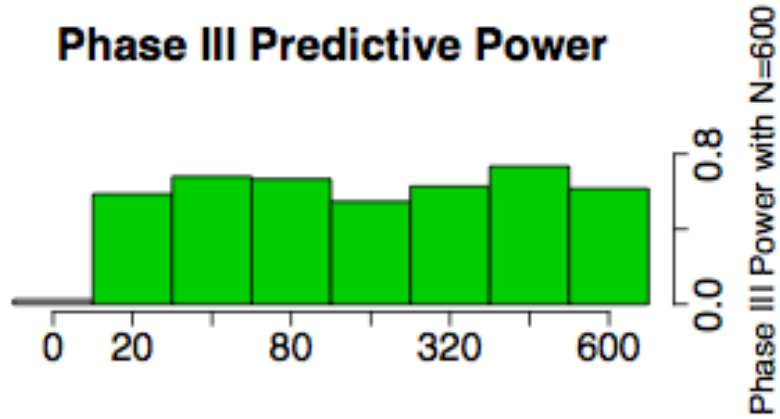

Search for the MED
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is MED)

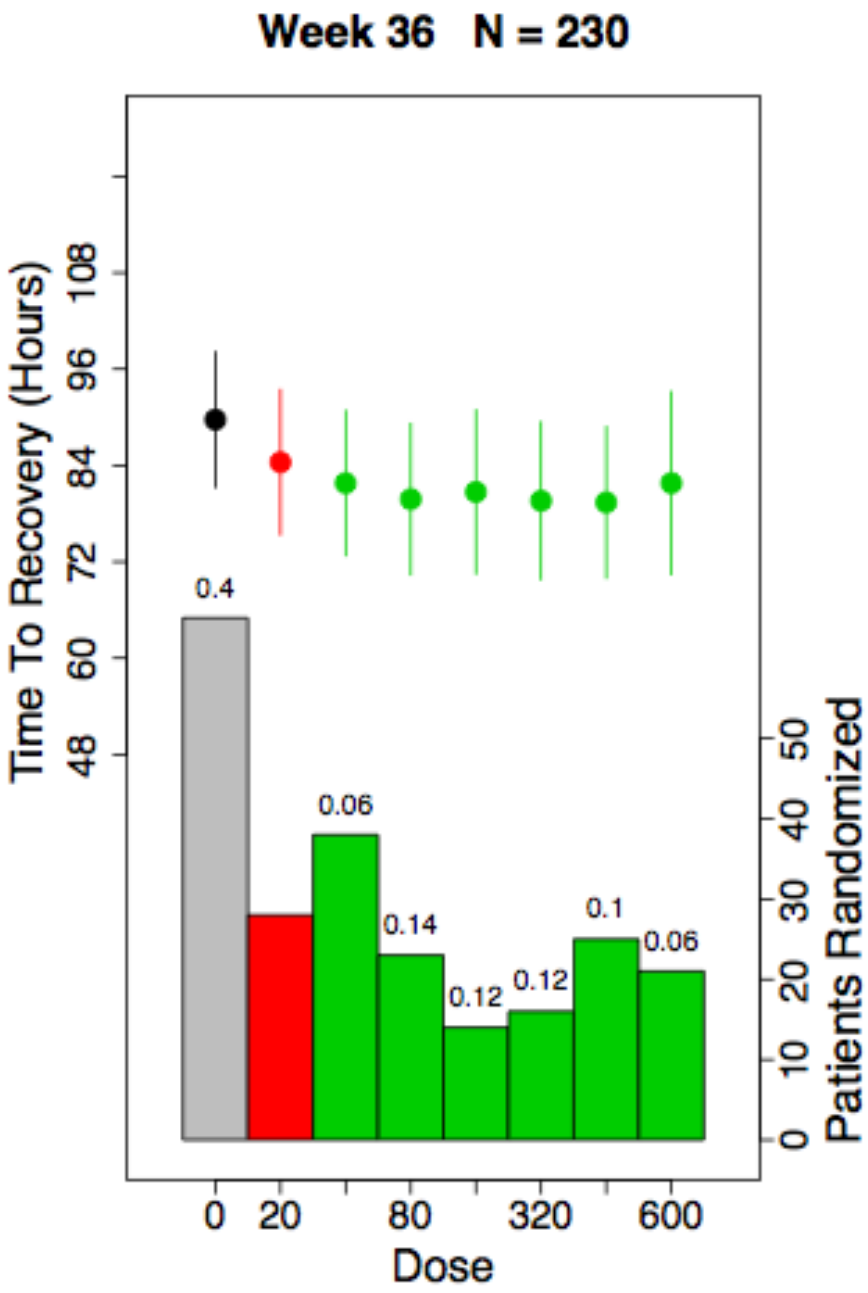

Phase III Predictive Power
0
20
80
320
600
0.0
0.8
Phase III Power with N=600

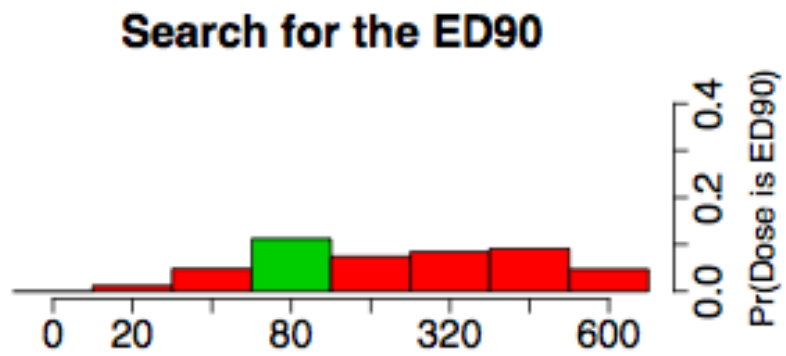

Week 30 N = 215
Time To Recovery (Hours)
48
60
72
84
96
108
0.4
0.09
0.09
0.09
0.06
0.08
0.12
0.07
0
10
20
30
40
50
Patients Randomized
0
20
80
320
600
Dose

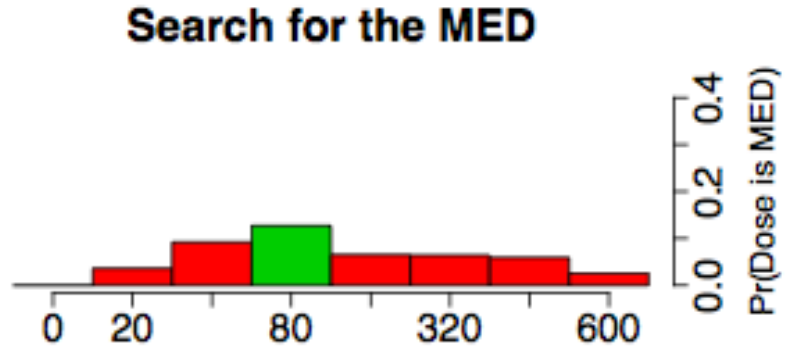

Search for the ED90
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is ED90)

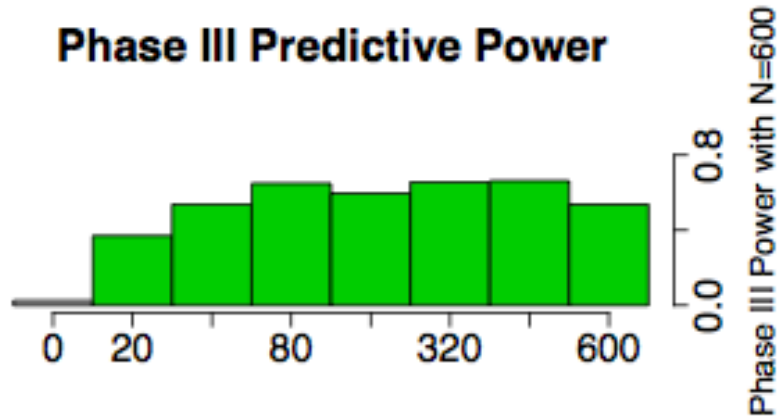

Search for the MED
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is MED)

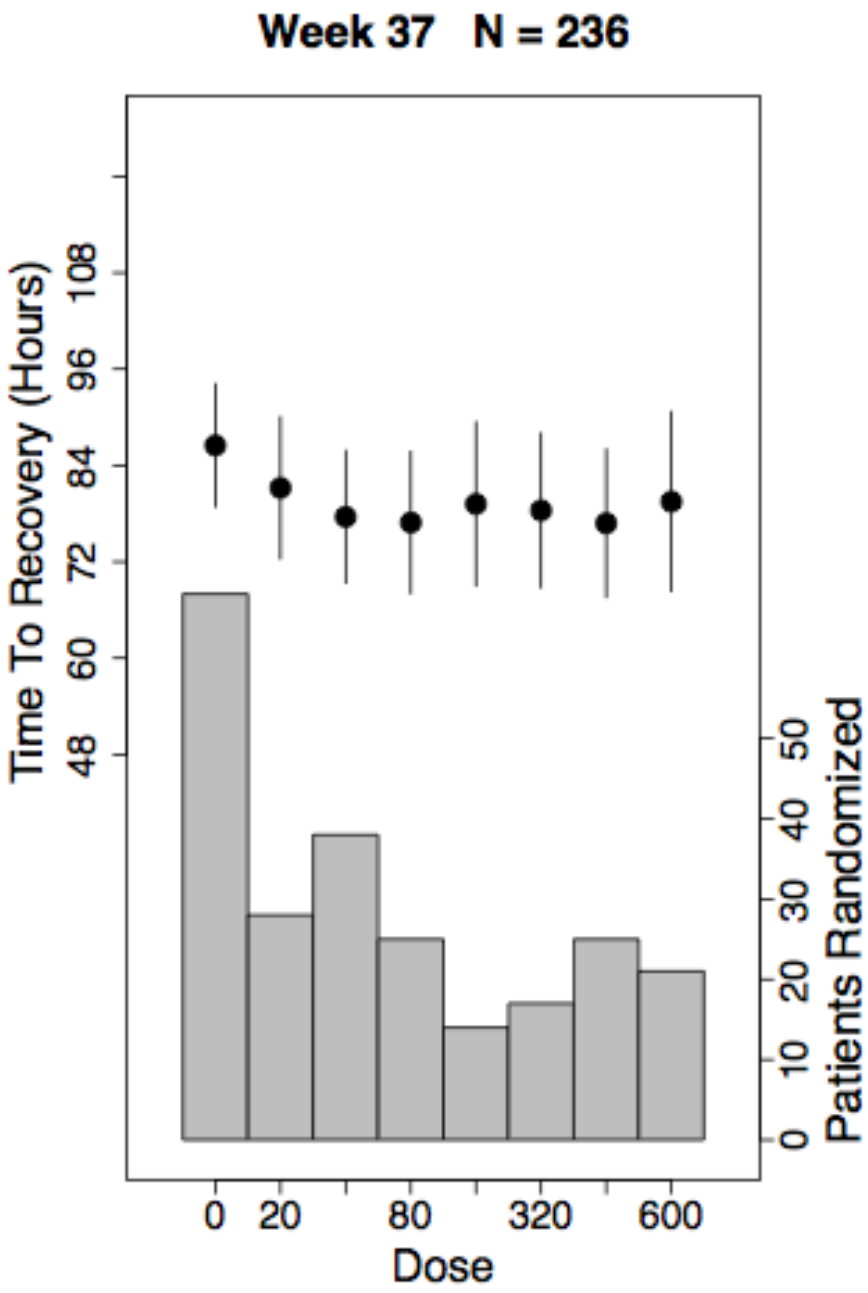

Phase III Predictive Power
0
20
80
320
600
0.0
0.8
Phase III Power with N=600

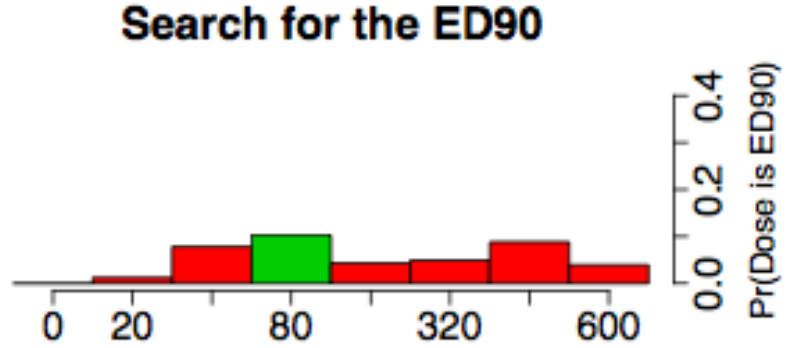

Week 36 N = 230
Time To Recovery (Hours)
48
60
72
84
96
108
0.4
0.06
0.14
0.12
0.12
0.1
0.06
0
10
20
30
40
50
Patients Randomized
0
20
80
320
600
Dose

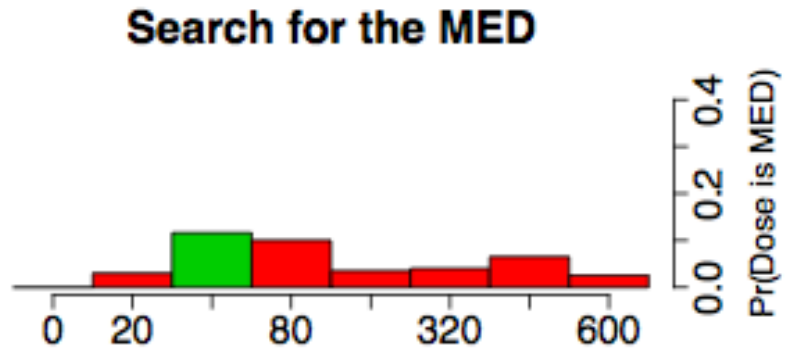

Search for the ED90
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is ED90)

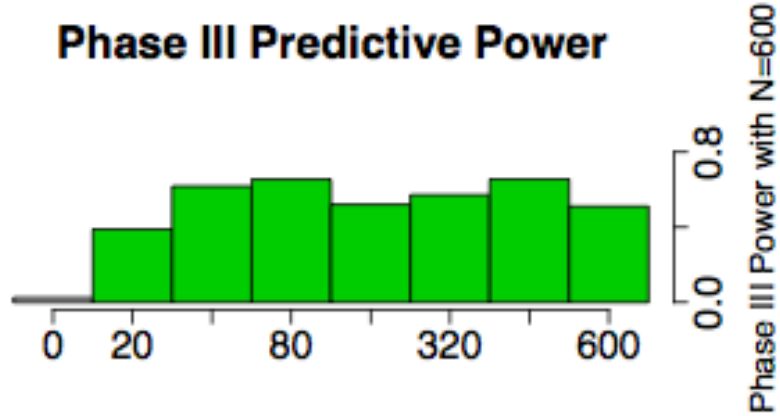

Search for the MED
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is MED)

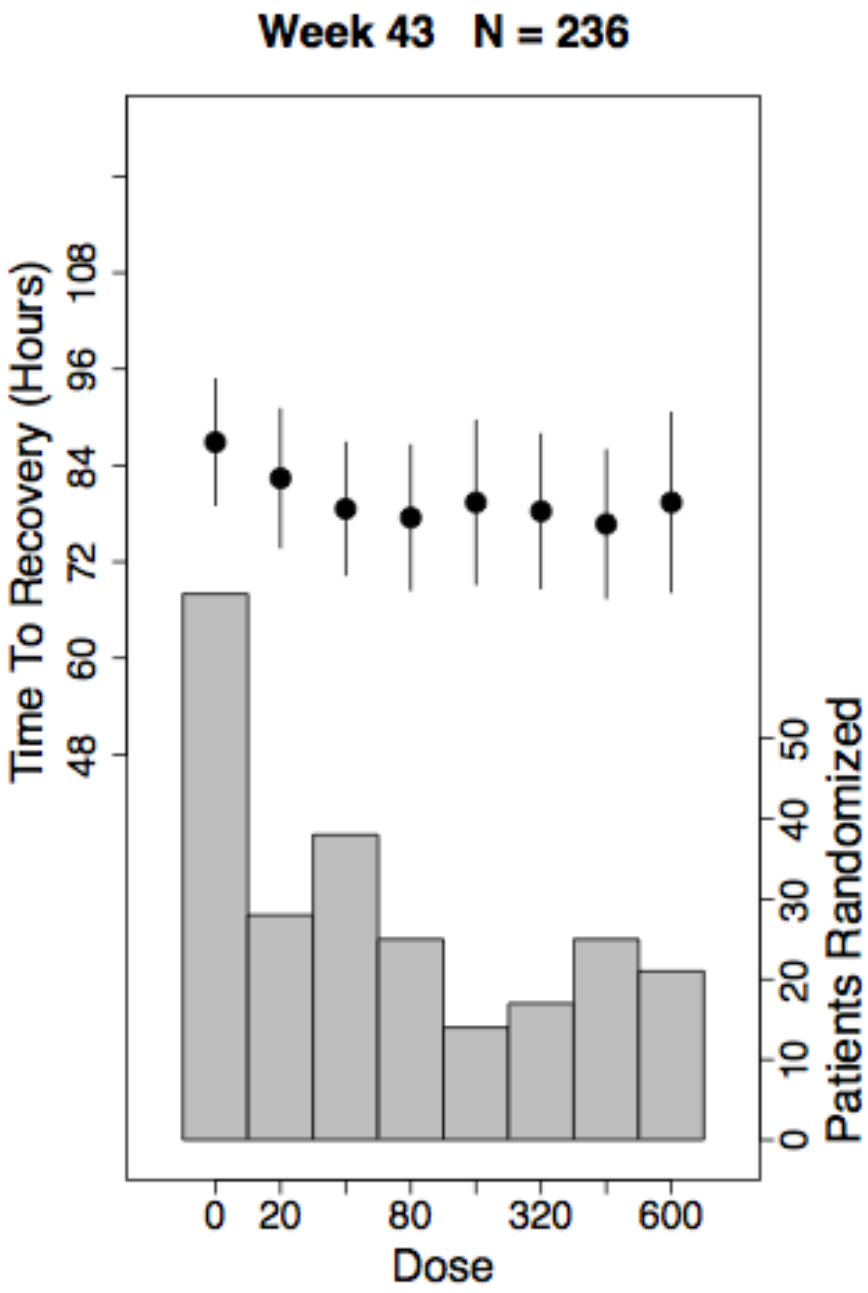

Phase III Predictive Power
0
20
80
320
600
0.0
0.8
Phase III Power with N=600

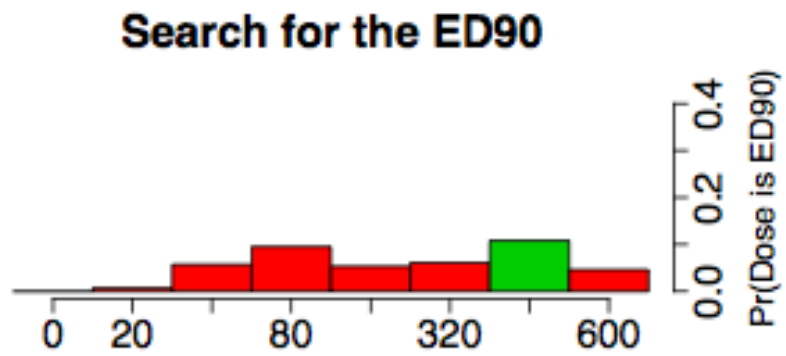

Week 37 N = 236
Time To Recovery (Hours)
48
60
72
84
96
108
0
10
20
30
40
50
Patients Randomized
0
20
80
320
600
Dose

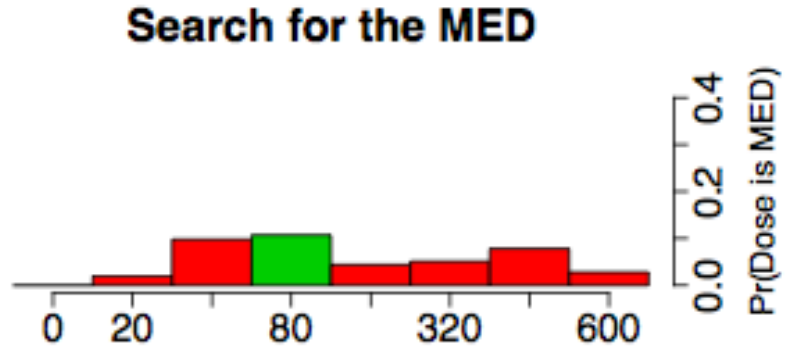

Search for the ED90
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is ED90)

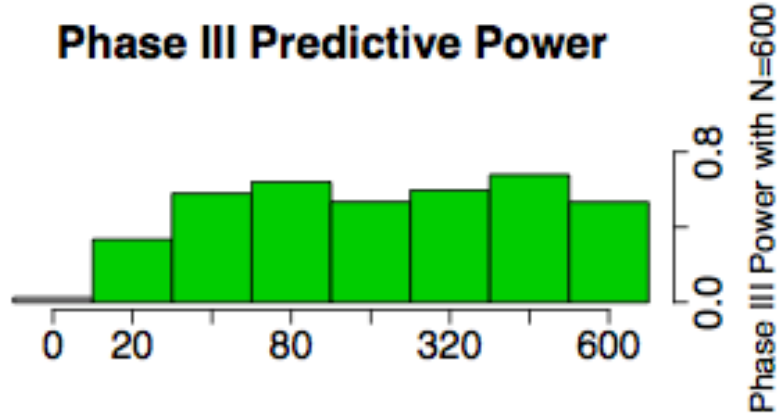

Search for the MED
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is MED)

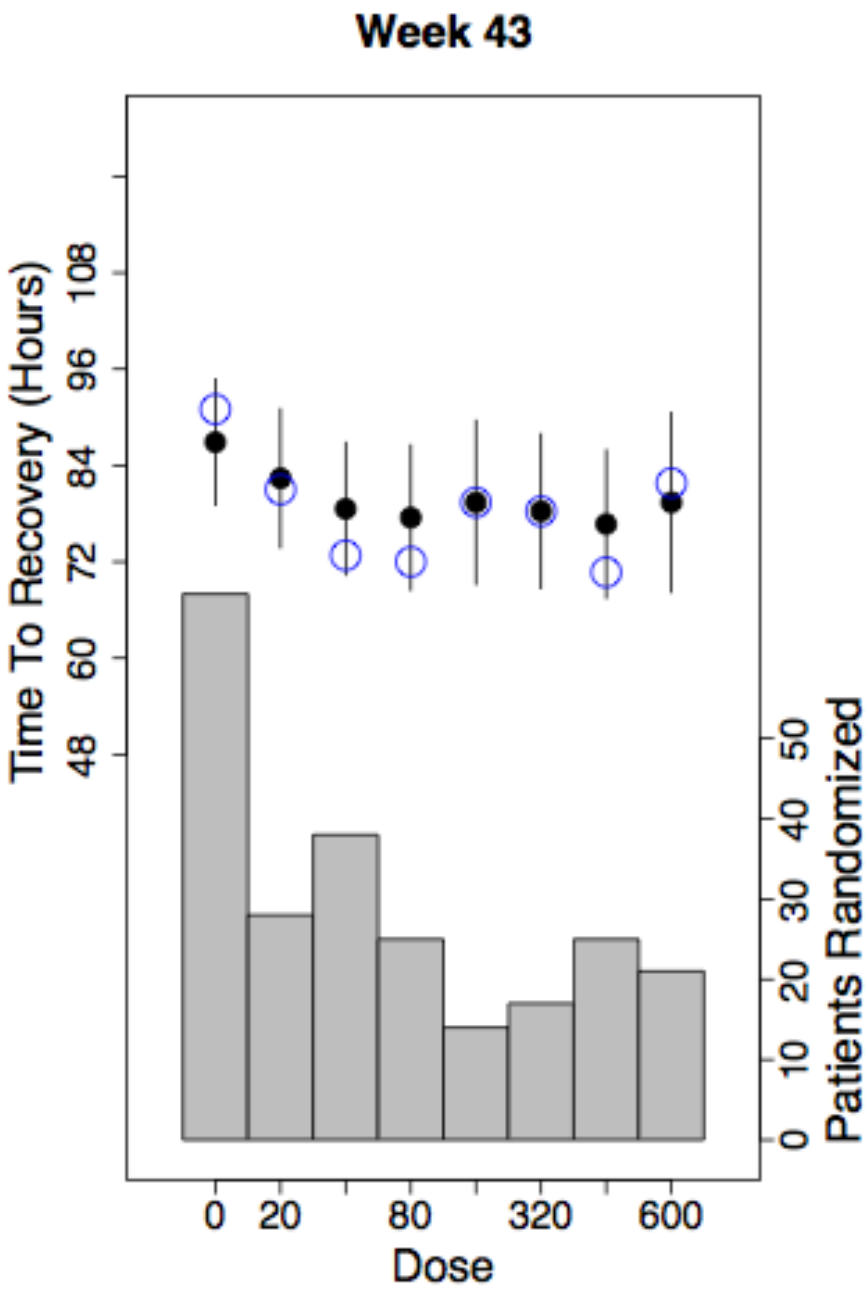

Phase III Predictive Power
0
20
80
320
600
0.0
0.8
Phase III Power with N=600

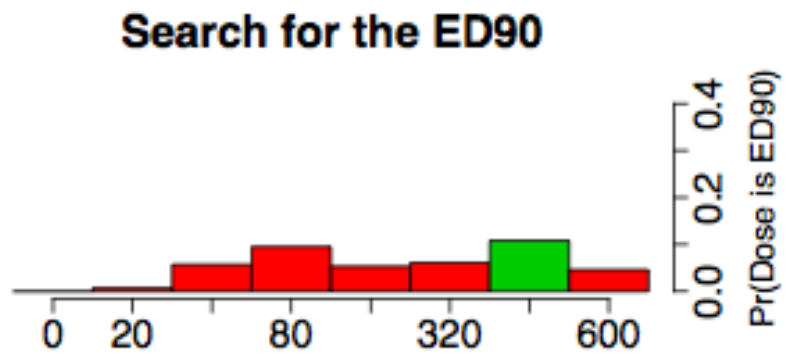

Week 43 N = 236
Time To Recovery (Hours)
48
60
72
84
96
108
0
10
20
30
40
50
Patients Randomized
0
20
80
320
600
Dose

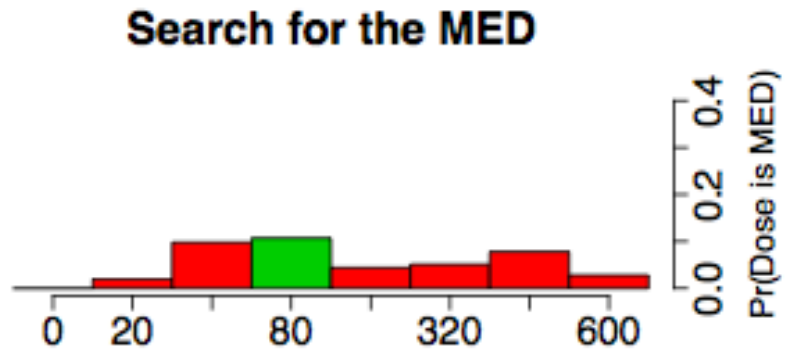

Search for the ED90
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is ED90)

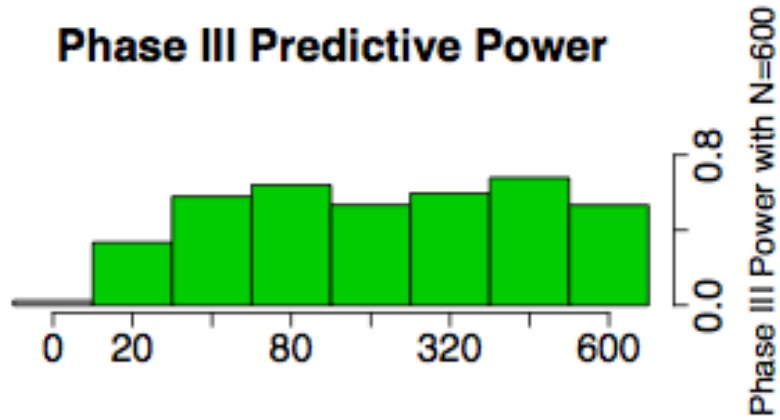

Search for the MED
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is MED)

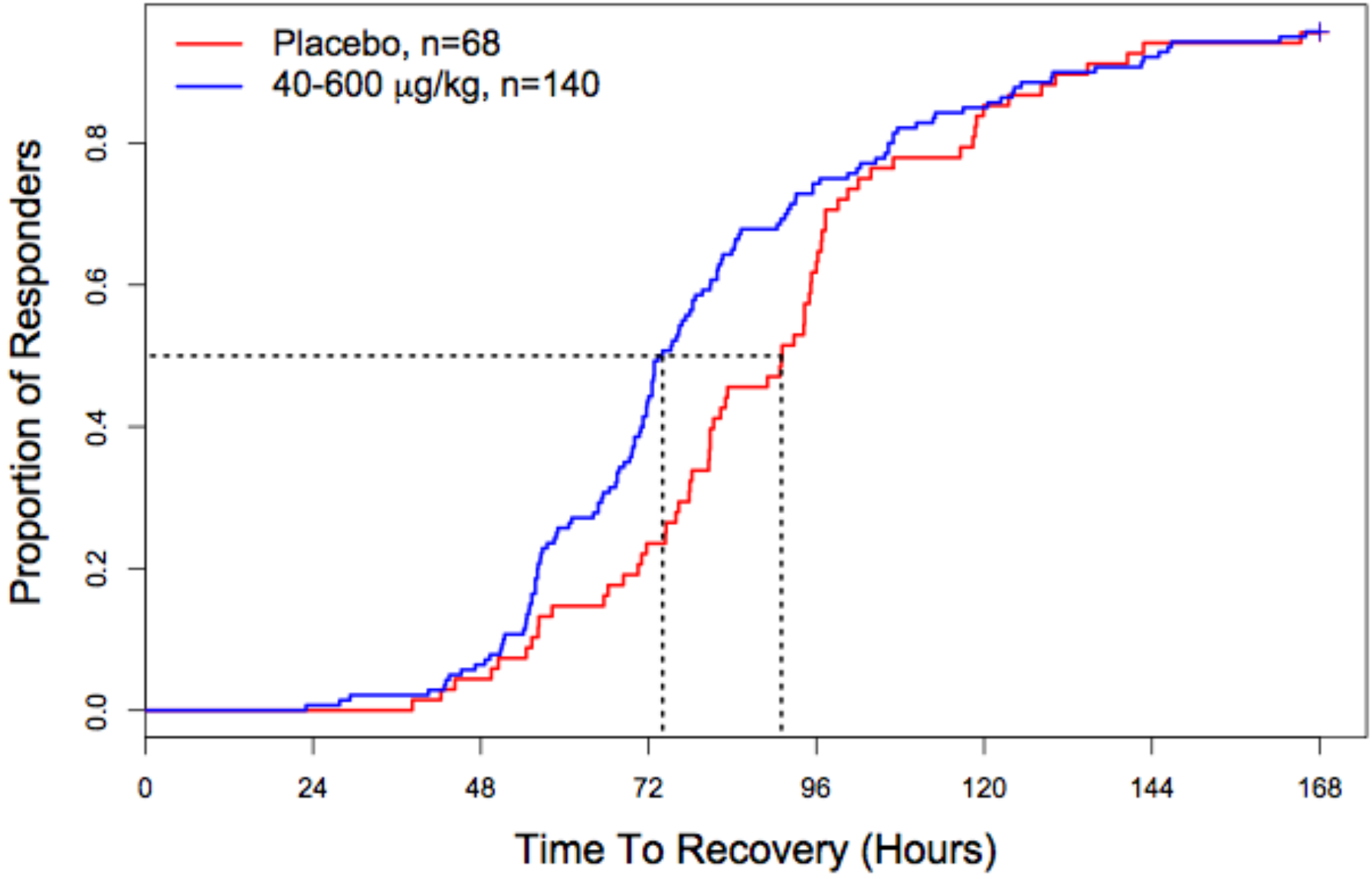

Phase III Predictive Power
0
20
80
320
600
0.0
0.8
Phase III Power with N=600


Week 43
Time To Recovery (Hours)
48
60
72
84
96
108
0
10
20
30
40
50
Patients Randomized
0
20
80
320
600
Dose


Search for the ED90
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is ED90)


Search for the MED
0
20
80
320
600
0.0
0.2
0.4
Pr(Dose is MED)


Phase III Predictive Power
0
20
80
320
600
0.0
0.8
Phase III Power with N=600


Placebo, n=68
40-600 µg/kg, n=140
Proportion of Responders
0.0
0.2
0.4
0.6
0.8
0
24
48
72
96
120
144
168
Time To Recovery (Hours)

Please find below (and attached) the latest news from Tranzyme Pharma.
We thank you for your continued interest.

PRESS RELEASE

For Immediate Release – October 1, 2008

Tranzyme Pharma Announces Positive Phase IIb Results with Its Ghrelin Agonist, TZP-101, for Postoperative Ileus (POI)

Phase III Initiation Targeted Q1 2009

RESEARCH TRIANGLE PARK, N.C. (October 1, 2008) - Tranzyme Pharma today announced positive Phase IIb results for its first-in-class, highly potent and selective ghrelin agonist, TZP-101, for the management of postoperative ileus (POI). Results demonstrated that TZP-101 was both safe and highly effective in reducing the duration of ileus following surgery in patients undergoing open bowel resection.

Over 200 patients were enrolled in an adaptive, multinational, double-blind, placebo-controlled Phase IIb clinical trial designed to assess the time to recovery of gastrointestinal (GI) function. Either TZP-101 or placebo was administered intravenously within the first hour after surgery, followed by once-daily dosing for up to seven days. The primary study endpoint was time to first bowel movement (BM), also known as "GI1". Given its pharmacoeconomic importance, a key secondary endpoint was the percentage of patients that achieved GI recovery within 72 hours of surgery.

Phase 2 Dose Finding Trials

- Target randomization to therapies performing best
- Target randomization to doses that will provide the most statistical information
- Perform pre-defined analyses early & often
- Need infrastructure in place to ensure all components work smoothly
  - Monitor! Monitor! Monitor!

Phase 3 / Confirmatory Trials

- CDER / CBER: Phase 3
- CDRH: Confirmatory
- The final test before market
- Control of Type I error rate very important
- Tend not to adaptively randomize
  - usually two-arm trial (no benefit to power to do adaptive randomization with two arms)
  - fear of drift

What is different about confirmatory trials?

- Type I error is dominant factor
- Adjustments to the design in order to accommodate adaptive aspects must still control type I error
- Predictive probabilities much more relevant than posterior probabilities for making adaptive decisions
- A very well-defined goal
  - a "game" you win or lose

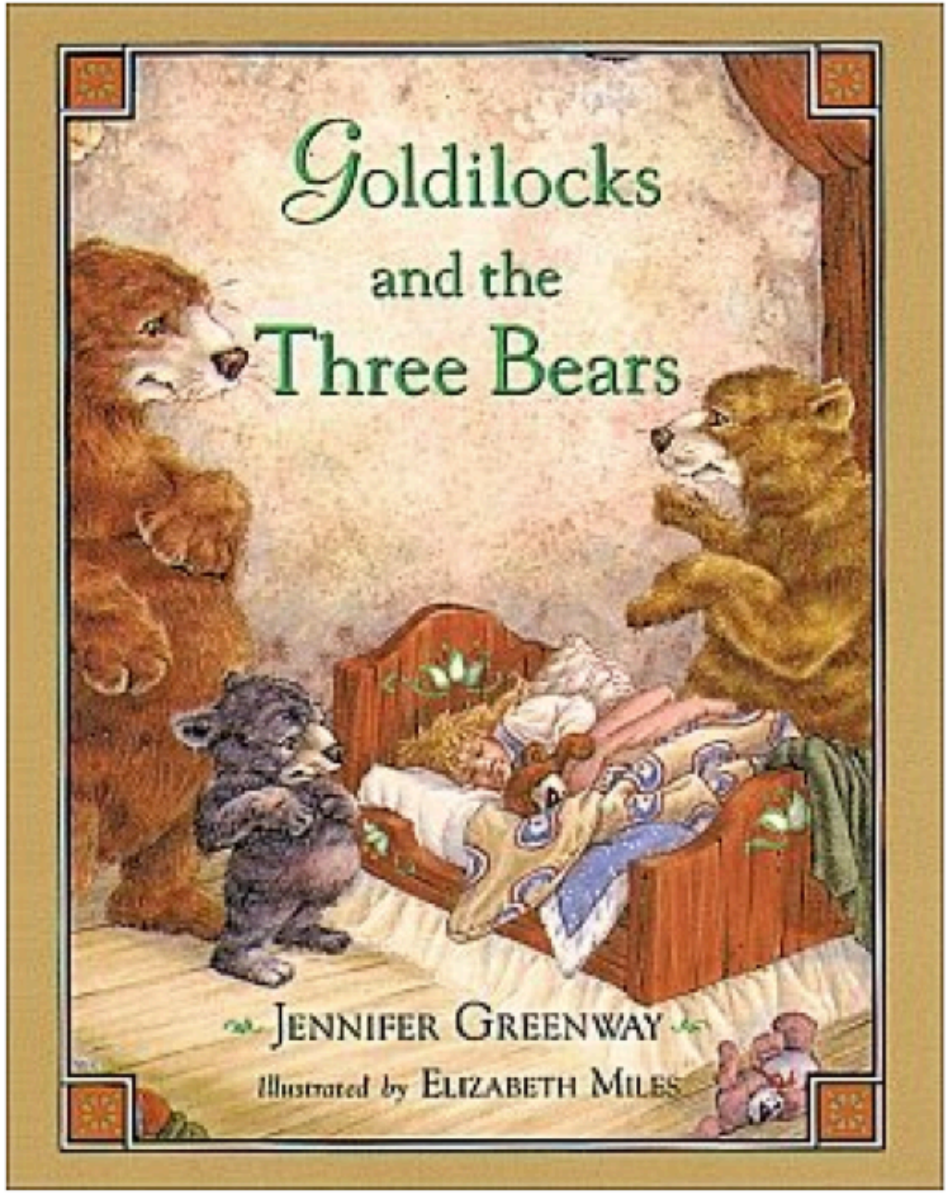
Posterior vs. Predictive

- *Posterior probability*
  - tells you something about the drug
  - how likely is it that the response rate is greater than 50%?

- *Predictive probability*
  - tells you something about the ability of the drug to *accomplish a task*
  - how likely is it that the drug can *win this trial*?

Porridge: Not too hot,
or too cold,
but *just right*

Trial: Not too big,
too small,
But *just right*



Journal of Biopharmaceutical Statistics, 24: 685–705, 2014

ISSN: 1054-3406 print/1520-5711 online
DOI: 10.1080/10543406.2014.888569

Taylor & Francis
Taylor & Francis Group

NOT TOO BIG, NOT TOO SMALL: A GOLDILOCKS APPROACH TO SAMPLE SIZE SELECTION

Kristine R. Broglio[1], Jason T. Connor[1,2], and Scott M. Berry[1]

[1]*Berry Consultants, LLC, Austin, Texas, USA*
[2]*University of Central Florida, Orlando, Florida, USA*

We present a Bayesian adaptive design for a confirmatory trial to select a trial's sample size based on accumulating data. During accrual, frequent sample size selection analyses are made and predictive probabilities are used to determine whether the current sample size is sufficient or whether continuing accrual would be futile. The algorithm explicitly accounts for complete follow-up of all patients before the primary analysis is conducted. We refer to this as a Goldilocks trial design, as it is constantly asking the question, "Is the sample size too big, too small, or just right?" We describe the adaptive sample size algorithm, describe how the design parameters should be chosen, and show examples for dichotomous and time-to-event endpoints.

Key Words: Bayesian adaptive trial design; Predictive probabilities; Sample size; Sequential design.

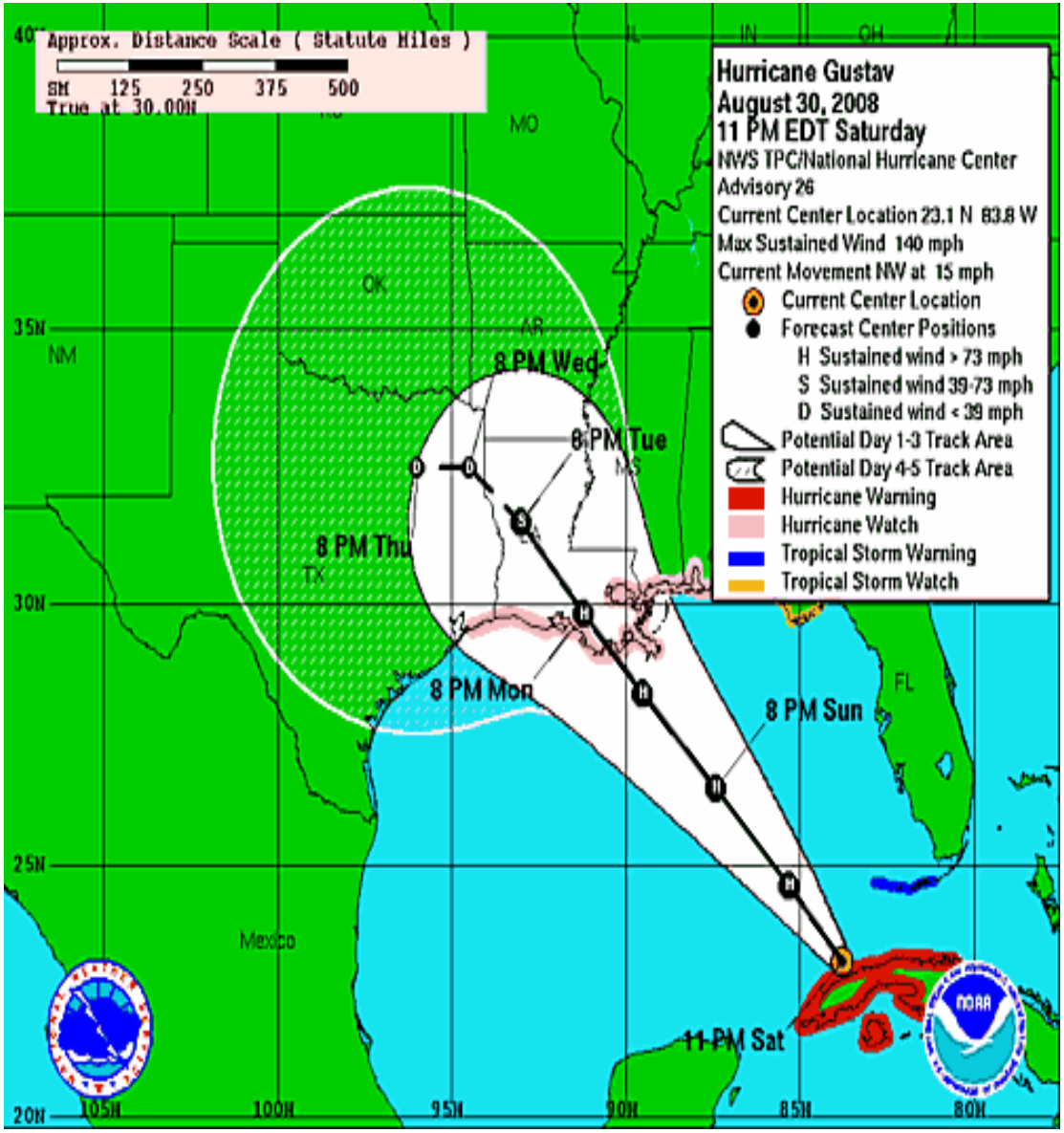
Goldilocks Sample Size

- Stop accrual for expected success
  - "If we stop enrolling now & track enrolled patients until complete data, will we have sufficient evidence?
  - If yes, stop accrual, wait, perform the decisive analysis
- Stop for futility
  - "If we enroll to the max will we have high chance of achieving goal?"
  - If not stop for futility

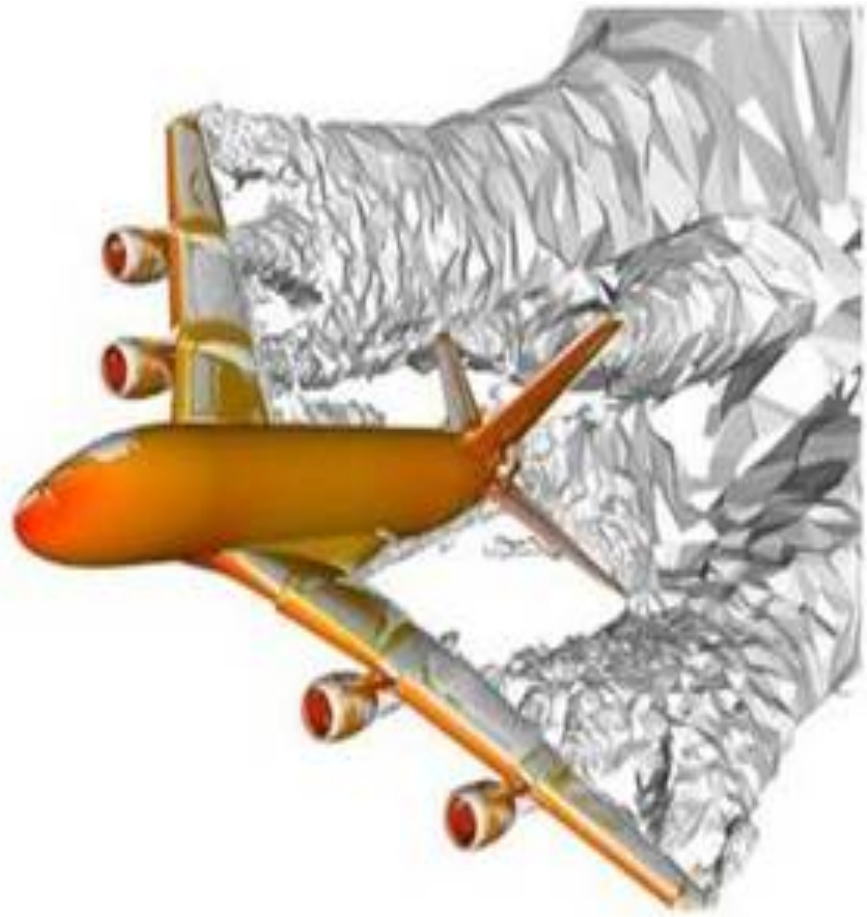
Computer Simulations

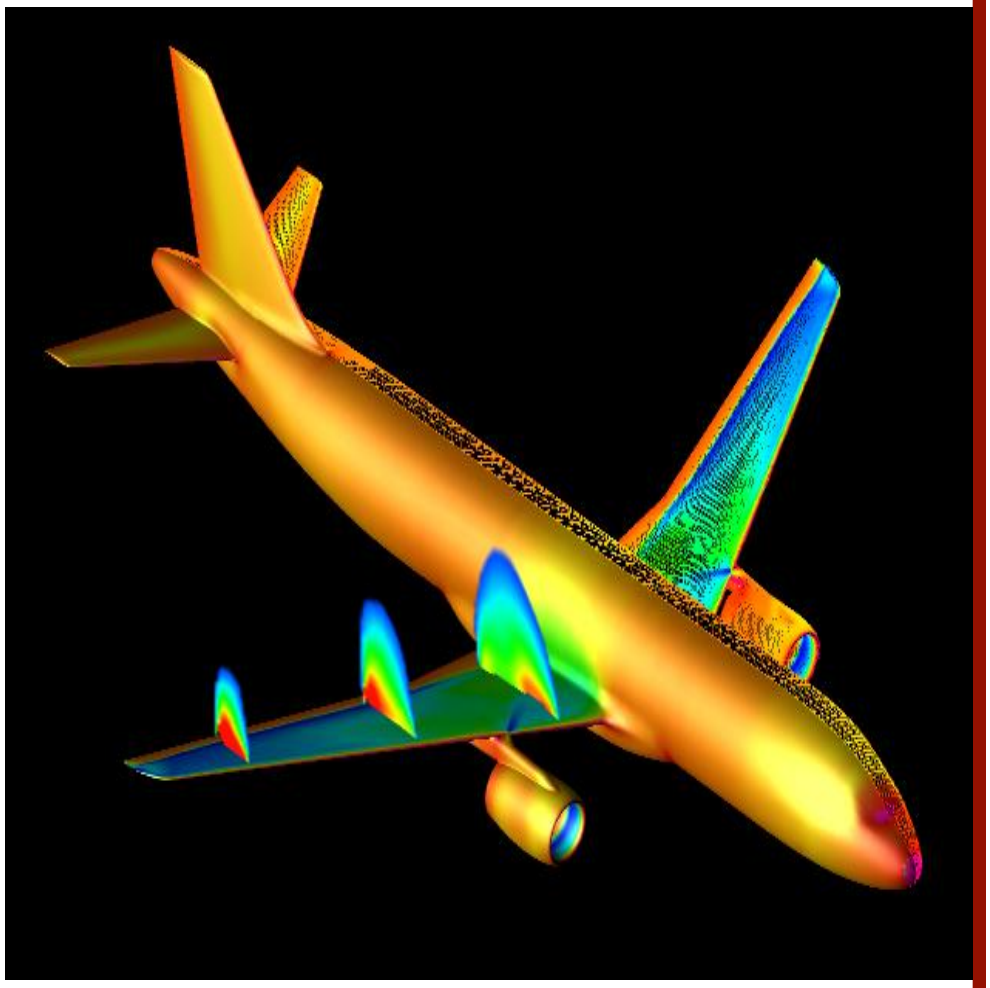
- We simulate the behavior of a design in order to find its performance on various metrics
- In this way it is a complex mathematical calculation as opposed to a prediction system
- This is numerical integration!
- Allows fully vetting the design as an instrument to learn about a medical therapy
- Not trying to predict outcome of a specific trial...

Weather Forecasts?



Building Airplanes?





Simulations and Regulatory Agencies

- How is a regulatory agency to evaluate trials built and maximized via simulation?
  - Can't be "if" but when and how
- We've been doing this for 10 years+ with CDRH
  - Enrichment, Unblinded Sample Size Selection, Many Interims, . . .
  - Type I error, Bias, Power, Simulation Code, . . .
  - Done with CDER, CBER, Case-by-case . . .
- Procedures for providing simulation code/ software/validation

Airplanes

- What would happen if we didn't simulations for building airplanes???

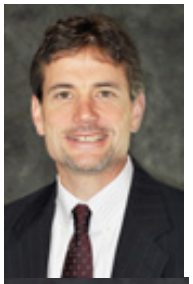

HL 7742

Future

- Where are we going?
- What is the future for randomized clinical trials?

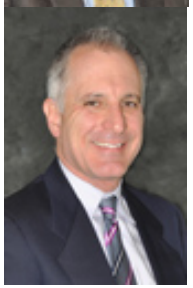
VIEWPOINT

The Platform Trial
An Efficient Strategy for Evaluating Multiple Treatments



Scott M. Berry, PhD
Berry Consultants LLC, Austin, Texas; and Department of Biostatistics, University of Kansas Medical Center, Kansas City.

Jason T. Connor, PhD
Berry Consultants LLC, Austin, Texas; and University of Central Florida College of Medicine, Orlando.



Roger J. Lewis, MD, PhD
Department of Emergency Medicine, Harbor-UCLA Medical Center, Torrance, California; and Berry Consultants LLC, Austin, Texas.

The drug development enterprise is struggling. The development of new therapies is limited by high costs, slow progress, and a high failure rate, even in the late stages of development. Clinical trials are most commonly based on a "one population, one drug, one disease" strategy, in which the clinical trial infrastructure is created to test a single treatment in a homogeneous population.

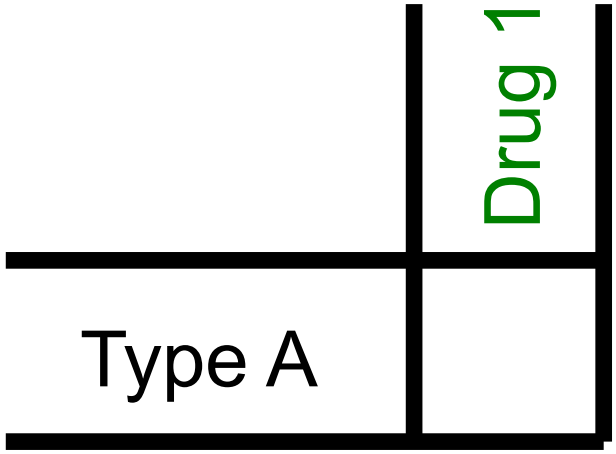
This approach has been largely unsuccessful for multiple diseases, including sepsis, dementia, and stroke. Despite promising preclinical and early human trials, there have been numerous negative phase 3 trials of treatments for Alzheimer disease[1] and more than 40 negative phase 3 trials of neuroprotectants for stroke.[2] Effective treatments for such diseases will likely require combining treatments to affect multiple targets in complex cellular pathways and, perhaps, tailoring treatments to subgroups defined by genetic, proteomic, metabolomic, or other markers.[3]

There has been increasing interest in efficient trial strategies designed to evaluate multiple treatments and benefits when evaluating potentially synergistic combination treatments (eg, treatment A, treatment B, treatment C, and all combinations) if the starting point is the testing of each treatment in isolation.

What Is a Platform Trial?

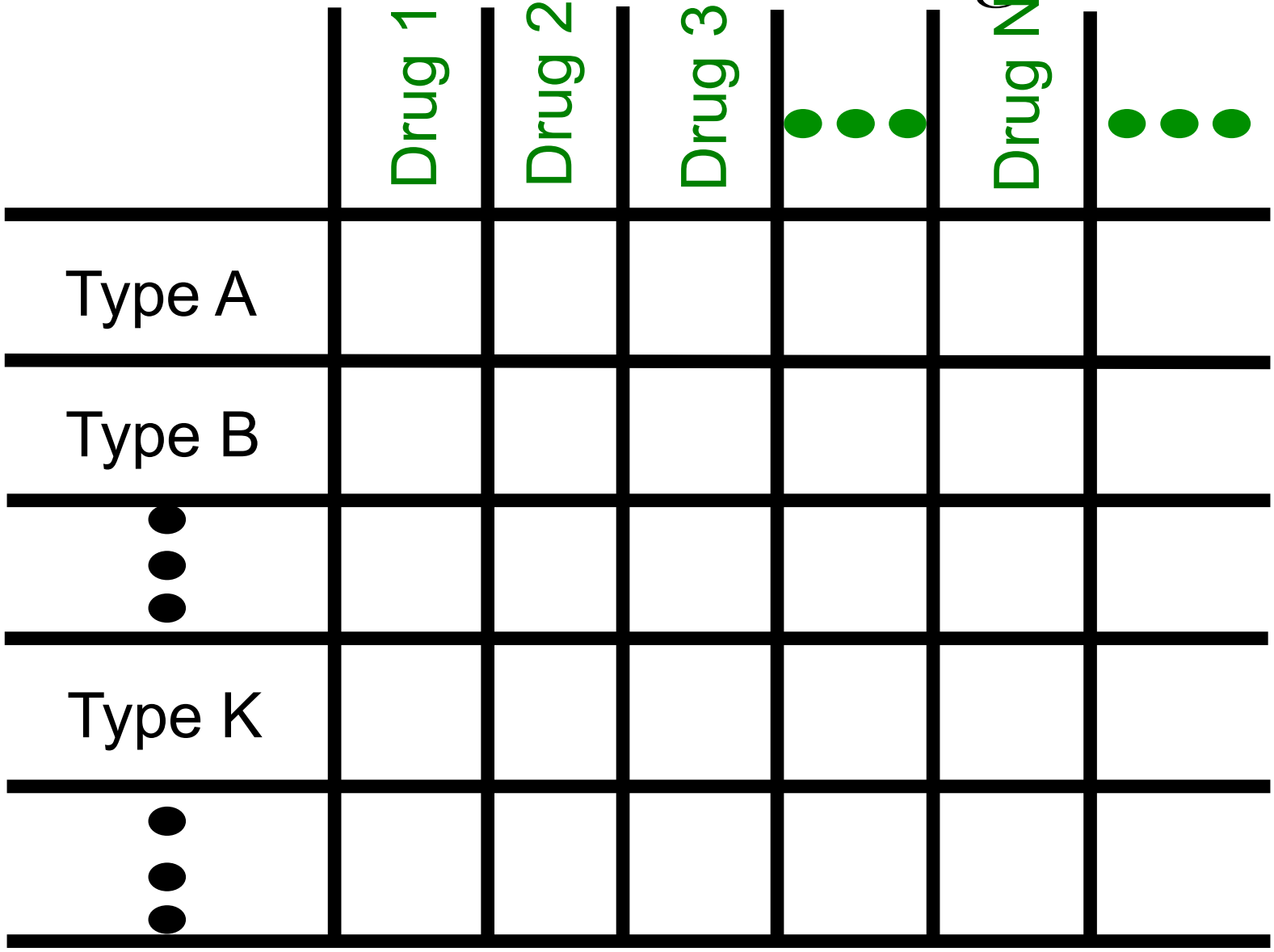
A platform trial is defined by the broad goal of finding the best treatment for a disease by simultaneously investigating multiple treatments, using specialized statistical tools for allocating patients and analyzing results. The focus is on the disease rather than any particular experimental therapy. A platform trial is often intended to continue beyond the evaluation of the initial treatments and to investigate treatment combinations, to quantify differences in treatment effects in subgroups, and to treat patients as effectively as possible within the trial. Although some of the statistical tools used in platform trials are frequently used in other settings and some less so, it is the integrated application of multiple tools that allows a platform trial to address its multiple goals. The **Table** summarizes the general differences be-

Traditional Trial: Focus on Treatment



"Standard Trial: Single treatment,
Homogeneous patients, Single question "

Platform Trial Designs



Adaptive Platform Trial Designs

- Master Protocol
- Focus is on the Disease
  - "What is the best treatment for a unique patient with this disease?
- Typical Innovations
  - Response Adaptive Randomization (RAR)
  - Patient heterogeneity (hierarchical modeling)
  - Combination treatments
  - Graduation/Removal, "Perpetual" trials
  - Statistical Modeling
- Bayesian methods and modeling are key for adaptations

Platform Trials

- Community Acquired Pneumonia
  - (PREPARE REMAP-CAP)
- Influenza (PREPARE ALICE)
- Breast Cancer (I-Spy2)
- Brain Cancer (GBM-AGILE)
  - *Google this, nice videos
- Pancreatic Cancer
- Antibiotics
- Alzheimer's (EPAD, DIAN)
- Lung Cancer (LUNG-MAP)
- Ebola
- Cystic Fibrosis
- Several rare diseases… and others "in the works"!

Summary

- Platform trials changing the landscape of clinical trials
- Changes trials from a focus on does X work, to treating patients better!
  - More effective, better treatment of patients, cheaper, faster…
  - BETTER SCIENCE!
- Are single sponsor trials dodos?
  - home phones, cable, newspapers, …?
  - 25 years from now what will the landscape be?