Understanding Simulations and Their Value in Clinical Trial Planning

William Meurer, MD, MS
Scott Berry, PhD

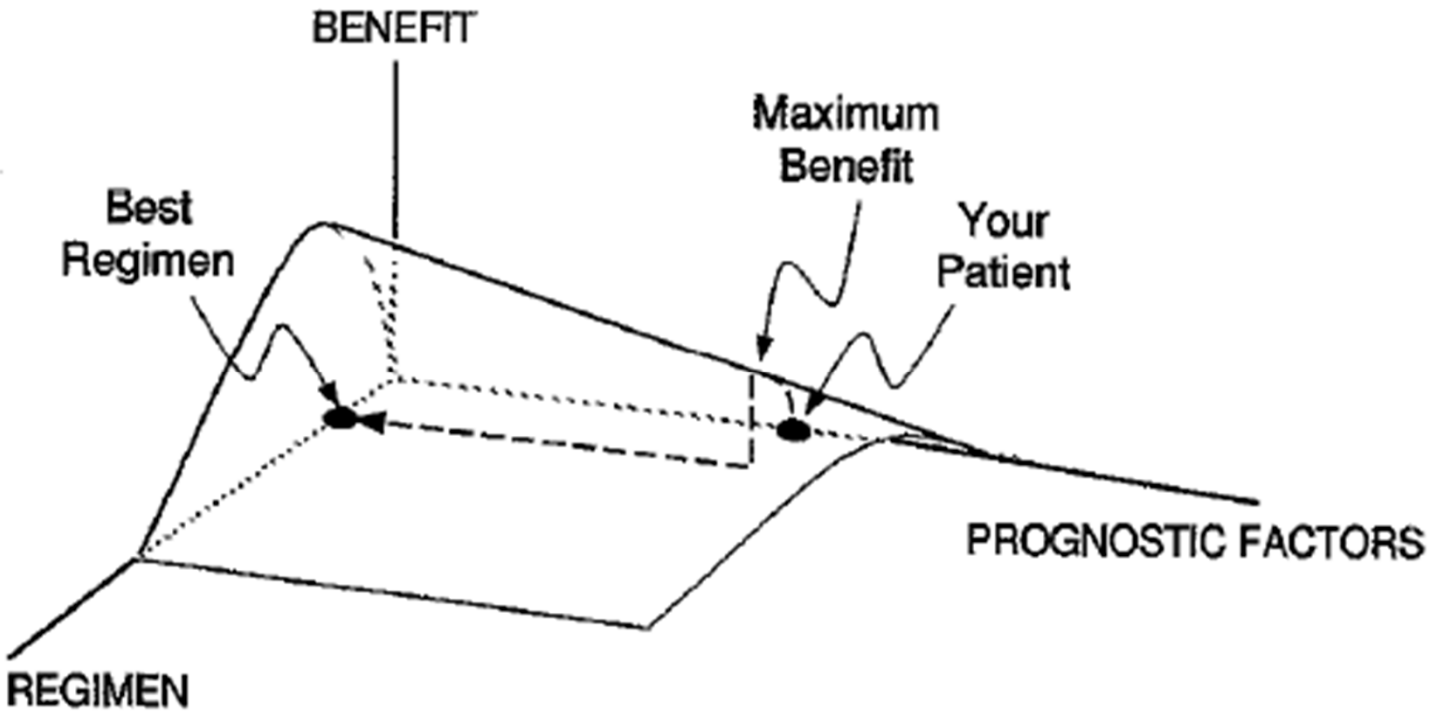
Objectives

- Understand why clinical trial simulation is needed
- Have familiarity with the general conduct of clinical trial simulations
- Be able to interpret clinical trial simulation results.

Learning vs. Confirming

- Learn to treat patients
  - Who
  - How
  - When
  - How long…
- Confirm treatment works

Therapeutic Response Surface



Clin Pharmacol Ther 1997;61:275-91

“I have always considered it more desirable to kill computer-generated patients than real ones when calibrating design parameters.”
Peter Thall

Chance 2001;14:23-8

Flexible Adaptive Designs

- May not have a direct analytical method for evaluating Type I and Type II error
- Simulation also allows estimation of the impact of various real-life clinical trial problems *(not limited to adaptive designs)*
  - Missing data
  - Choice of endpoint
  - Patient population
  - Covariate impact

Stat Methods Med Res. 2011 Dec;20(6):613-22

INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

PROBABILITY AND STATISTICS IN COMPLEX SYSTEMS:

GENOMICS, NETWORKS, AND FINANCIAL ENGINEERING

ORGANIZING COMMITTEE

MARCO AVELLANEDA	COURANT INSTITUTE, NYU
BRUCE HAJEK	UNIVERSITY OF ILLINOIS, URBANA
RICHARD KARP	UC BERKELEY
SALLIE KELLER-MCNULTY	LOS ALAMOS NATIONAL LABS
TOM KURTZ	UNIVERSITY OF WISCONSIN, MADISON
ANDREW LO	MIT
MICHAEL NEWTON	UNIVERSITY OF WISCONSIN, MADISON
SIMON TAVARÉ	UNIVERSITY OF SOUTHERN CALIFORNIA
WALTER WILLINGER	AT&T LABS-RESEARCH

APPLICATIONS ARE INVITED FOR POSTDOCTORAL MEMBERSHIPS (TWO YEAR APPOINTMENTS) AND SENIOR MEMBERSHIPS (FOR ONE TO NINE MONTHS), AND FOR INVITATIONS TO INDIVIDUAL WORKSHOPS. SEE WWW.IMA.UMN.EDU/DOCS/FORMS.HTML FOR APPLICATION INSTRUCTIONS AND FORMS.

INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS
400 LIND HALL, 207 CHURCH ST. S.E.
MINNEAPOLIS, MN 55455-0436
TELEPHONE: (612) 624-6066
FAX: (612) 626-7370
E-MAIL: IMA-STAFF@IMA.UMN.EDU
URL: WWW.IMA.UMN.EDU

SEPTEMBER

S	M	T	W	TH	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

OCTOBER

S	M	T	W	TH	F	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

NOVEMBER

S	M	T	W	TH	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30						

DECEMBER

S	M	T	W	TH	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

JANUARY

S	M	T	W	TH	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

FEBRUARY

S	M	T	W	TH	F	S
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29						

MARCH

S	M	T	W	TH	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

APRIL

S	M	T	W	TH	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

MAY

S	M	T	W	TH	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

JUNE

S	M	T	W	TH	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

SEPTEMBER 1, 2003 - JUNE 30, 2004

COMPLETE IMA PROGRAM INFORMATION IS ON-LINE AT: WWW.IMA.UMN.EDU/COMPLEX

FALL QUARTER (SEPTEMBER - DECEMBER, 2003)

MATHEMATICAL & STATISTICAL PROBLEMS IN GENOME SCIENCES

SEPTEMBER 15-19, 2003
SEPTEMBER 29 - OCTOBER 3, 2003
OCTOBER 20 - 24, 2003
NOVEMBER 17 - 21, 2003

OPENING WEEK TUTORIALS
WORKSHOP 1: STATISTICAL METHODS FOR GENE EXPRESSION MICROARRAYS AND PROTEOMICS
WORKSHOP 2: COMPARATIVE GENOMICS
WORKSHOP 3: NETWORKS AND THE POPULATION DYNAMICS OF DISEASE TRANSMISSION

WINTER QUARTER (JANUARY - MARCH, 2004)

COMMUNICATION NETWORKS

JANUARY 7 - 9, 2004
JANUARY 11, 2004
JANUARY 12 - 16, 2004
FEBRUARY 8, 2004
FEBRUARY 9 - 13, 2004
MARCH 7, 2004
MARCH 8 - 12, 2004

SHORT COURSE: THE INTERNET FOR MATHEMATICIANS
TUTORIAL: MEASUREMENT, MODELING AND ANALYSIS OF THE INTERNET
WORKSHOP 4: MEASUREMENT, MODELING AND ANALYSIS OF THE INTERNET
TUTORIAL: ROBUSTNESS AND THE INTERNET: DESIGN, EVOLUTION, AND THEORETICAL FOUNDATIONS
WORKSHOP 5: ROBUSTNESS IN COMPLEX SYSTEMS
TUTORIAL: CONTROL AND PRICING IN COMMUNICATION AND POWER NETWORKS
WORKSHOP 6: CONTROL AND PRICING IN COMMUNICATION AND POWER NETWORKS

SPRING QUARTER (APRIL - JUNE, 2004)

QUANTITATIVE MODELING IN FINANCE AND ECONOMETRICS

MARCH 29 - APRIL 2, 2004
APRIL 12 -16, 2004
MAY 3 - 7, 2004
MAY 24 - 28, 2004

SHORT COURSE: TOOLS FOR MODELING AND DATA ANALYSIS IN FINANCE/ASSET PRICING
WORKSHOP 7: RISK MANAGEMENT AND MODEL SPECIFICATION ISSUES IN FINANCE
WORKSHOP 8: MODEL IMPLEMENTATION, ALGORITHMS AND SOFTWARE ISSUES
WORKSHOP 9: FINANCIAL DATA ANALYSIS AND APPLICATIONS

The Presentation Of Simulation Results – By Phenotype

Statistician/Quantitative

-Data generation
- – Realistic
- – Transparent

-Analysis Methods
- – Robust
- – Precise
- – Unbiased
- – Reproducible

Clinician/Sponsor

-Decision making
- – Trial output

- Performance
- – Competing designs
- – Sample size
- – Type I and II error
- – Answering the question?

Name this car

Barriers

- Up front cost
- In academia, no funding for this sort of rigorous planning
- Simulation has occurred haphazardly in past (diminishing its value in some eyes)
- Reporting of simulation studies in biomedical literature often incomplete*

*Statist. Med 2006; 25:4279-4292

venturebeat.com/2013/02/22/big-data-is-dead-whats-next/

Work | Home | Fun | Temporary | Gmail - Inbox (1) - will... | Home - Google Docs | Lib Proxy | School

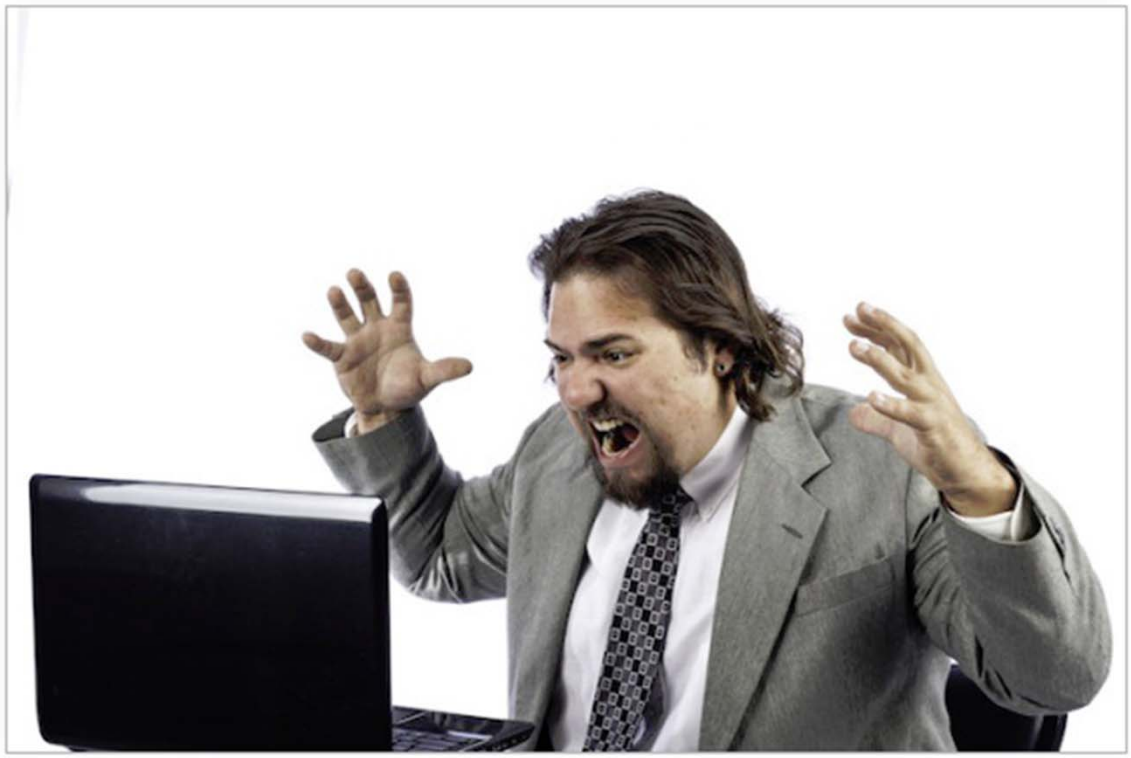
VB *you are reading...*
'Big data' is dead. What's next?

channels

- Main
- Big Data
- Business
- Cloud
- Deals
- DEMO
- Dev
- Entrepreneur
- Gadgets
- Green
- Health
- Lifestyle
- Media
- Mobile
- New York
- Science
- Security
- Small Biz
- Social
- GAMESBEAT

GUEST POST

'Big data' is dead. What's next?

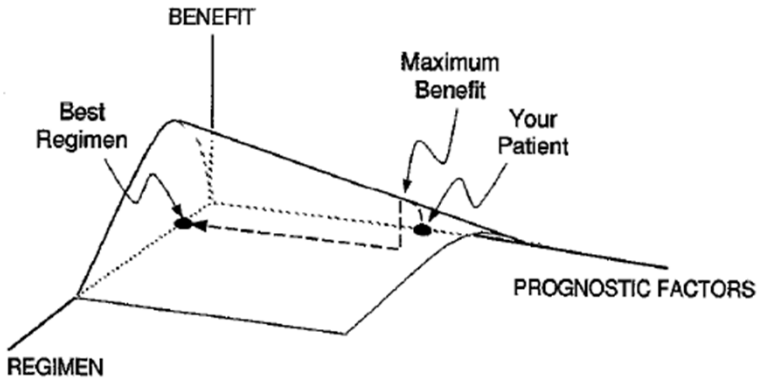


February 22, 2013 2:40 PM
John De Goes

23 Comments

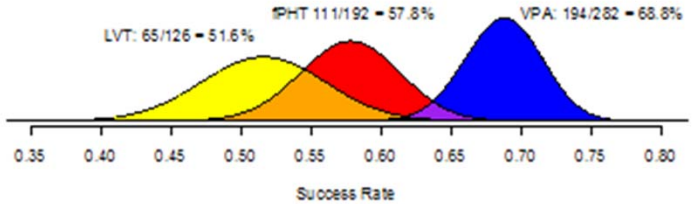
Simulations, Scenarios, Sample Trials

- Adaptive designs – simulate trials to see how “machine” works
- Scenarios – stress test the machine under different assumed truths (all drugs the same, 1 really good, etc)
- Sample Trial – watch progress of virtual trial (as a DSMB would)
- Simulations reports aggregate the results of MANY sample trials



Remember these from yesterday

Example Trial: Final Evaluation



Treatment	Observed	%	95% CI	Pr(Best)	Pr(Worst)
VPA	194/282	68.8%	(.632, .739)	0.992	0.0005
fPHT	111/192	57.8%	(.507, .646)	0.007	0.138
LVT	65/126	51.6%	(.429, .601)	0.0005	0.862

Difference	Observed	95% CI	Pairwise Comparison
VPA – fPHT	0.110	(0.022, 0.197)	Pr(VPA>fPHT) = 0.993
VPA – LVT	0.172	(0.069, 0.272)	Pr(VPA>LVT) > 0.999
fPHT - LVT	0.062	(-0.049, 0.172)	Pr(fPHT>LVT) = 0.862

Operating Characteristics

Scenario 3 Efficacy Rates	Pr(ID Best) Early-End	Pr(Randomize To Best)	Mean N
Null 0.5 – 0.5 – 0.5	0.020 0.019 0.001	100%	545
One Good 0.5 – 0.5 – 0.65	0.939 0.893 0.046	48%	494
Two Good 0.5 – 0.65 – 0.65	0.109 0.099 0.010	87%	753
One Middle One Good 0.5 – 0.575 – 0.65	0.536 0.473 0.063	48%	635
All Bad 0.10 – 0.10 – 0.10	0.005 0.005 0.000	100%	400

Simulations

Scott M. Berry

April 10, 2013

berryconsultants.com

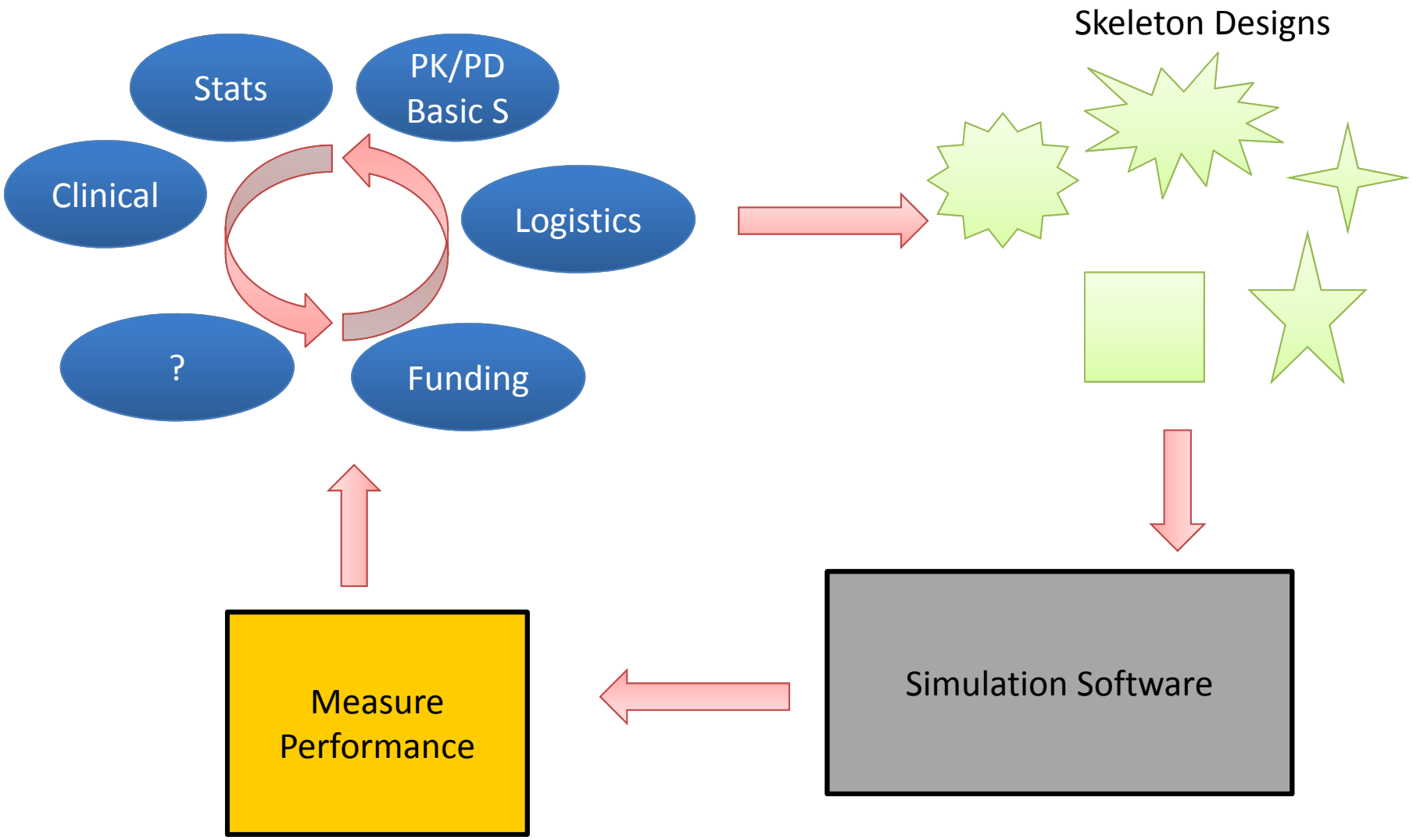
What Stage/Phase of CT?

- Phase I:
  - Sample size
  - Dose escalation
  - Combination of arms
  - Seamless phase I-II
- Phase II/Pilot:
  - Sample size
  - Dose allocation
  - Introduce/Drop arms
  - Enrichment
  - Prediction of Phase III
  - Seamless phase II-III
- Phase III/Confirmatory:
  - Sample size
  - Multiple Arms
  - Accrual Interim Analyses
  - Futility Analyses
  - Timing of Conclusions
  - Enrichment
- Phase IV:
  - Sample size
  - Timing of Conclusions
  - Indications

Therapeutic Areas/Diseases

- Oncology
- Migraine
- Lupus
- Sepsis
- Diabetes
- Obesity
- Stroke
- Tinnitus
- MS
- CHD
- Smoking Cessation
- Gastroparesis
- Alzheimers
- Atrial Fibrillation
- Cancer diagnostic
- Disc Disease
- Contraceptives
- Valves/stents
- Asthma
- Emphysema
- PFO
- RA
- Sleep Apnea
- Osteoparesis
- Parkinsons
- Pain
- Hydrocephalus
- HIV
- Schizophrenia
- Crohns
- Spinal Cord Injury
- Hep C
- Preterm Labor
- Constipation
- Micturition
- Drooling
- PO Ileus
- DVT
- Sexual health
- Emesis
- Statins
- Infections
- OAB
- TB
- Head Trauma
- Cardiac Arrest
- ALS
- Alcohol Abuse
- SARI

Design Process



ICECAP

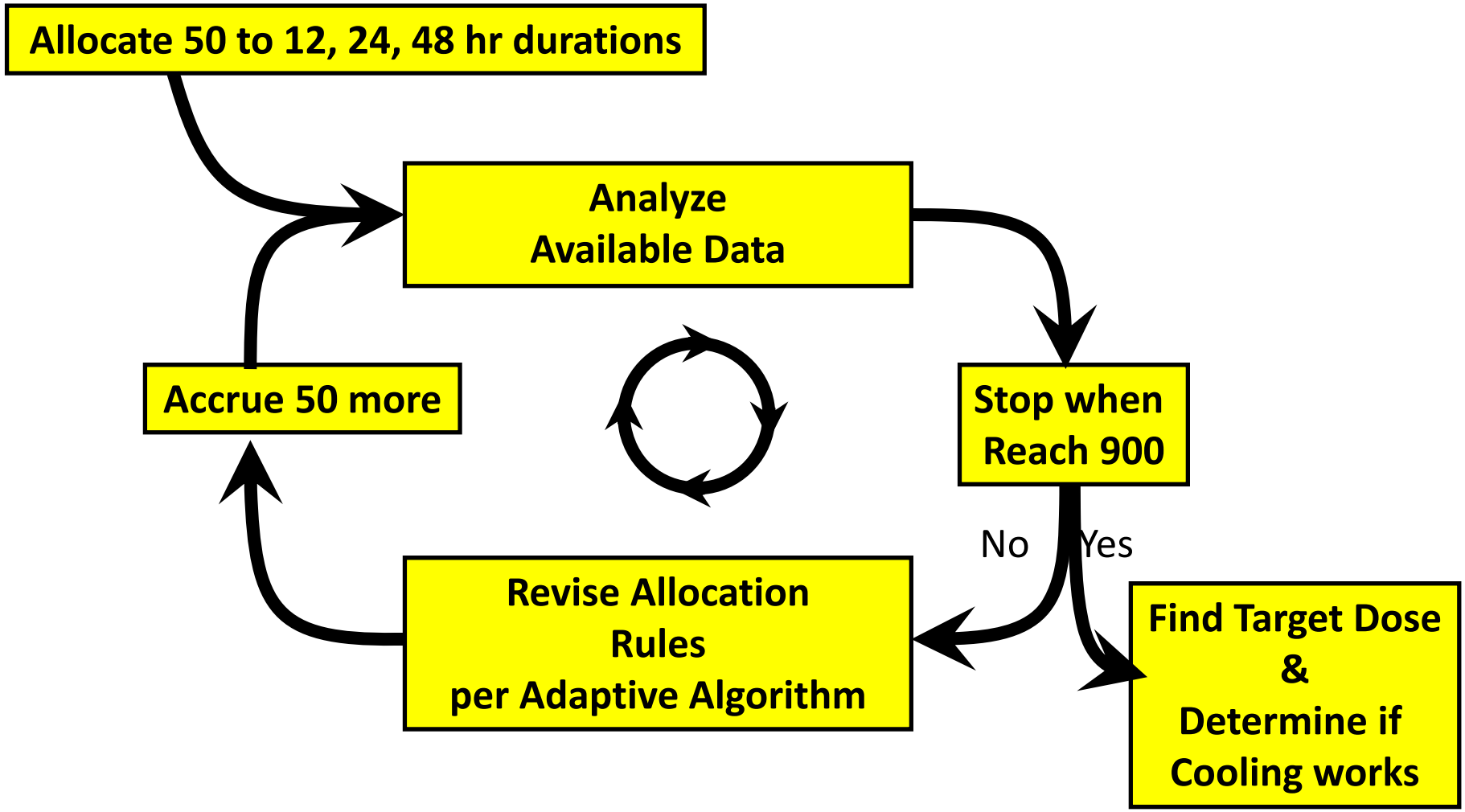
"Under Construction"

- ICECAP – Hypothermia after post cardiac arrest coma
  - Background
    - Two small surface cooling trials demonstrated efficacy (different durations and endovascular cooling more frequently used)
    - Medically accepted that this works
    - No FDA approval
  - Goals
    - To identify optimum cooling duration
    - Gain additional insight into efficacy (functional form of duration response model)
    - What types of strokes vs. duration
  - Fixed Design:
    - 300? On 12, 24, 48 hours cooling

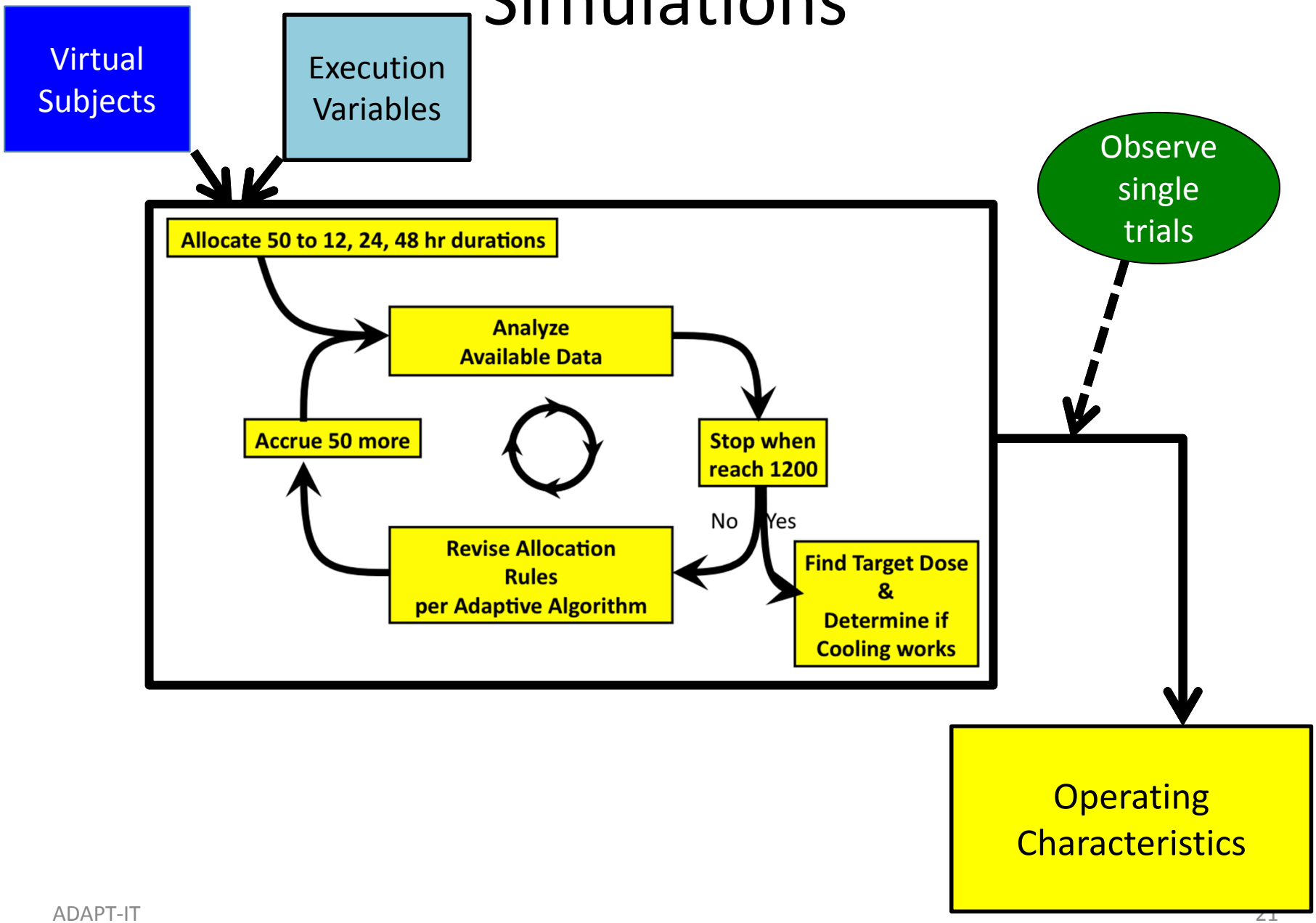
Initial skeleton

- Start with 12, 24, 48-hour durations (say 50/arm)
- Then analyze data and randomize to the best duration
  - Allow randomization to a much wider grid:
  - 6, 12, 18, 24, 30, 36, 42, 48, 60, 72
- Continue updating, say every 50 patients
- Continue to end of trial

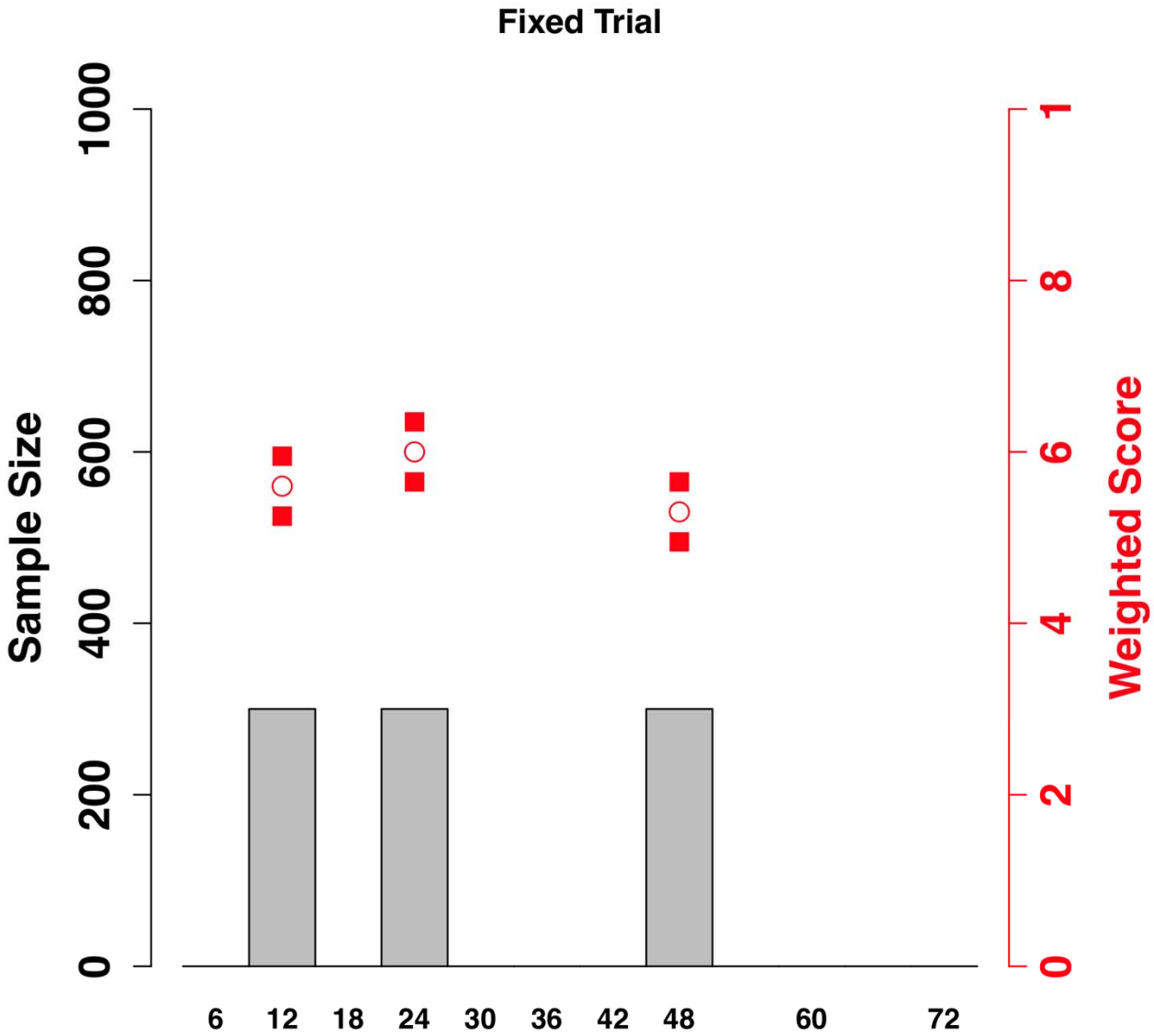
Adaptive Algorithm



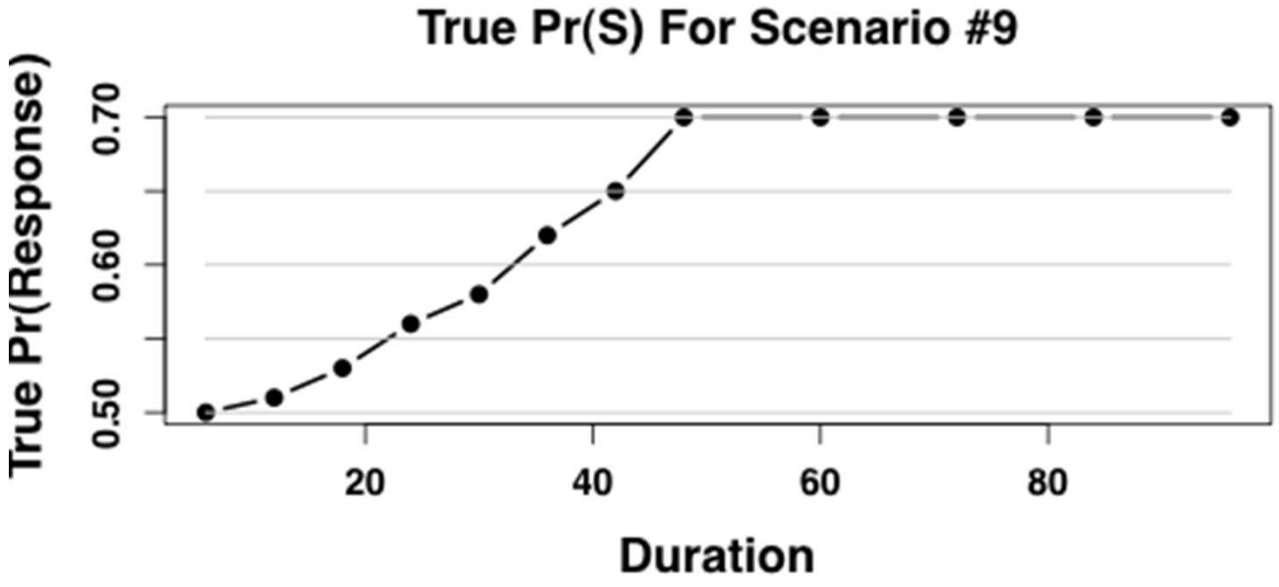
Simulations

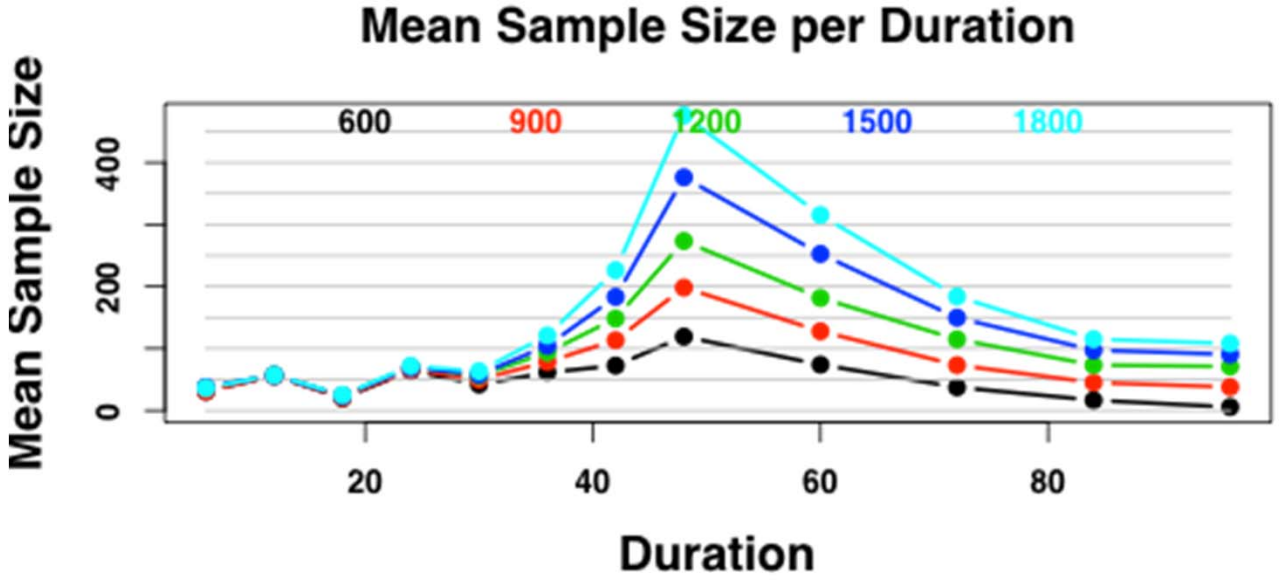


Example Outcome of Fixed



- Idealized Outcome?
- Answer All your questions?
- Do anything differently?


True Pr(S) For Scenario #9
True Pr(Response)
0.50
0.60
0.70
20
40
60
80
Duration


Mean Sample Size per Duration
Mean Sample Size
0
200
400
600
900
1200
1500
1800
20
40
60
80
Duration

Role Simulations

- Incredible Learning Tool
  - Team, Regulators, Funders, DSMB, Operations
- Changed Models
- Changed measures of success
- Endpoint (dichotomous) wasn't correct
  - Weighted one
- Needed both rhythm types (shockable and non-shockable)
  - Possibly different duration, relative efficacy
- All recognized through flight simulator
  - Single example trials critical

Value Added

Journal of Diabetes Science and Technology
Volume 6, Issue 6, November 2012

ORIGINAL ARTICLE

Application of Adaptive Design Methodology in Development of a Long acting Glucagon-like Peptide-1 Analog (Dulaglutide): Statistical Design and Simulations

Zachary Skrivanek, Ph.D.,[1] Scott Berry, Ph.D,[2] Don Berry, Ph.D.,[2,3] Jenny Chien, Ph.D.,[1]
Mary Jane Geiger, M.D., Ph.D.,[1] James H. Anderson, Jr., M.D.,[4] and Brenda Gaydos, Ph.D.[5]

- Lilly (seamless) Diabetes Trial
  - Trial went from 3 to 7 doses
    - Automatic selection of 2 doses (utility function)
  - Signaled additional phase III trials to start (doses)
  - Accrual rates 6-10/week
  - Control of Type I error

Value Added

- Phase I – II Seamless Oncology
  - Created hundreds of movies of escalation rules
    - Combined Adults/Kids
  - Simulations separated “rules” from “model borrowing”
  - Added Utility function for Tolerability & Efficacy

Value Added

- X Tumor Agnostic
  - Rules for approval
    - By simulating many trials we could show FDA exactly what “success” meant
    - Can we approve with 1/1 ? Okay?
  - Added rules for minimum information needed to gain approval

Value Added

- ARCTIC Trial
  - 3 durations of cooling for spinal cord injury vs. No Cooling
    - Adaptive randomization for full trial? Find and confirm best duration
    - Compared to AR, followed by 1:1 comparison phase (same maximum sample size)
  - Despite better performance, acceptability by community very important – Two stage
    - Final results, trial examples

Value Added

- SHINE Trial
  - Tight glucose control in hyperglycemic acute ischemic stroke patients
  - Use of blinded sample size re-estimation
  - During simulations of the procedure we noticed that when there is a treatment effect the sample size was almost always increased – then the trial may stop for superiority, or be unnecessarily large
    - Algorithm confused between treatment effect and larger variance

Value Added

- Very Common:
  - We describe the design, and the first comment is: "Wow, that is way too complex"
  - We then show simulations of example trials:
    - "Could you add X, Y, and Z"
  - Brings a great deal of comfort!
    - You can do this!

Conclusions

- The trial is ready to run – code written, structure ready
  - What data in needed?
- Risks for execution parameters known
- Trial has been carried out millions of times before it is run
  - It's as though team is adjusting the trial exactly as they should/would!
- The real trial shouldn't be the first time your trial is run.