Bayesian Adaptive Clinical Trial Design

Jason Connor

ConfluenceStat

Jason@ConfluenceStat.com

412-860-3113

Introduction

Introductions

Decision Problem 1: Pandemic!

- A pandemic just hit the USA!!
- Patients are dying from a deadly disease
- 7-day survival rate is estimated to be less than 50% with standard care
- Patients who are alive at 7 days after initial symptoms typically have full recovery

Decision Problem 1: Pandemic!

- We need to determine best treatment of infected people
- Currently available therapies
  - Standard care with aforementioned ~50% mortality
  - 3 experimental anti-virals are ready to go
  - Each experimental arm is a novel anti-viral drug plus standard care
- Primary Endpoint:
  - Alive at 7 days after randomization (yes/no)

Allocation of Patients

- An effective treatment is any treatment that is better than standard care
- We will design the trial in stages, lets say we can enroll 80 patients per month
- You tell me where you want to assign patients
- I'll tell you how many on each drug survived

Interim Analyses

- At each interim analysis, you will receive efficacy data and will have to decide one of three things:
    1. Terminate the trial for futility, choose standard care as best option
    2. Stop the trial for success, choose optimal drug to treat all future patients
    3. Continue to collect data, allocating the next 80 patients to the four arms however you choose

Contest Points

- Team Competition
  - Each deceased patient costs 5 points
  - Every minute it takes to make a final decision costs 50 points (e.g., 20 minutes costs 1000 points)
  - If you claim a drug is superior to standard care (successful trial):
    1. If (in truth) the chosen drug is not superior to standard care, you lose 1,000 points
    2. If (in truth) the chosen drug is superior to standard care, you receive 2,000 points plus 200 for each % efficacy compared to control
  - If you claim standard care is best (futile trial):
    1. If (in truth) at least one of the drugs is superior to standard, you lose 1,000 points
    2. If (in truth) all drugs are not superior to standard, you receive 2,000 points

Instructions

- Write on a piece of paper how many patients (80 total) you would like to allocate to
  - Standard Care
  - Drug 1
  - Drug 2
  - Drug 3
- Bring to me to receive instant patient results
  - But don't get too close
- Repeat process until treatment is selected by group

Assign Teams

Go!

REMINDER TO SELF:
START CLOCK

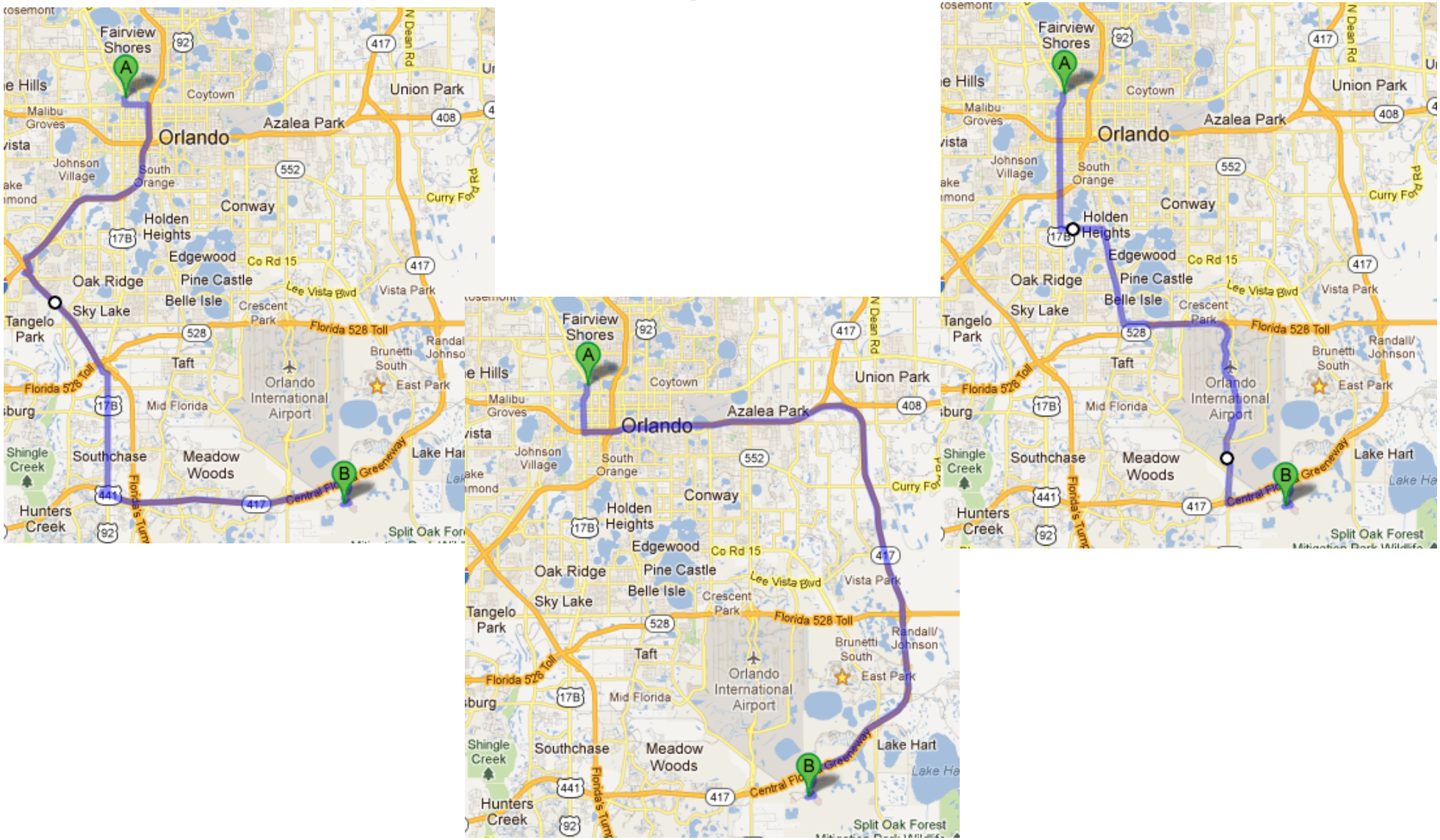
Discussion

Decision Problem #2*

- You move to a new city and start a new job
- Want the fastest way to work
  - Could take the highway (smaller μ, larger σ?)
  - Could take surface roads (larger μ, smaller σ?)
- How do you decide which way to go?
- How do you decide how to decide?

* Only for those who've moved to a city & started a new job prior to owning a smartphone

A common CER trial: How do I get to work?



Desirable Qualities of an RCT

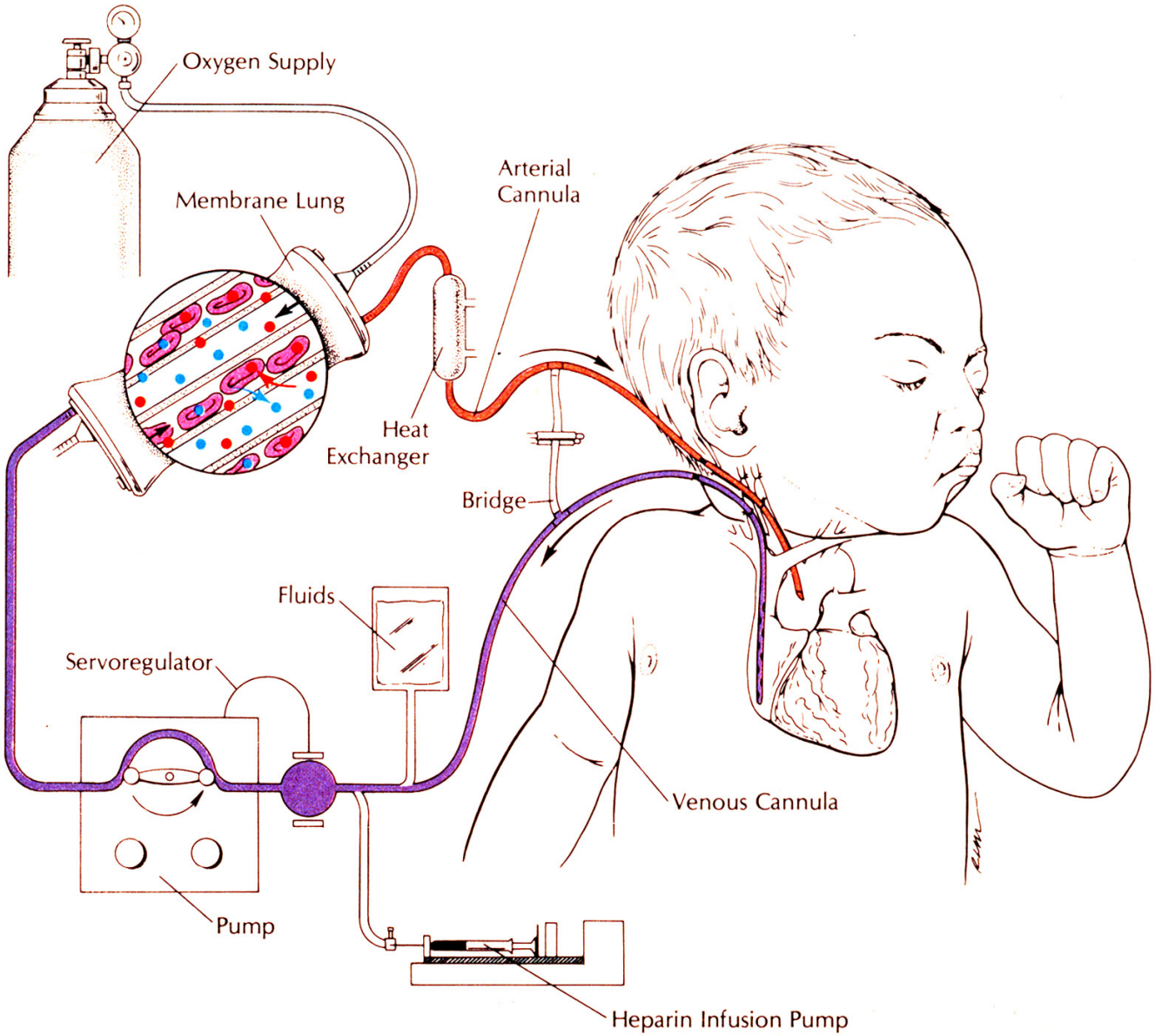
- Row 1
- Row 2

Decision Problem #3

- New device to assist pre-mature infants
- Historical mortality rate >75%
- How to decide if new device is better than standard of care?

Decision Problem 3: ECMO

- Extracorporeal membrane oxygenation
- Oxygenates babies' blood & gives underdeveloped lungs & heart time to heal or grow
- Historical survival rates $\leq 25\%$
- Michigan trial: Randomized play the winner strategy
  - Bartlett, *Pediatrics,* 1985, 76: 479-487


Oxygen Supply
Membrane Lung
Arterial Cannula
Heat Exchanger
Bridge
Fluids
Servoregulator
Venous Cannula
Pump
Heparin Infusion Pump

Randomization Rules

- Randomize first patient 1:1 to treatment *t*
- If survives on treatment *t*, add 1 "*t*-colored" ball
- If dies on treatment *t*, add 1 other colored ball
- Treat 10 patients this way

- Expected number patients treated with better treatment > 5, "ethical"

ECMO Results

	Prob to ECMO	TRT	Result	Balls in Urns CMT	ECMO
Start				1	1
1	0.50				
2					
3					
4					
5					
6					
7					
8					
9					
10					

ECMO Results

	Prob to ECMO	TRT	Result	Balls in Urns CMT	ECMO
Start				1	1
1	0.50	ECMO			
2					
3					
4					
5					
6					
7					
8					
9					
10					

ECMO Results

	Prob to ECMO	TRT	Result	Balls in Urns CMT	ECMO
Start				1	1
1	0.50	ECMO	Lived		
2					
3					
4					
5					
6					
7					
8					
9					
10					

ECMO Results

	Prob to ECMO	TRT	Result	Balls in Urns CMT	ECMO
Start				1	1
1	0.50	ECMO	Lived	1	2
2	0.67				
3					
4					
5					
6					
7					
8					
9					
10					

ECMO Results

	Prob to ECMO	TRT	Result	Balls in Urns CMT	ECMO
Start				1	1
1	0.50	ECMO	Lived	1	2
2	0.67	CMT	Died	1	3
3	0.75				
4					
5					
6					
7					
8					
9					
10					

ECMO Results

	Prob to ECMO	TRT	Result	Balls in Urns CMT	ECMO
Start				1	1
1	0.50	ECMO	Lived	1	2
2	0.67	CMT	Died	1	3
3	0.75	ECMO	Lived	1	4
4	0.80				
5					
6					
7					
8					
9					
10					

ECMO Results

	Prob to ECMO	TRT	Result	Balls in Urns CMT	ECMO
Start				1	1
1	0.50	ECMO	Lived	1	2
2	0.67	CMT	Died	1	3
3	0.75	ECMO	Lived	1	4
4	0.80	ECMO	Lived	1	5
5	0.83				
6					
7					
8					
9					
10					

ECMO Results

	Prob to ECMO	TRT	Result	Balls in Urns CMT	ECMO
Start				1	1
1	0.50	ECMO	Lived	1	2
2	0.67	CMT	Died	1	3
3	0.75	ECMO	Lived	1	4
4	0.80	ECMO	Lived	1	5
5	0.83	ECMO	Lived	1	6
6	0.86	ECMO	Lived	1	7
7	0.88	ECMO	Lived	1	8
8	0.89	ECMO	Lived	1	9
9	0.90	ECMO	Lived	1	10
10	0.91	ECMO	Lived	1	11

What Would You Decide?

- ECMO 9/9 CMT 0/1*

  * The 1 on CMT was the sickest of all patients

- As a statistician / clinical trialist do you have sufficient information to declare ECMO more efficacious than standard of care?

What Would You Decide?

- ECMO 9/9 CMT 0/1*

  * The 1 on CMT was the sickest of all patients

- As a statistician / clinical trialist do you have sufficient information to declare ECMO more efficacious than standard of care?

- As a parent would you dare *not* request ECMO for your premature baby?

Lessons of ECMO

- Questions the trials designers should have asked *before* the trial
  - How do we calculate a p-value?

Lessons of ECMO

- Questions the trials designers should have asked *before* the trial
  - How do we calculate a p-value?
  - Published p-values for this data (Stat Sci Nov 1989)

0.00049	0.051
0.001	0.083[F]
0.003	0.280
0.009	0.500
0.038	0.617
0.045	1.000
undefined	

Lessons of ECMO

- Questions the trials designers should have asked *before* the trial
  - How do we calculate a p-value?
  - Will the medical community believe our results?
    - Will we have enough data to sway opinions of people with a wide range of prior beliefs
  - What are trial results likely to look like?
  - What if everyone is randomized to ECMO?
    - If CMT success = 30% and ECMO success = 90%

      6% chance all 10 patients will be randomized to ECMO

Follow-Up Trials

- Harvard
  - Stage 1: randomize equally until 4 deaths in one arm
  - Stage 2: assign all to other arm until 4 deaths or stat sig.
  - 6/10 conventional therapy (60%)
  - 9/9 & 19/20 on ECMO (97%)
  - *Pediatrics,* 1989, 84: 957-963
- U.K
  - 63/93 on ECMO (68%)
  - 38/92 on conventional therapy (41%)

    Lancet, 1996, 348: 75-82
- Were these study designs ethical?
- Do we have an irrational commitment to blinded RCTs?
- Do we have an irrational commitment to $p < 0.05$?
- Does lack of $p<0.05$ mean equipoise until we see $p<0.05$?

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

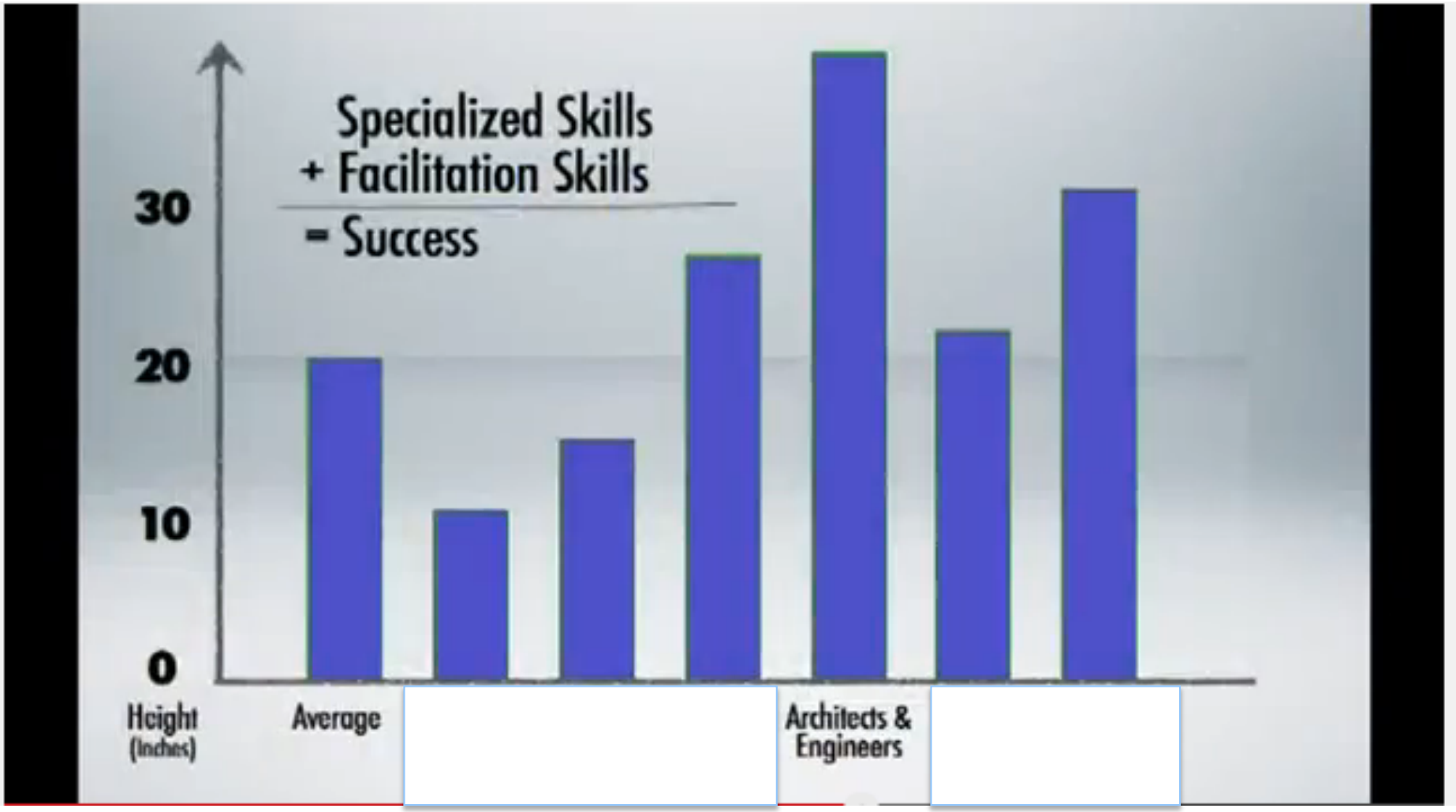
Digression:
The Marshmallow Design Challenge

The Marshmallow Design Challenge
Peter Skillman

- 4-person team
- 18 minutes
- 20 pieces of raw spaghetti
- 1 meter of tape
- 1 meter of string
- 1 marshmallow

Peter Skillman Marshmallow Design Challenge
https://www.youtube.com/watch?v=1p5sBzMtB3Q

The Marshmallow Design Challenge

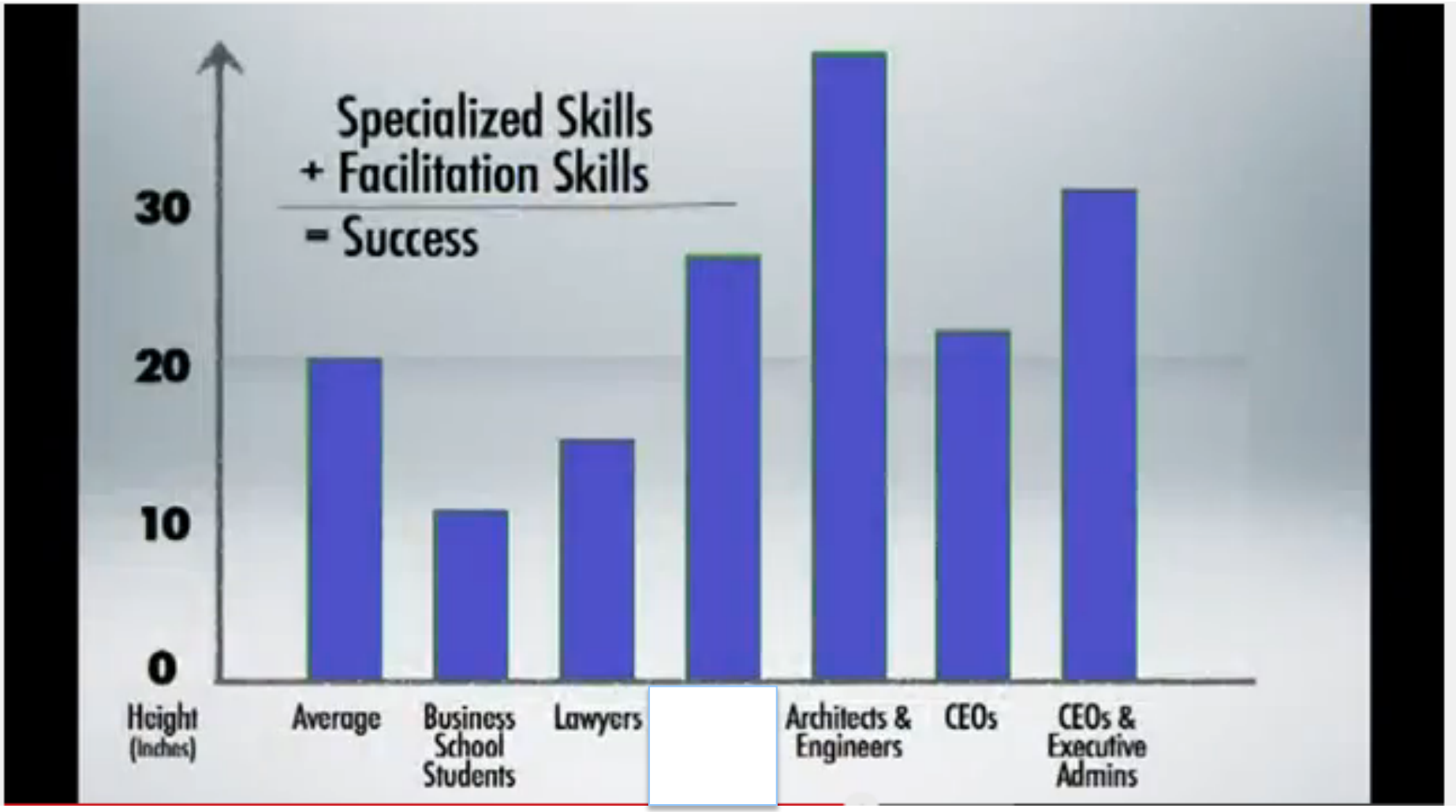


Tom Wujec: Build a tower, build a team.
https://www.youtube.com/watch?v=H0_yKBitO8M

The Marshmallow Design Challenge

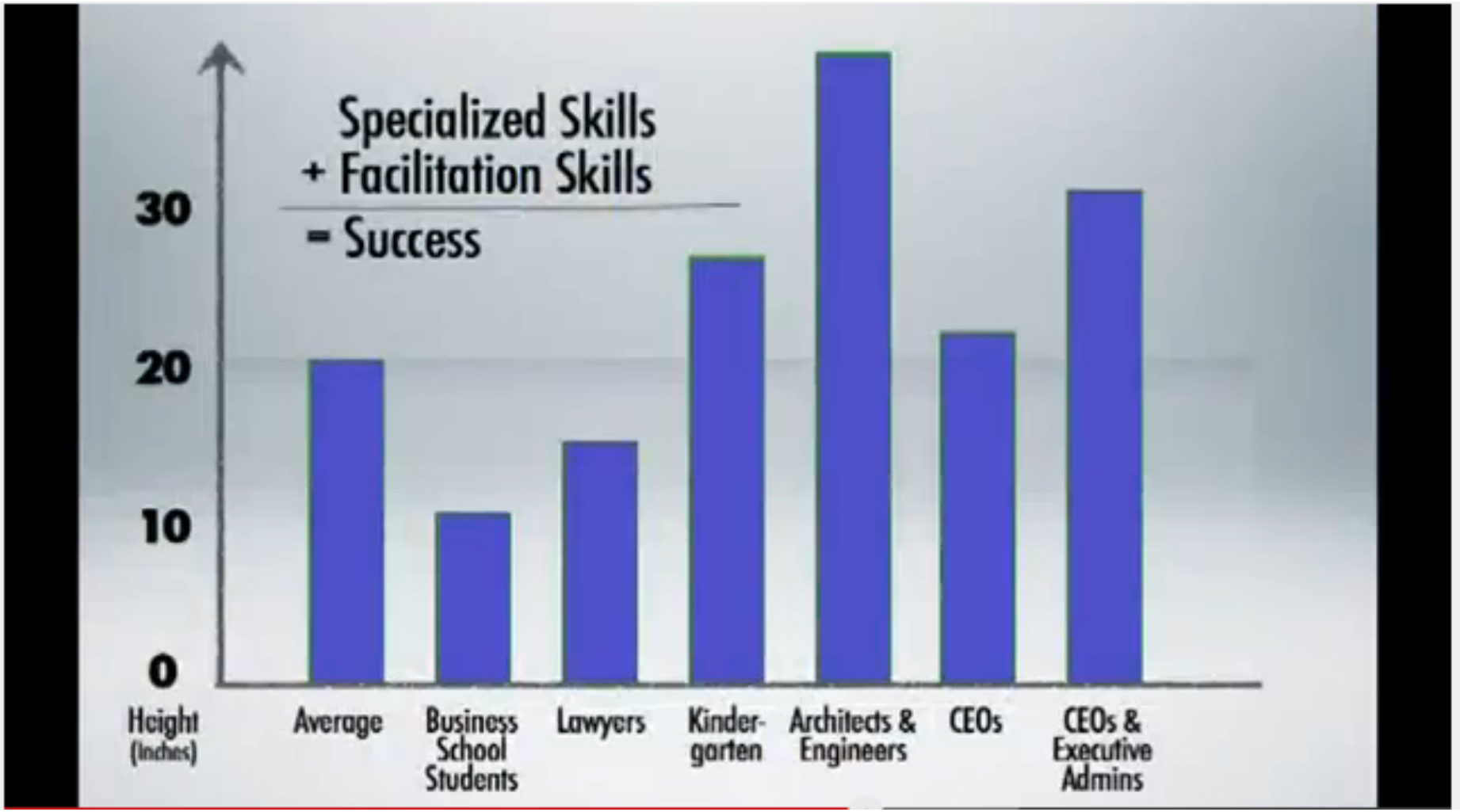
Tom Wujec: Build a tower, build a team.
https://www.youtube.com/watch?v=H0_yKBitO8M

The Marshmallow Design Challenge



Tom Wujec: Build a tower, build a team.
https://www.youtube.com/watch?v=H0_yKBitO8M

The Marshmallow Design Challenge



Tom Wujec: Build a tower, build a team.
https://www.youtube.com/watch?v=H0_yKBitO8M

The Marshmallow Design Challenge
Peter Skillman

- Kindergartners
  - Don't waste time seeking power
  - Don't sit around talking about the problem
  - Try, fail, try, fail until time runs out
  - They all grab stuff and try things
  - Usually keep the marshmallow on top when trying

- MBA grads
  - Spend a lot of time talking
  - Trained to find single best plan
  - Trained never to fail
  - Last thing they do it put the marshmallow on top (and often watch the whole tower collapse)

The Marshmallow Design Challenge
Peter Skillman

- You learn by doing and failing & redoing
- Work in parallel
- Doing multiple iterations is good
- All projects have resource constraints

ECMO: Trial & Error Design by Simulation

```
p.ecmo <- 0.75;      p.cmt <- 0.25

group.vec <- NULL;    outcome.vec <- NULL
outcome <- matrix(nrow=100000, ncol=5)

for(s in 1:100000){
urn <- c(1,1)
for(pt in 1:10){
  group <- sample(c("C","E"), 1, prob=urn)
  result <- rbinom(1, 1, ifelse(group=="C",p.cmt, p.ecmo))
  if(group=="C"){
     if(result==1){
          urn[1] <- urn[1] + 1
     }else{
          urn[2] <- urn[2] + 1
     }
  }else{
     if(result==1){
          urn[2] <- urn[2] + 1
     }else{
          urn[1] <- urn[1] + 1
     }
}
group.vec[pt] <- group
outcome.vec[pt] <- result
}
tab <- table(factor(group.vec, levels=c("C","E")), factor(outcome.vec,
levels=0:1))
outcome[s,] <- c(c(tab), fisher.test(tab, alternative='greater')$p.value)
print(s)
}
```

```
### Pr no patients on control
mean((outcome[,1]+outcome[,3]) == 0)
### Pr no patients on ECMO
mean((outcome[,2]+outcome[,4]) == 0)
### Pr more on ECMO than control
mean((outcome[,1]+outcome[,3]) < (outcome[,2]+outcome[,4]))
### Pr more equal on each
mean((outcome[,1]+outcome[,3]) == (outcome[,2]+outcome[,4]))
### Pr more on control than ECMO
mean((outcome[,1]+outcome[,3]) > (outcome[,2]+outcome[,4]))

### More ECMO than control success
mean((outcome[,3]) < (outcome[,4]))
### 4 or more ECMO than control successes
mean((outcome[,3] + 4) <= (outcome[,4]))
```

ECMO: Prospective Simulation

Operating Characteristics	CMT 25% ECMO 75%	CMT 25% ECMO 25%
Pr(All patients randomized to ECMO)	2.5%	0.04%
Pr(All patients randomized to CMT)	0.04%	0.04%
Pr(Majority to ECMO)	72%	36%
Pr(5 ECMO & 5 CMT)	14%	27%
Pr(Majority to CMT)	14%	36%
Pr(Fisher P-value < 5%) Pr(Chi-square P-value < 5%)	12% 32%	0.1% 1.9%
Pr(# ECMO Success > # CMT Successes)	89%	38%
Pr(# ECMO Success ≥ # CMT Success + 4)	59%	2.7%

ECMO: Prospective Simulation

Operating Characteristics	CMT 25% ECMO 75%	CMT 25% ECMO 25%
Pr(All patients randomized to ECMO)	2.5%	0.04%
Pr(All patients randomized to CMT)	0.04%	0.04%
Pr(Majority to ECMO)	72%	36%
Pr(5 ECMO & 5 CMT)	14%	27%
Pr(Majority to CMT)	14%	36%
Pr(Fisher P-value < 5%)	12%	0.1%
Pr(Chi-square P-value < 5%)	32%	1.9%
Pr(# ECMO Success > # CMT Successes)	89%	38%
Pr(# ECMO Success ≥ # CMT Success + 4)	59%	2.7%

Power

Type I error

ECMO: Prospective Simulation

Operating Characteristics	CMT 25% ECMO 75%	CMT 25% ECMO 25%
Pr(All patients randomized to ECMO)	2.5%	0.04%
Pr(All patients randomized to CMT)	0.04%	0.04%
Pr(Majority to ECMO)	72%	36%
Pr(5 ECMO & 5 CMT)	14%	27%
Pr(Majority to CMT)	14%	36%
Pr(Fisher P-value < 5%)	12%	0.1%
Pr(Chi-square P-value < 5%)	32%	1.9%
Pr(# ECMO Success > # CMT Successes)	89%	38%
Pr(# ECMO Success ≥ # CMT Success + 4)	59%	2.7%

Power

Type I error

ECMO Iterate Design

N	Decision Rule # ECMO Successes vs. # CMT Successes	Power when ECMO 75% CMT 25%	Type I error ECMO 25% CMT 25%
10	1 or more	89%	38%
10	4 or more	59%	2.7%
10	3 or more	72%	8.1%

ECMO Iterate Design

N	Decision Rule # ECMO Successes vs. # CMT Successes	Power when ECMO 75% CMT 25%	Type I error ECMO 25% CMT 25%
10	4 or more	59%	2.7%
10	3 or more	72%	8.1%
15	4 or more	79%	5.9%
15	5 or more	71%	2.3%

ECMO Iterate Design

N	Decision Rule # ECMO Successes vs. # CMT Successes	Power when ECMO 75% CMT 25%	Type I error ECMO 25% CMT 25%
10	4 or more	59%	2.7%
10	3 or more	72%	8.1%
15	4 or more	79%	5.9%
15	5 or more	71%	2.3%
16	4 or more	82%	6.7%
16	5 or more	74%	2.8%

ECMO Iterate Design

N	Decision Rule # ECMO Successes vs. # CMT Successes	Power when ECMO 75% CMT 25%	Type I error ECMO 25% CMT 25%
10	4 or more	59%	2.7%
10	3 or more	72%	8.1%
15	4 or more	79%	5.9%
15	5 or more	71%	2.3%
16	4 or more	82%	6.7%
16	5 or more	74%	2.8%
18	5 or more	80%	3.5%

Fisher's exact test: 59% power @ 1-sided 5.0%.

ECMO Iterate Design

N	Decision Rule ECMO v CMT	Power 75v25	ECMO S/N	CMT S/N	T1error 25v25	ECMO S/N	CMT S/N
10	4 or more	59%	4.9 / 6.5	0.9 / 3.5	2.7%	1.25 / 5	1.25 / 5
10	3 or more	72%			8.1%		
18	5 or more	80%	9.2 / 12.2	1.4 / 5.8	3.5%	2.25 / 9	2.25 / 9

8 more patients

5.7 more

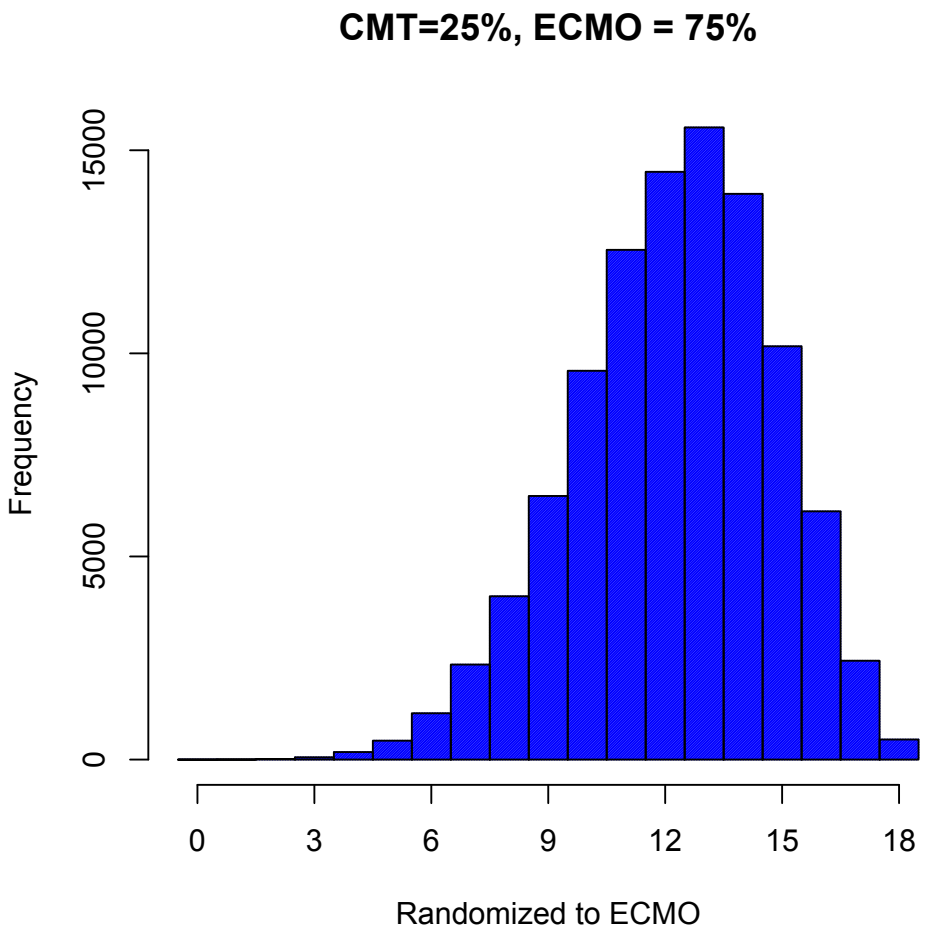
2.3 more

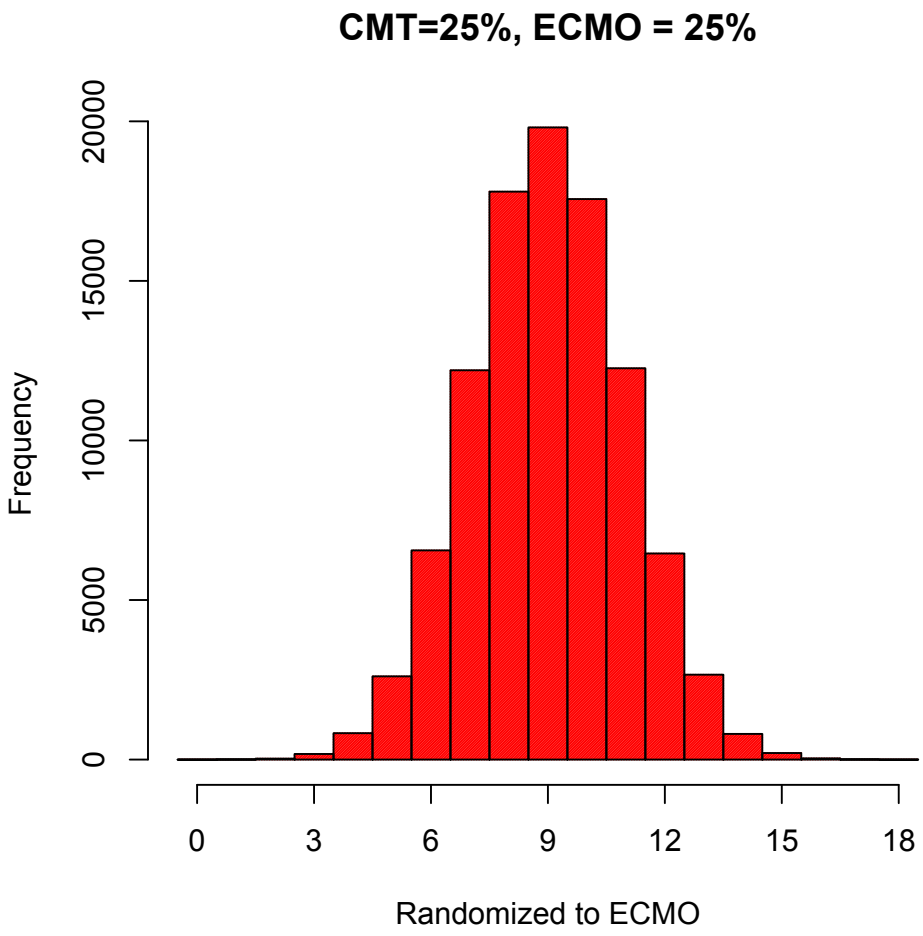
4 more

4 more

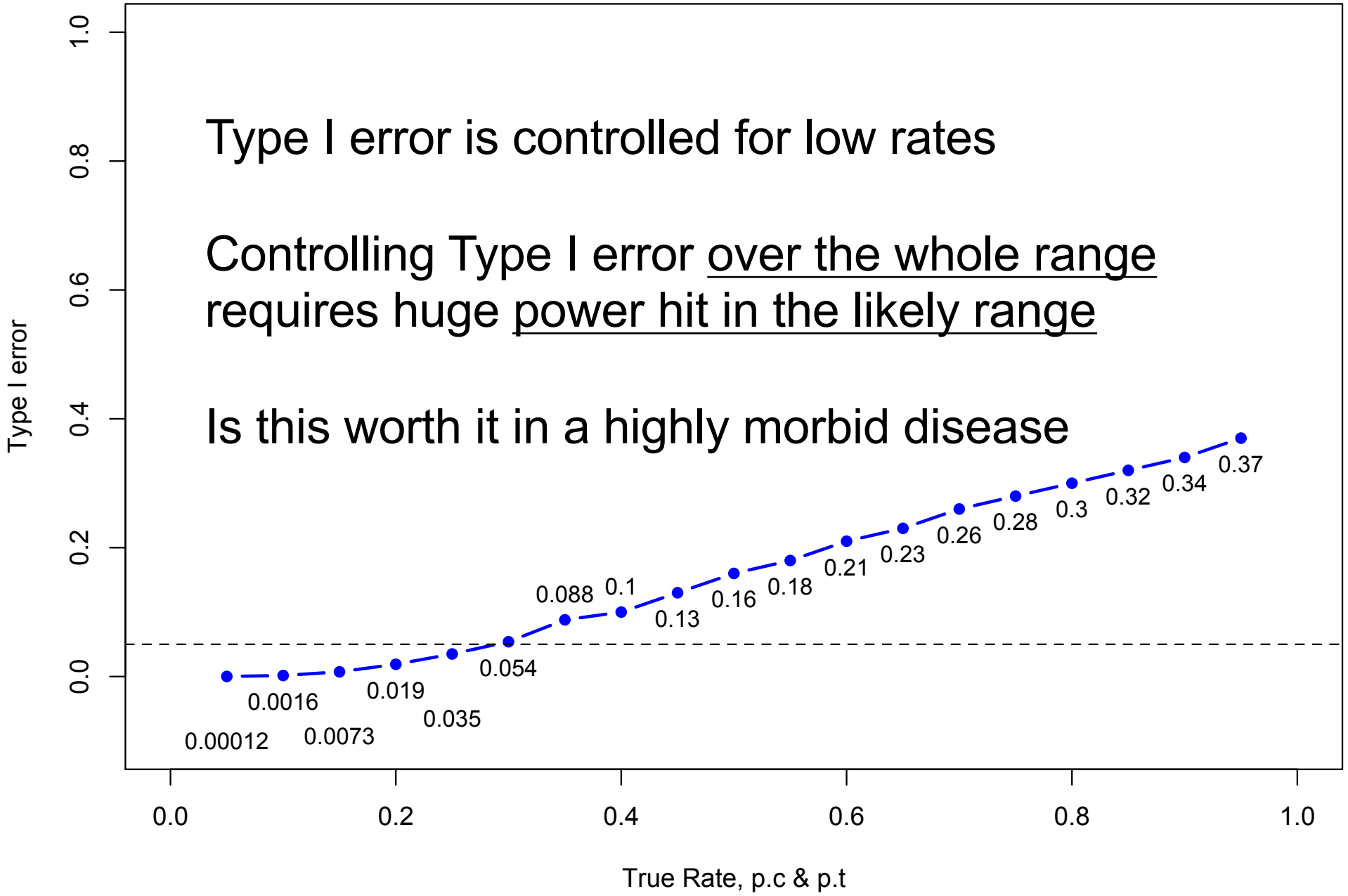
Standard trial with 18 patients has 58% power with 5% Type I error
Always randomized half to CMT; E(survive) = 10.6 vs. 9

ECMO with 18 patients





No free lunch



When designing trials I believe we should

- Remember that most 'standard' methods were developed for agriculture
- Remember that current trialists were trained by people who were trained by people who had seeds as patients
- Remember most statistical methodology is based on asymptotic theory
  - Because we couldn't do math then that we can do now
- Forget much of what we know about clinical trials & hypothesis testing & asymptotic theory
- Hire smart people with their heart in the right place
- Balance treating the next patient well & producing valuable long-term evidence
- Think much harder about the 'right' Type I error rate
- Design trials by trial & error by using simulation, iterate designs with doctors, patients, payers, regulators
- Not let within-trial patient benefit be a side effect of quality research

When designing trials I believe we should

- Remember that most 'standard' methods were developed for agriculture
- Remember that current trialists were trained by people who were trained by people who had seeds as patients
- Remember most statistical methodology is based on asymptotic theory
  - Because we couldn't do math then that we can do now
- Forget much of what we know about clinical trials &
- hypothesis testing & asymptotic theory
- Hire smart people with their heart in the right place
- Balance treating the next patient well & producing valuable long-term evidence
- Think much harder about the 'right' Type I error rate
- Design trials by trial & error by using simulation,
- iterate designs with doctors, patients, payers, regulators
- Not let within-trial patient benefit be a side effect of quality research

Decision Problem #4:
ESETT Trial

A multicenter, randomized, double-blind,

comparative effectiveness study of

fos-phenytoin, levetiracetam, and valproic acid

in subjects with benzodiazepine-refractory

Status Epilepticus:

The Established Status Epilepticus Treatment

Trial

Acknowledgements

ADAPT-IT U01-NS073476

- William Barsan, MD
  University of Michigan
- Donald Berry, PhD
  MD Andersson, Berry Consultants
- Roger Lewis, MD, PhD
  UCLA Harbor
- Scott Berry, PhD
  Berry Consultants
- Valerie Durkalski, PhD, Yuko Palesch, PhD
  Medical University of South Carolina
- Michael Fetters, MD
  University of Michigan
- Shirley Fredrickson, RN, MS
  University of Michigan
- Will Meurer, MD
  University of Michigan
- Robin Conwit, MD, Scott Janis, MD
  NINDS

ESETT U01-NS088034

- Jaideep Kapur, MD
  University of Virginia
- Kristine Broglio, MS
  Berry Consultants
- Jordan Elm, PhD, Wenle Zhao, PhD
  Medical University of South Carolina
- James Chamberlain, MD
  Children's National Med Center
- Nathan Fountain, MD
  University of Virginia
- Daniel Lowenstein, MD
  UCSF
- Shlomo Shinnar, MD, PhD
  Albert Einstein COM
- Rob Silbergliet, MD
  University of Michigan
- David Treiman, MD
  Barrow Neurological Institute

Research Question

- How to treat seizing patients who've failed benzodiazapine?
  - fosphenytoin (fPHT)
  - levetiracetam (LVT)
  - valproic acid (VPA)

Comparative Effectiveness

- No control group
  - Three drugs start out equal
  - Want to know which is best
- What is Type I error in CER?
  - Consequence of Type I error less in CER
- Really want to know
  - Which drug is best … with measure of certainty
  - Which drug is worst … with measure of certainty

Trial Overview

- Primary endpoint
  - cessation of seizure within 20 minutes
  - no further intervention within 1 hour
  - no significant adverse event
- Powered to identify 15% difference in response rate
  - Min 400, Max 795 Patients (to get 720)
- Stratify randomization by age

Bayesian Adaptive Design Features

- Adaptively allocate to favor better treatments
- Drop poor performing arms
  - Relative to one another
  - Relative to 25% goal
- Stop early if we know the answer

  or know we won't know
  - Efficacy stop if treatment clearly better
  - Futility stop if unlikely to ID a 'best' or 'worst'
    - Do not stop if 1 worse and other 2 equally good
  - Futility stopping if all arms bad

Adaptive Allocation

- Randomize 300 patients equally
- At 300 & then every 100 adaptively allocate to

$$r_t \propto \sqrt{\frac{\Pr\left(p_t = \max(p)\right) Var(p_t)}{n_t}}$$

  - Favor better performing treatments
  - Favor treatments with greater uncertainty
  - Every 100 = About every 6 months | expected accrual
- If allocation probability < 5%, suspend accrual
- If Pr(Success > 0.25) < 0.05 drop arm

Early Stopping

- Analyses begin after 400 patients and repeat every additional 100 patients accrued
- Early Success Stopping:
  - If arm has 97.5% probability of having highest success rate
    - i.e. $\Pr(p_t = max(p)) > 0.975$
- Early Futility Stopping
  - If all doses have Pr(Success > 0.25) < 0.05
  - If predicted probability of success (ID 'winner' or 'loser' at the max N=795) < 0.05

Example Trial: 300 pt analysis

	N Enrolled Observed Response Rate			Pr(Max Effective Trt)			Pr(Allocation)			Pred Prob
Look	LVT	fPHT	VPA	LVT	fPHT	VPA	LVT	fPHT	VPA	
300	51/100 51%	55/100 55%	64/100 64%	0.025	0.092	0.88	0.12	0.22	0.66	0.71

Example Trial: 400 pt analysis

	N Enrolled Observed Response Rate			Pr(Max Effective Trt)			Pr(Allocation)			Pred Prob
Look	LVT	fPHT	VPA	LVT	fPHT	VPA	LVT	fPHT	VPA	
300	51/100 51%	55/100 55%	64/100 64%	0.025	0.092	0.88	0.12	0.22	0.66	0.71
Next 100	6/11 55%	19/26 73%	39/63 62%							
400	57/111 51%	74/126 59%	105/163 64%	0.01	0.16	0.83	0.09	0.34	0.57	0.50

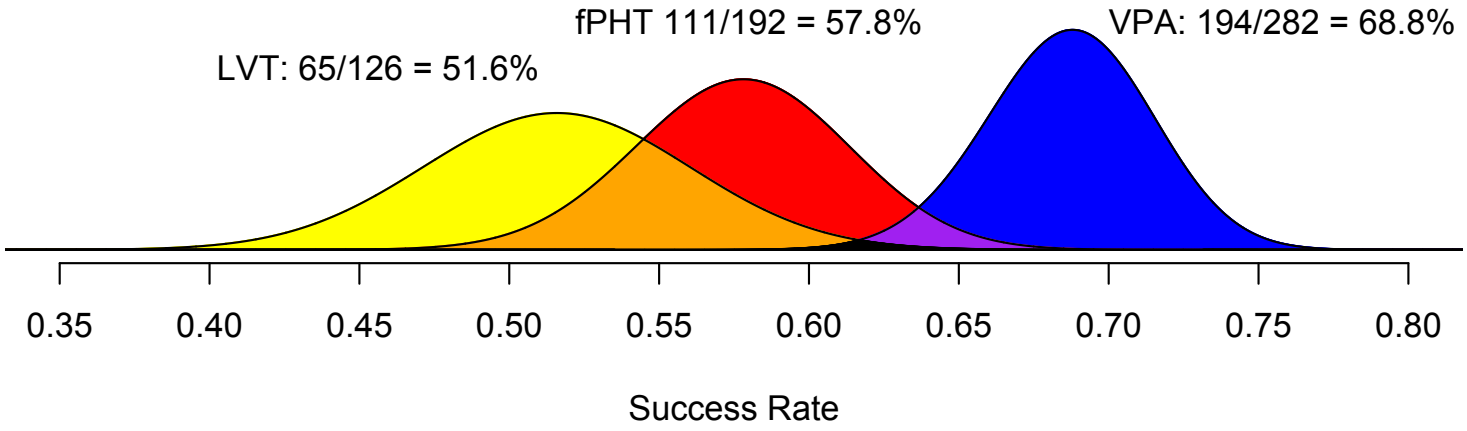
Example Trial: 500 pt analysis

	N Enrolled Observed Response Rate			Pr(Max Effective Trt)			Pr(Allocation)			Pred Prob
Look	LVT	fPHT	VPA	LVT	fPHT	VPA	LVT	fPHT	VPA	
300	51/100 51%	55/100 55%	64/100 64%	0.025	0.092	0.88	0.12	0.22	0.66	0.71
400	57/111 51%	74/126 59%	105/163 64%	0.01	0.16	0.83	0.09	0.34	0.57	0.50
Next 100	5/12 42%	20/38 53%	34/50 68%							
500	62/123 50%	94/164 57%	139/213 65%	0.004	0.056	0.94	0.08	0.23	0.69	0.59

Example Trial: 600 pt analysis

	N Enrolled Observed Response Rate			Pr(Max Effective Trt)			Pr(Allocation)			Pred Prob
Look	LVT	fPHT	VPA	LVT	fPHT	VPA	LVT	fPHT	VPA	
300	51/100 51%	55/100 55%	64/100 64%	0.025	0.092	0.88	0.12	0.22	0.66	0.71
400	57/111 51%	74/126 59%	105/163 64%	0.01	0.16	0.83	0.09	0.34	0.57	0.50
500	62/123 50%	94/164 57%	139/213 65%	0.004	0.056	0.94	0.08	0.23	0.69	0.59
Next 100	3/3 100%	17/28 61%	55/69 80%							
600	65/126 52%	111/192 58%	194/282 69%	0.000 0.87	0.008 0.13	0.992 0.00	Trial Stops Early for Identifying Best Treatment			

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>LVT) = 0.993
VPA – LVT	0.172	(0.069, 0.272)	Pr(VPA>fPHT) > 0.999
fPHT - LVT	0.062	(-0.049, 0.172)	Pr(LVT>fPHT) = 0.862

Comparison to without Adaptive Randomization

	Adaptive Randomization			Fixed Randomization		
Scenario 3 Efficacy Rates	**Power Best/Wst**	**Mean N**	**% to Best**	**Power Best/Wst**	**Mean N**	**% to Best**
Null 0.5 – 0.5 – 0.5	0.013 0.018	507		0.023 0.007	499	
One Good 0.5 – 0.5 – 0.65	0.89 0.03	483	48	0.87 0.04	497	33
Two Good 0.5 – 0.65 – 0.65	0.11 0.67	679	84	0.10 0.79	687	67
One Middle One Good 0.5 – 0.575 – 0.65	0.50 0.25	586	47	0.44 0.31	599	33
All Bad 0.25– 0.25 – 0.25	0.011 0.020	524		0.023 0.008	509	
All Very Bad 0.10 – 0.10 – 0.10	0.006 0.01	400		0.008 0.02	400	

Comparison to without Adaptive Randomization

	Adaptive Randomization			Fixed Randomization		
Scenario 3 Efficacy Rates	**Power Best/Wst**	**Mean N**	**% to Best**	**Power Best/Wst**	**Mean N**	**% to Best**
Null 0.5 – 0.5 – 0.5	0.013 0.018	507		0.023 0.007	499	
One Good 0.5 – 0.5 – 0.65	0.89 0.03	483	48	0.87 0.04	497	33
Two Good 0.5 – 0.65 – 0.65	0.11 0.67	679	84	0.10 0.79	687	67
One Middle One Good 0.5 – 0.575 – 0.65	0.50 0.25	586	47	0.44 0.31	599	33
All Bad 0.25– 0.25 – 0.25	0.011 0.020	524		0.023 0.008	509	
All Very Bad 0.10 – 0.10 – 0.10	0.006 0.01	400		0.008 0.02	400	

Comparison to without Adaptive Randomization

	Adaptive Randomization			Fixed Randomization		
Scenario 3 Efficacy Rates	**Power Best/Wst**	**Mean N**	**% to Best**	**Power Best/Wst**	**Mean N**	**% to Best**
Null 0.5 – 0.5 – 0.5	0.013 0.018	507		0.023 0.007	499	
One Good 0.5 – 0.5 – 0.65	0.89 0.03	483	48	0.87 0.04	497	33
Two Good 0.5 – 0.65 – 0.65	0.11 0.67	679	84	0.10 0.79	687	67
One Middle One Good 0.5 – 0.575 – 0.65	0.50 0.25	586	47	0.44 0.31	599	33
All Bad 0.25– 0.25 – 0.25	0.011 0.020	524		0.023 0.008	509	
All Very Bad 0.10 – 0.10 – 0.10	0.006 0.01	400		0.008 0.02	400	

Comparison to without Adaptive Randomization

	Adaptive Randomization			Fixed Randomization		
Scenario 3 Efficacy Rates	**Power Best/Wst**	**Mean N**	**% to Best**	**Power Best/Wst**	**Mean N**	**% to Best**
Null 0.5 – 0.5 – 0.5	0.013 0.018	507		0.023 0.007	499	
One Good 0.5 – 0.5 – 0.65	0.89 0.03	483	48	0.87 0.04	497	33
Two Good 0.5 – 0.65 – 0.65	0.11 0.67	679	84	0.10 0.79	687	67
One Middle One Good 0.5 – 0.575 – 0.65	0.50 0.25	586	47	0.44 0.31	599	33
All Bad 0.25– 0.25 – 0.25	0.011 0.020	524		0.023 0.008	509	
All Very Bad 0.10 – 0.10 – 0.10	0.006 0.01	400		0.008 0.02	400	

Conclusions for Adaptive Designs in Comparative Effectiveness Research

- Adaptive trials / adaptive CER processes more closely mimic real-life human learning & decision making
- Ongoing projects: Learn & Adapt
  - randomize patients to best products
  - drop treatments/strategies that prove less effective
  - include new treatments as they come to market
  - provide constant sharing of information
  - encourage better patient management

Adaptive Randomization[1]

- Pros
  - Resolve conflict of healer vs. investigator
  - Maximize number of patients assigned more effective therapy
  - Consistent with current theories of continuous quality improvement
- Cons
  - Must be one (or few) outcome(s) of interest
  - Outcomes must be apparent in a short timeframe relative to accrual time
  - May be statistically less efficient
  - Estimates affected by population drift during accrual

[1] Used with permission, Robert Truog, http://www.bioethics.nih.gov/slides04/truog.ppt

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

Equipoise

- Would you rather be the last patient enrolled in a clinical trial or the first person treated after its results are published?

- Declaration of Helsinki:
  - "considerations related to the well-being of the human subject should take precedence over the interests of science and society"

FDA Critical Path Initiative

From FDA website:

> Many of the tools used today to predict and evaluate product safety and efficacy are badly outdated from a scientific perspective. We have not made a concerted effort to apply new scientific knowledge -- in areas such as gene expression, analytic methods, and bioinformatics -- to medical product development. There exists tremendous opportunities to create more effective tests and tools, if we focus on the hard work necessary to turn these innovations into reliable applied sciences.

http://www.fda.gov/scienceresearch/specialtopics/criticalpathinitiative/ucm077015.htm

FDA Critical Path Initiative

From FDA website:

Inefficient clinical trial designs. Innovative clinical trial design may make it possible to develop accepted protocols for smaller but smarter trials. For example, new statistical techniques may make it possible to reduce the number of people who need to receive placebo or to adaptively change the trial based on ongoing results.

50% of Phase 3 trials failing

$800 million per successful NME (new chemical entity)

Ann. Rev. Medicine, Woodcock & Woosley, 2008

Critical Path Initiative

- Areas of improvement
  - Development & use of biomarkers (for prediction)
    toward personalized medicine
  - Modernizing clinical trial methodologies & processes
  - Aggressive use of bioinformatics
    including disease modeling & trial simulation
  - Improvement in manufacturing technologies
  - 76 discrete projects that could improve product development & product use

US FDA 2006, "Innovation or Stagnation: Critical Path Opportunities Report & List."

www.fda.gov/oc/initiatives/criticalpath/reports/opp_report.pdf

Historical Context

- Historically, obtaining results that were "reliable and valid" required fixed study designs
- Allowed the determination of theoretical error rates
- Fundamental characteristic of the "culture" of biostatistics and clinical trial methodology

Why are Study Designs 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
- We could do the math 30 years ago
  - We still can but we can also do more way sophisticated calculations now

What are Adaptive Designs?

- Adaptive Design:
  - A design that "changes" depending on observed values in the trial
- Prospective Adaptive Design:
  - A design that has pre-specified dynamic aspects that are determined by the accruing information

Every time I say "Adaptive Design" I mean

"Prospectively Adaptive Design"

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

What are Adaptive Trials?

Trials that change based on prospective rules & the accruing information

- Adaptive sample sizes based on predictive probabilities
  - Stop early for success
  - Terminate early for futility
- Adaptive randomization
  - For statistical efficiency
  - For improved patient treatment
  - Drop/Re-enter arms or dose groups
- Adaptive accrual rate
- Combination therapies
- Adapt to responding sub-populations
- Adaptive borrowing of information
- Seamlessly combine phases of development
  - Phase 2/3 designs: Operationally vs. Inferentially seamless

Key Design Features

- Frequent interim analyses
- Predefined decision rules for adaptations
- Explicit longitudinal modeling of the accumulating data based upon interim outcomes

- Response-adaptive randomization
- Dose-response modeling using information from all patients
- Extensive simulation of trial performance
- Repeatedly ask when are primary questions answered

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

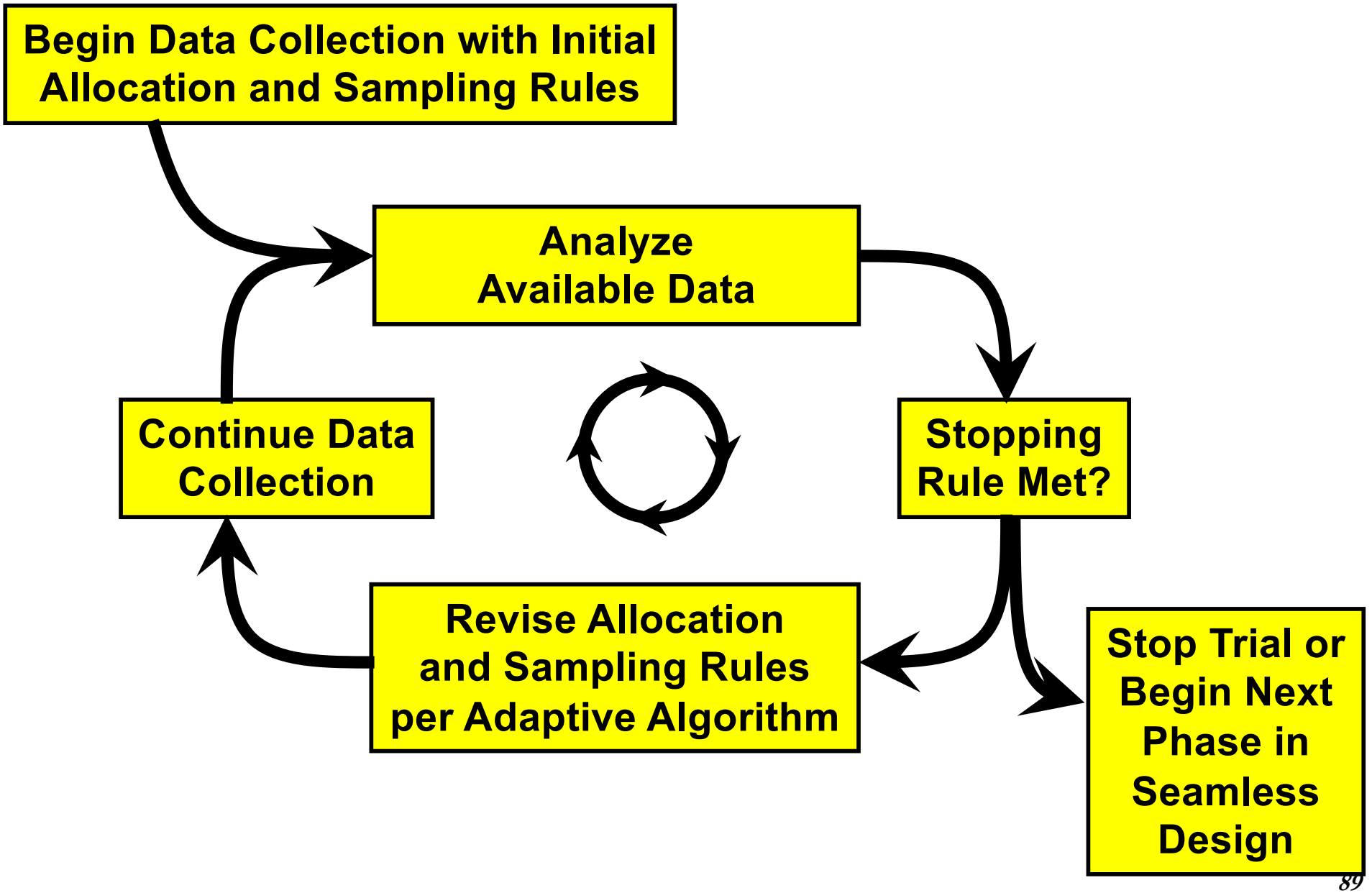
Drawbacks of Adaptation

- Infeasible if time from patient accrual to final outcomes long vs. total accrual time
- Adaptive design take much more forethought & buy-in from more stakeholders
- Determining traditional Type I and II error rates more difficult
  - Rely on simulation
- People fear new
  - Most statisticians have never designed or analyzed an adaptive trial
  - Some regulatory personnel unfamiliar with
  - Funders (e.g. venture capitalists and NIH) unfamiliar with
  - DMCs / IRBs may not understand
  - Clinicians may not understand

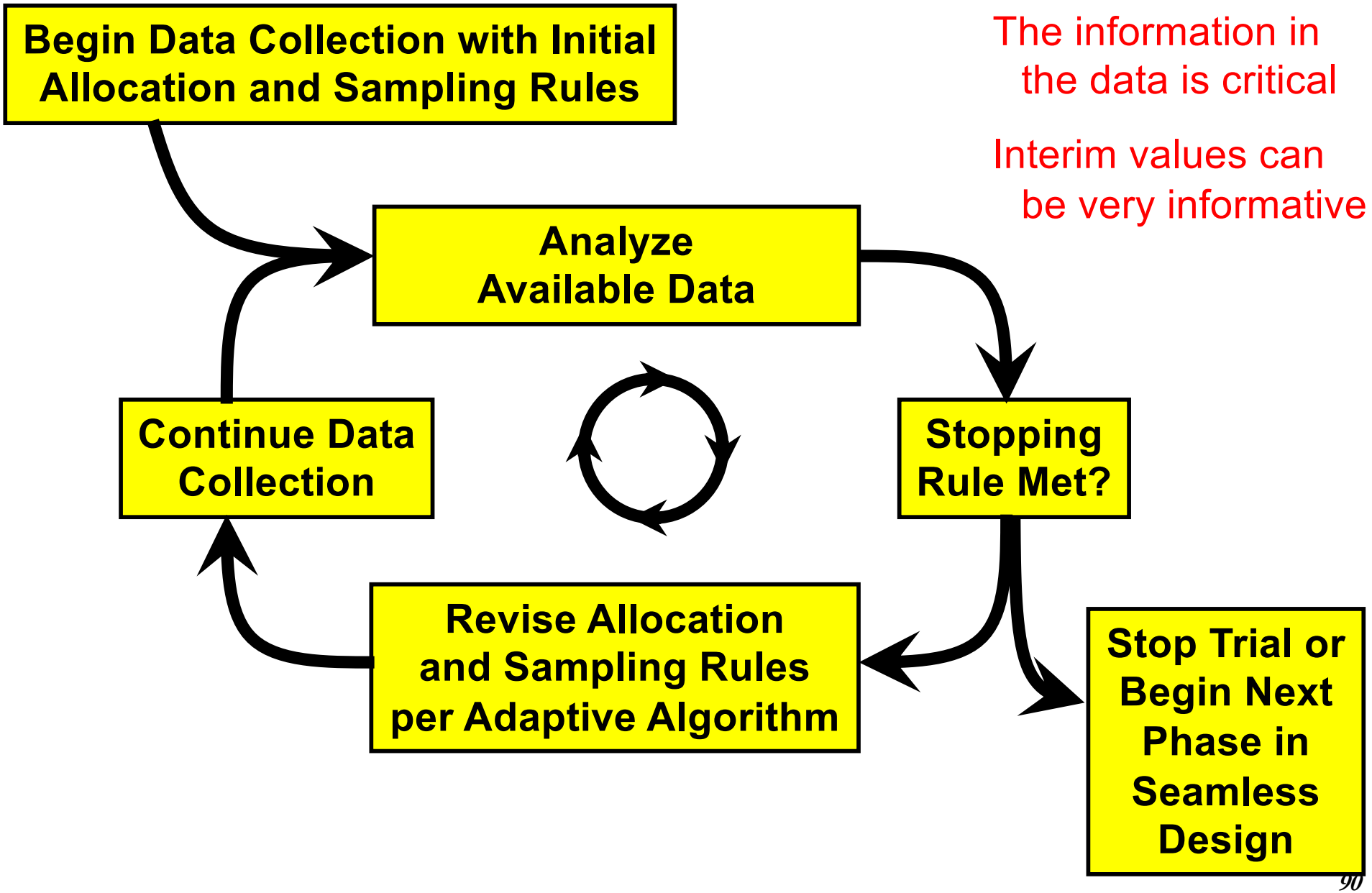
Drawbacks of Adaptation

- Logistical issues
  - Design stage is longer
  - Data needs to be entered & transmitted quickly
  - Data needs to be checked / validated quickly
  - Events need to be adjudicated quickly
  - Drug supply concerns for adaptive randomization
    - Fear of unblinding
  - Need centralized randomization
    - Use web or phone systems
  - Need to have lots of people / systems well & correctly connected

Typical Prospective Adaptive Design



Typical Prospective Adaptive Design



JAMA 2006;296:1955-1957.

Who To Involve

- Sponsor
  - Project leaders
  - Statisticians
  - PK/PD
  - Clinical experts
  - Business leaders
  - Patient advocates
- Clinical site IRBs
- Data Safety Monitoring Board
- IVRS/IWRS service
- CRO who will house data
- Regulatory agencies
- Patient advocacy groups?
  - Treat patients in trial best vs. get drug to market sooner?
- Payers

Adaptive Designs & Collaborators

- Requires buy-in and educating IRB, DSMB, decision-makers, study teams, investigators, and subjects
- Requires more time, resources, and upfront planning, especially at the protocol-design stage
- Show sponsor many many example trials
  - Also great for debugging
- Complex study designs typically require more statistical assumptions, rigorous calculations, and extensive simulations (operating characteristics)
- But also more robust to deviations from our assumptions
- Operationally challenging
  - Work with CROs as early as possible, fit statistical parts within infrastructure
- Make sure sponsors understands what adaptive designs are not

Components of an Adaptive Trial

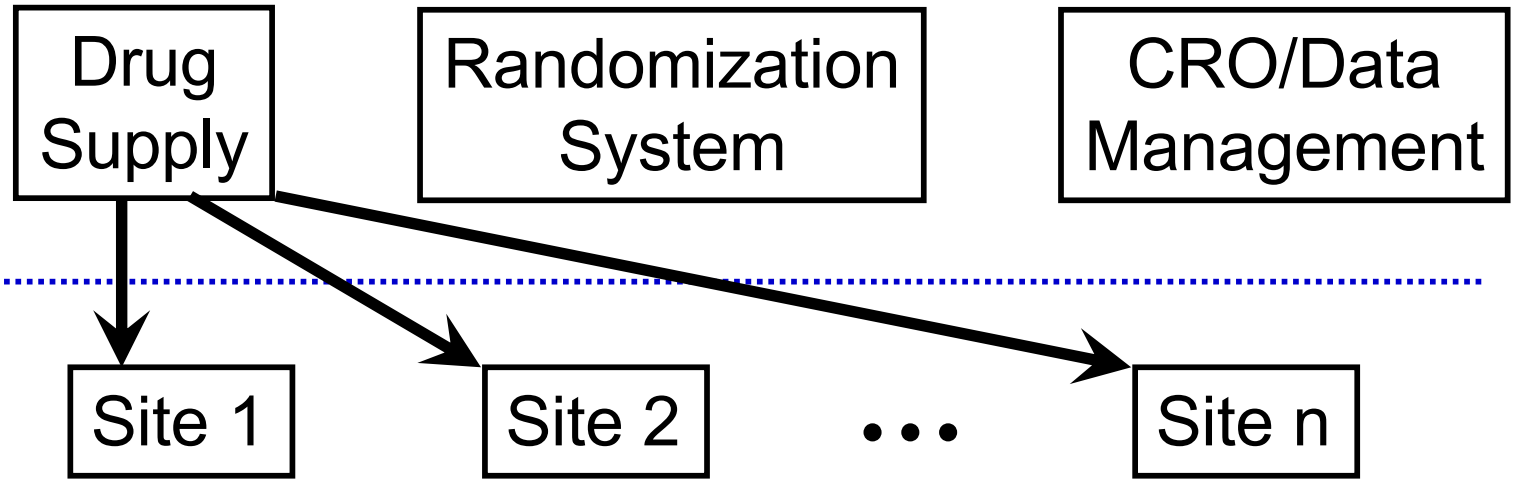
Thanks Roger Lewis

Management

Adaptive Machinery

Logistics

Clinical



Components of an Adaptive Trial

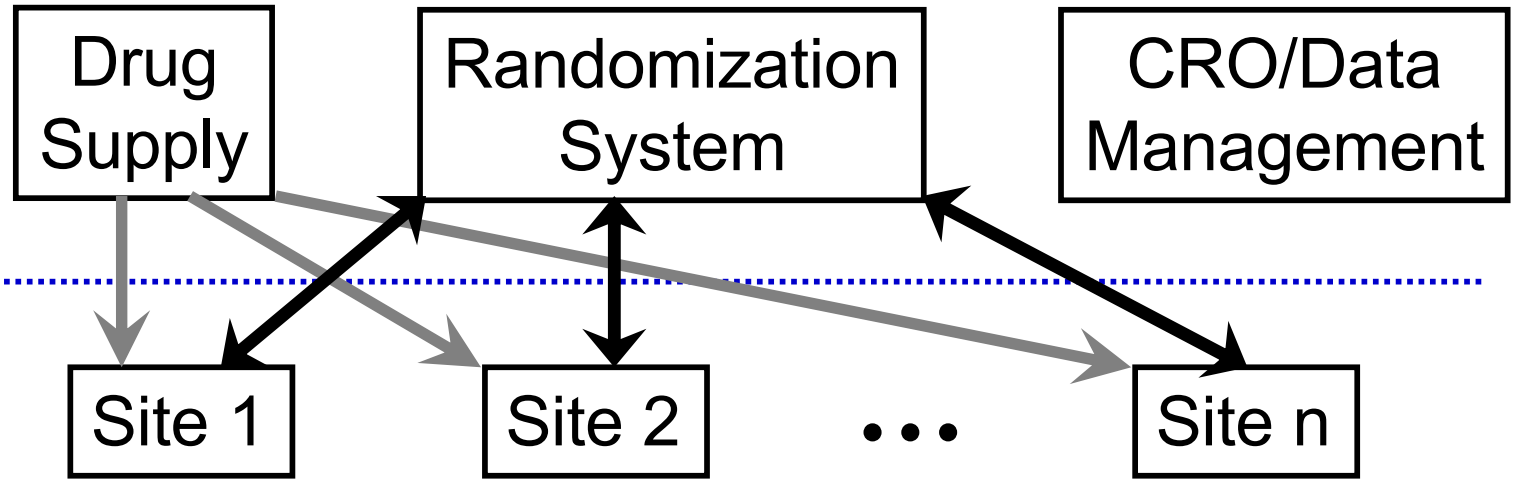
Thanks Roger Lewis

Management

Adaptive Machinery

Logistics

Clinical



Components of an Adaptive Trial

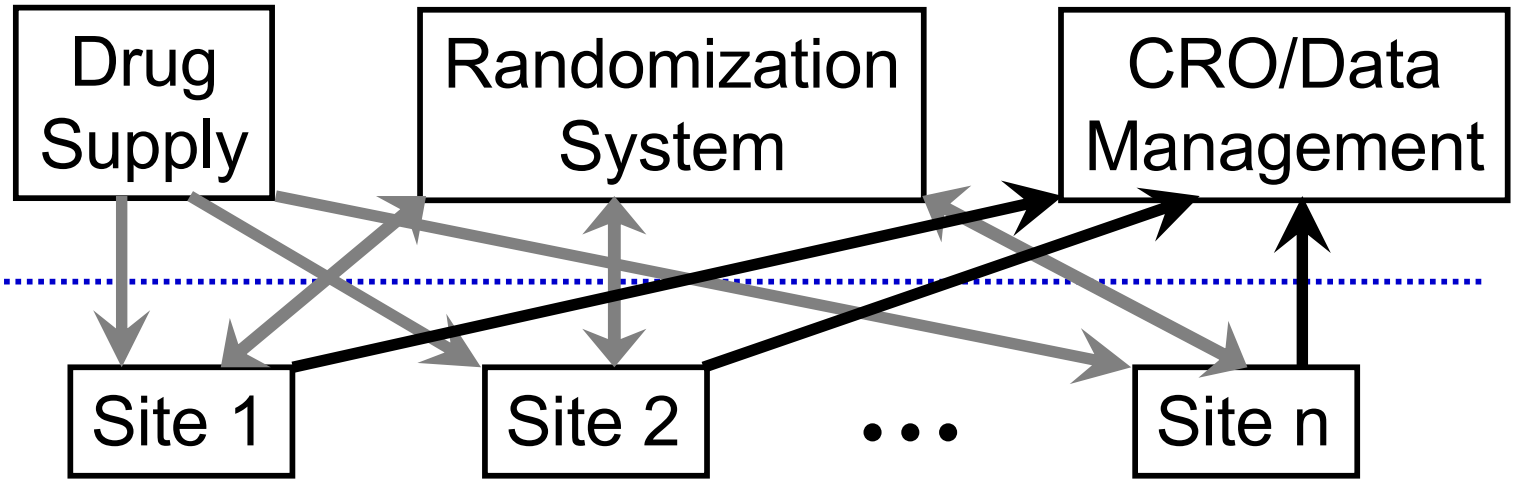
Thanks Roger Lewis

Management

Adaptive Machinery

Logistics

Clinical



Components of an Adaptive Trial

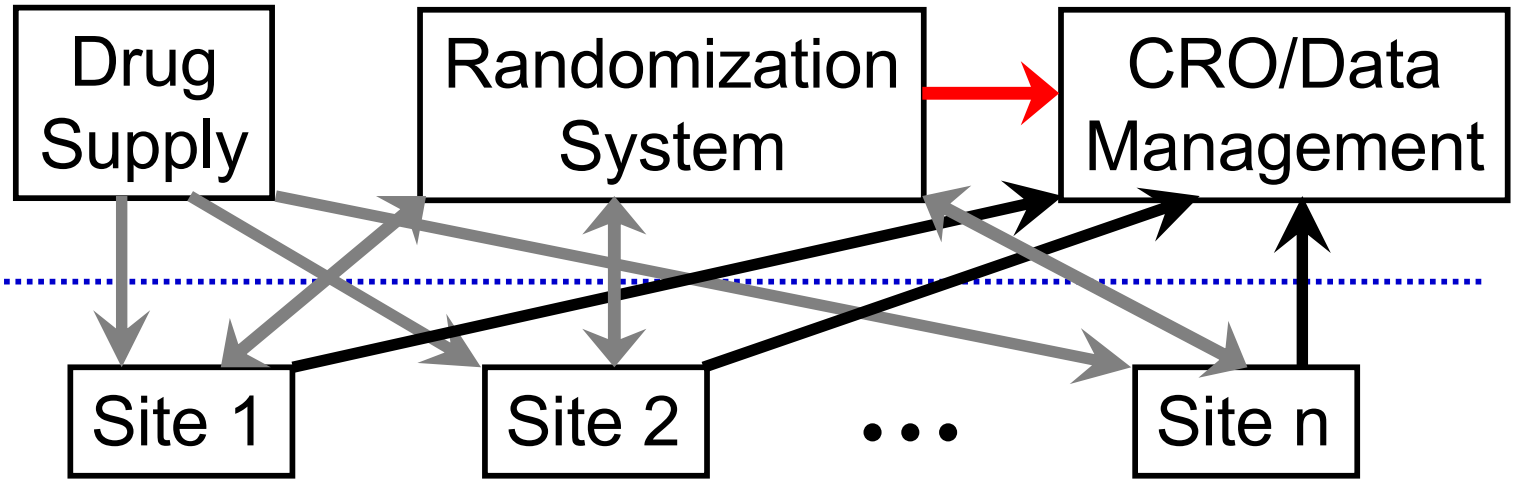
Thanks Roger Lewis

Management

Adaptive Machinery

Logistics

Clinical



Components of an Adaptive Trial

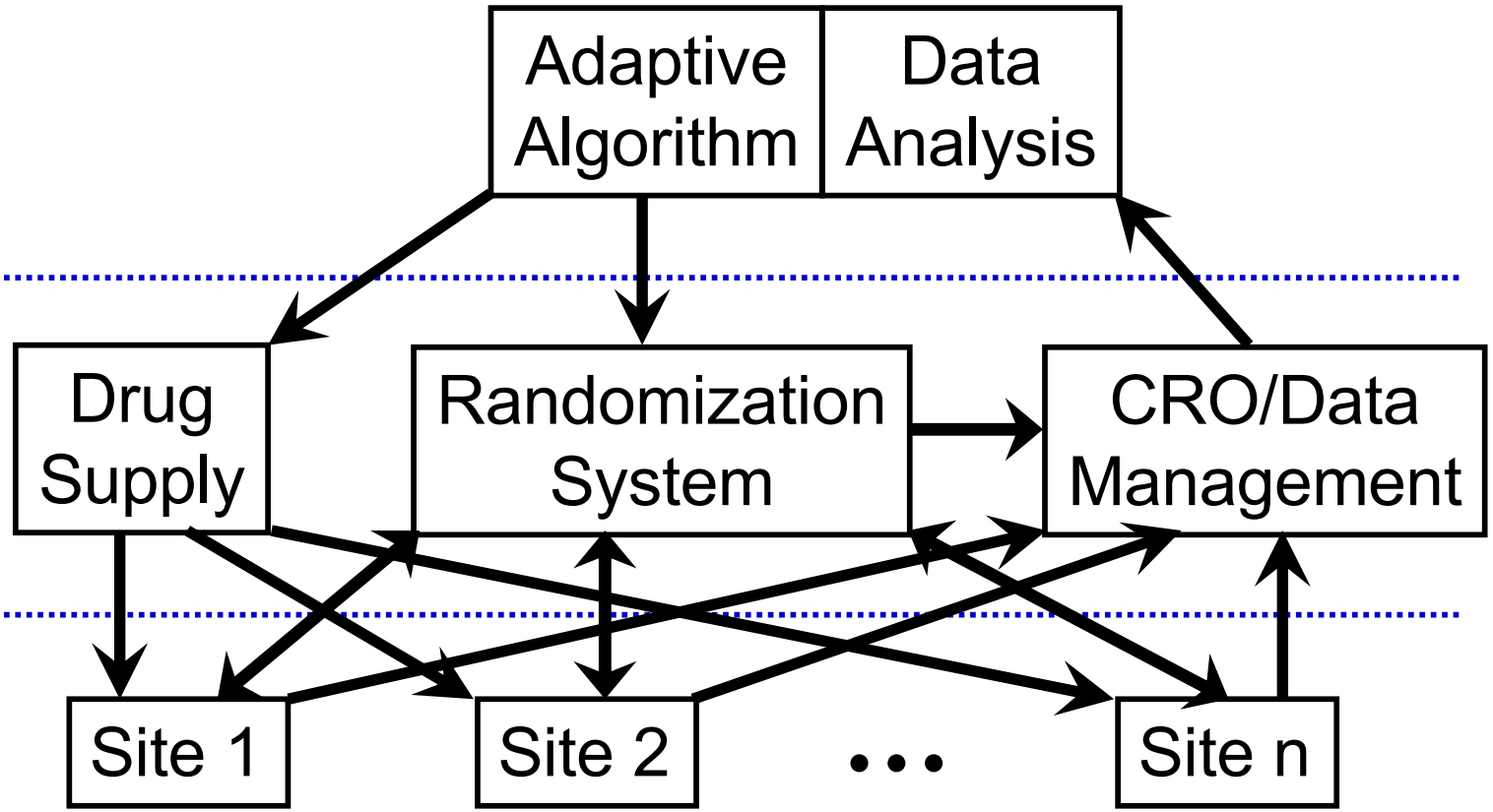
Thanks Roger Lewis

Management

Adaptive Machinery

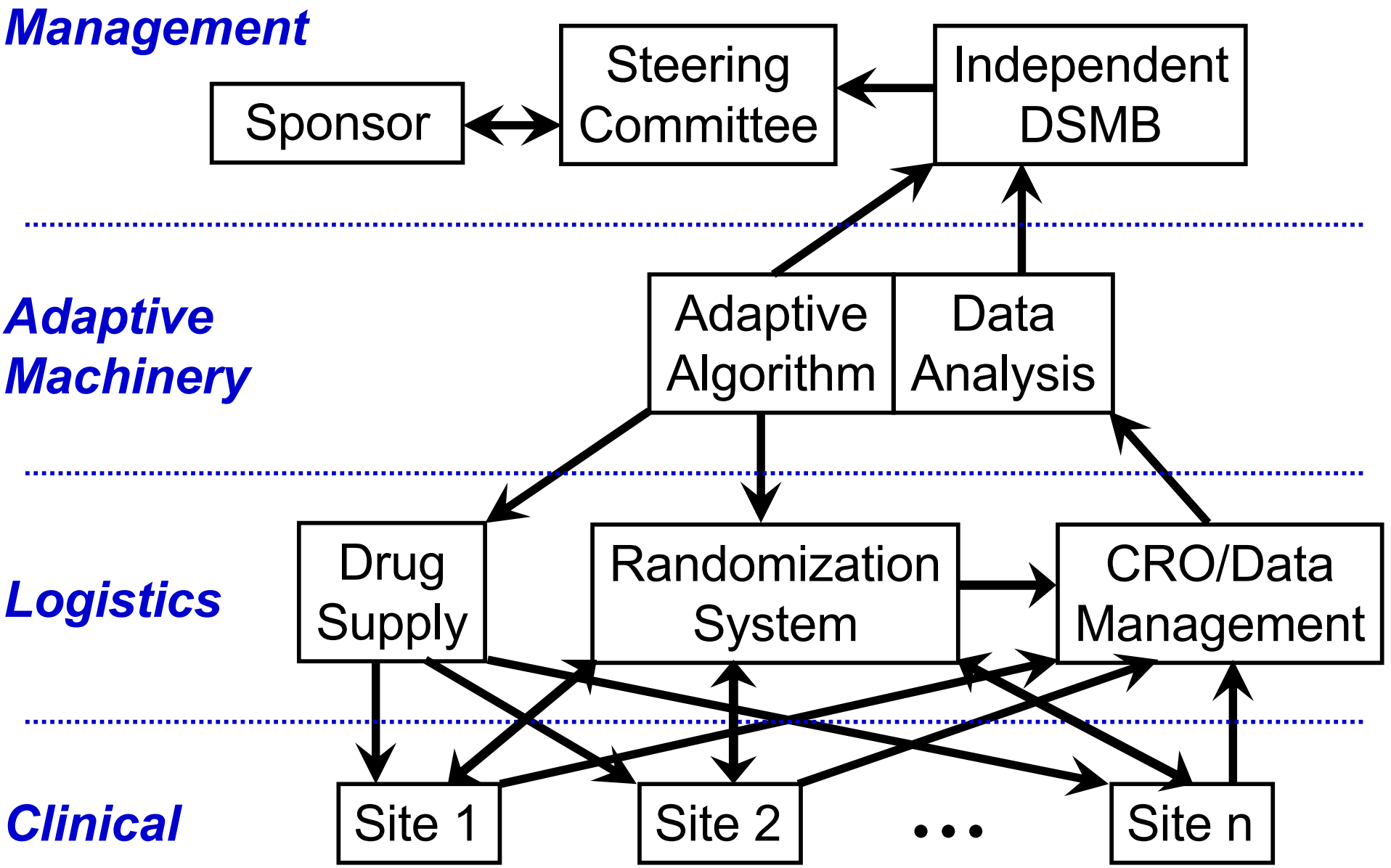
Logistics

Clinical



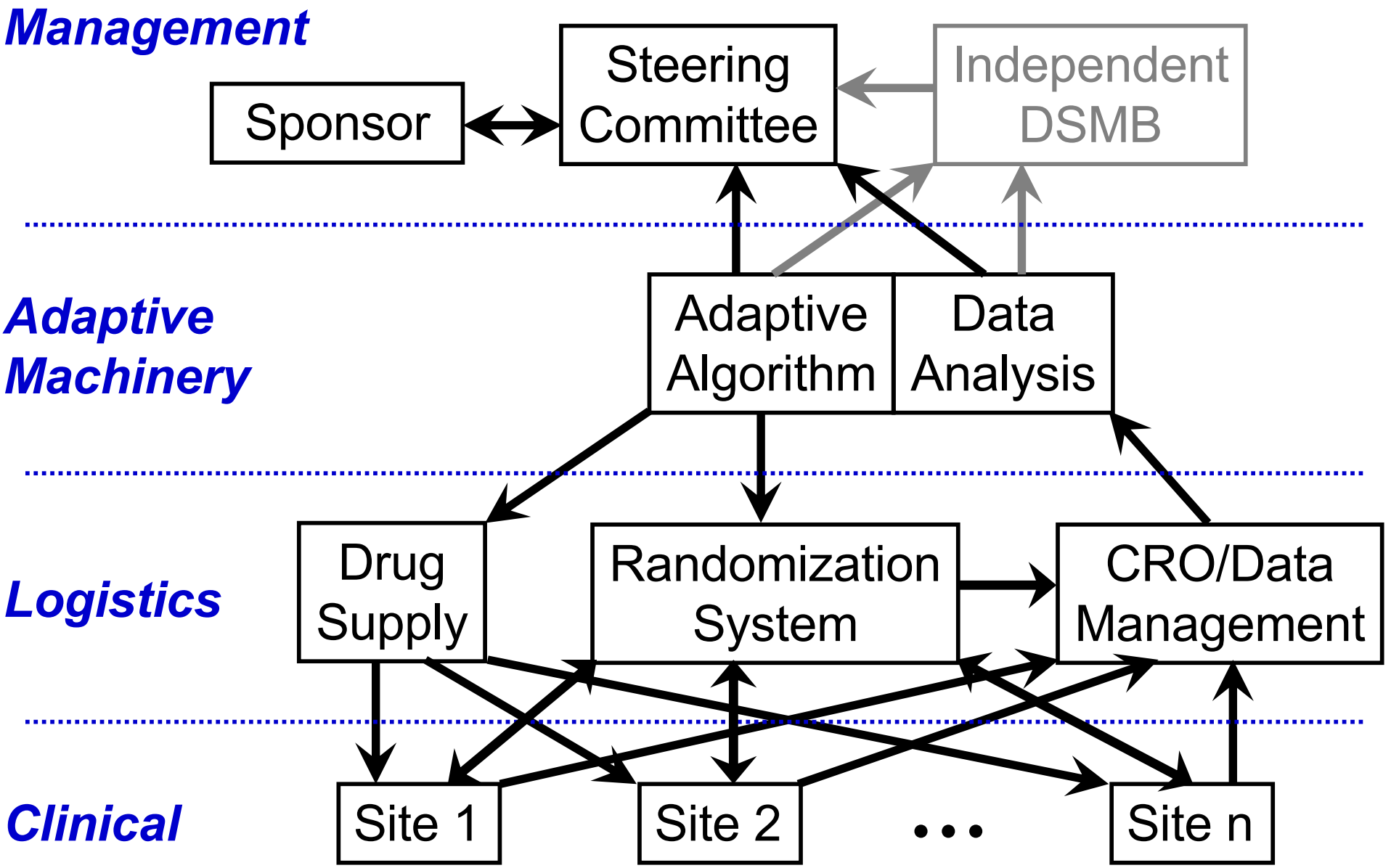
Components of an Adaptive Trial

Thanks Roger Lewis



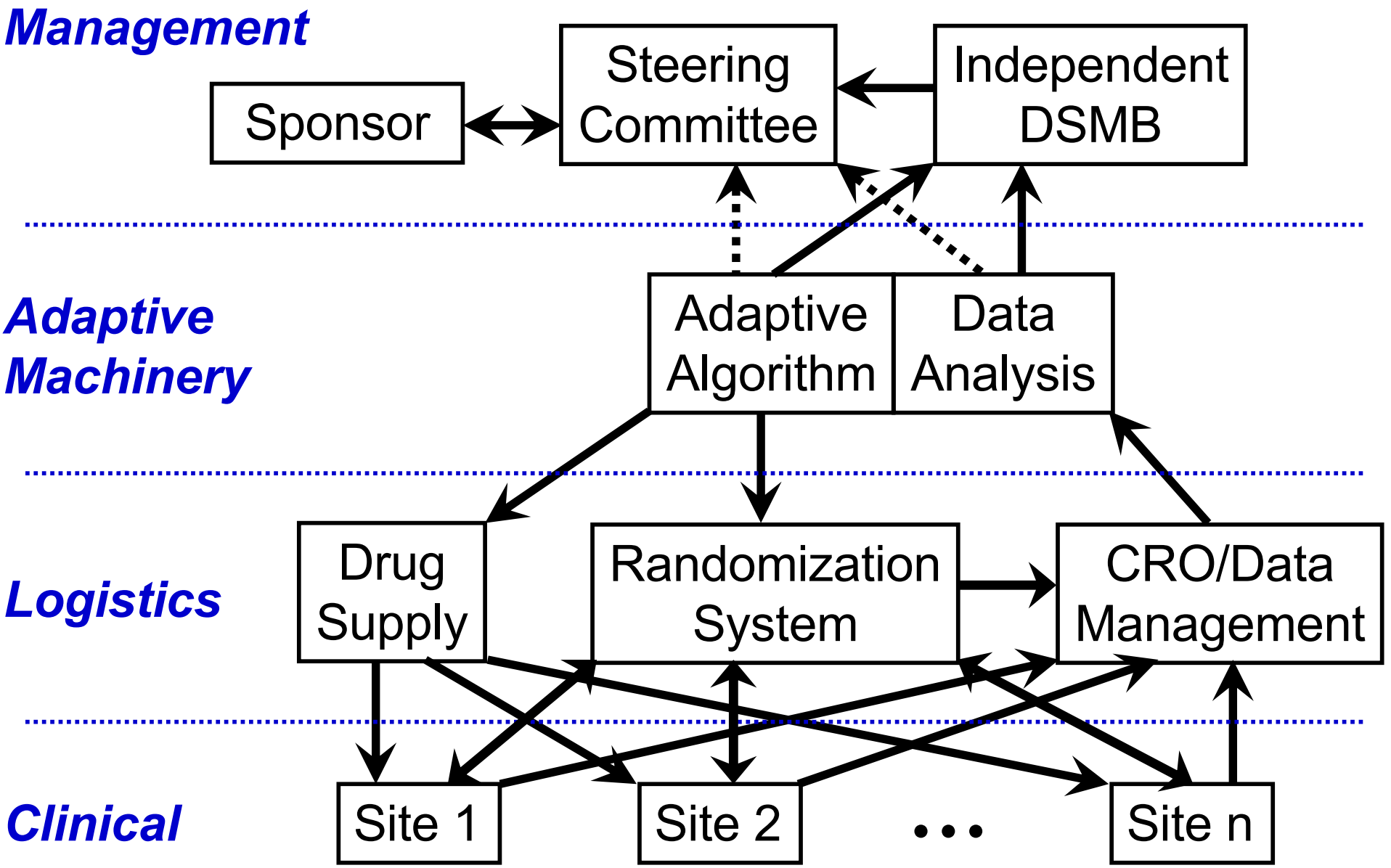
Components of an Adaptive Trial

Thanks Roger Lewis



Components of an Adaptive Trial

Thanks Roger Lewis



Data Safety Monitoring Boards

- Purpose
  - To ensure continued safety, validity, feasibility, and integrity of the clinical trial
  - To ensure the trial is conducted according to *a priori* plan, including adaptation
- Structure
  - Learn phase: usually includes internal personnel
  - Confirm phase: generally includes only independent, external members

Data Safety Monitoring Boards

- What's different in an adaptive trial?
  - Requires expertise to assess whether the planned adaptations continue to be safe and appropriate
  - May increase need to include sponsor personnel
  - Ideally expertise to ensure everything is working
- What's unchanged in an adaptive trial?
  - The DSMB ensures completion of the trial *as planned, including the adaptation*
  - It is the trial that's adaptive, not the DSMB

IRB Review

- IRBs review/approve the full protocol, including the planned adaptations
- No new review when adaptations made
  - IRBs may request to be informed (e.g., new sample size, dropping of a surgical arm)
- Amendments are different
  - Not preplanned
- Irony
  - Little changes (amendments) may require IRB review
  - Big changes (adaptations) are defined by design and only reviewed/approved once

Acceptability to Key Stakeholders

- FDA
  - FDA Critical Path Initiative
  - 2010 Guidance for the Use of Bayesian Statistics in Medical Device Trials
  - 2010 Draft Guidance for Adaptive Design Clinical Trials for Drugs and Biologics
  - Joint Regulatory Science initiative with NIH
  - Multiple adaptive trials accepted in development plans
- PhRMA
  - Highly active "working group" on adaptive trials → DIA
  - 2006 PhRMA/FDA Conference on Adaptive trials
  - Many adaptive trials designed or initiated in industry

Acceptability to Key Stakeholders

- NIH
  - Sponsored Scientific Advances in Adaptive Clinical Trial Designs Workshop, Fall 2009
  - ADAPT-IT sponsored by NIH Common Fund
    - Redesigning four neurologic emergency trials using adaptive designs
  - READAPT sponsored by
- Journals
  - Surprisingly clinical journals care little about design
    - Ever see a medical journal with smaller font for the methods?
  - We've had to argue to let journals give us <u>more</u> space for the design

Is Now a Prime Time for Adaptive Designs in Clinical Trials?

- It's well past time
- Virtually every large pharmaceutical company, 100+ device companies, and dozens of biotech companies are investing in adaptive designs
  - Many device companies have completed adaptive designs
- What is the likelihood that these designs will lead to regulatory approval when such approval is warranted?
- Is there a gap between perceived risk to sponsors and the real risk?
  - Does industry overestimate FDAs conservatism?

Time has been Right for Adaptive Designs

- Janet Woodcock, FDA's CDER Director, 2006
  - Improved utilization of adaptive and Bayesian methods could help resolve low success rate of and expense of phase 3 clinical trials
- Margaret Hamburg, FDA Commissioner 2010
  - "The final guidance on the use of Bayesian statistics is consistent with the FDA's commitment to streamline clinical trials, when possible, in order to get safe and effective products to market faster."
- CDRH produced guidelines for Bayesian statistics Feb 5, 2010
  - "Agency says Bayesian statistical methods could trim costs, boost efficiency" from press release
  - "They beauty is you do not end up doing a trial that is too big or too small; you end up doing a trial that is just right." Greg Campbell
- CDER/CBER produced draft guidance for adaptive designs Feb 2010
  - Generally supportive of well-characterized adaptation by design
  - Appropriately cautious

FDA Guidance Documents

Guidance for Industry and FDA Staff

Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials

Document issued on: February 5, 2010

The draft of this document was issued on 5/23/2006

For questions regarding this document, contact Dr. Greg Campbell (CDRH) at 301-796-5750 or greg.campbell@fda.hhs.gov or the Office of Communication, Outreach and Development, (CBER) at 1-800-835-4709 or 301-827-1800.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Devices and Radiological Health

Division of Biostatistics
Office of Surveillance and Biometrics

Center for Biologics Evaluation and Research

Guidance for Industry

Adaptive Design Clinical Trials for Drugs and Biologics

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document contact Robert O'Neill or Sue-Jane Wang at 301-796-1700, Marc Walton at 301-796-2600 (CDER), or the Office of Communication, Outreach and Development (CBER) at 800-835-4709 or 301-827-1800.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

February 2010
Clinical/Medical

Asking for Adaptive/Bayesian!

Guidance for Industry and FDA Staff

Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials

Document issued on: February 5, 2010

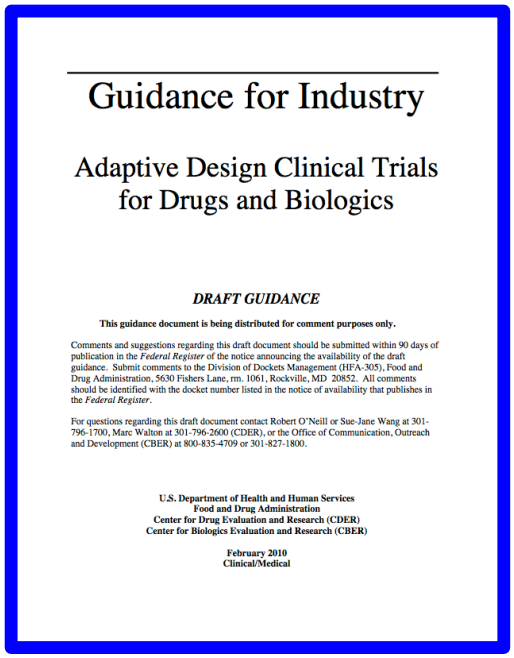
The draft of this document was issued on 5/23/2006

For questions regarding this document, contact Dr. Greg Campbell (CDRH) at 301-796-5750 or greg.campbell@fda.hhs.gov or the Office of Communication, Outreach and Development, (CBER) at 1-800-835-4709 or 301-827-1800.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Devices and Radiological Health

Division of Biostatistics
Office of Surveillance and Biometrics

Center for Biologics Evaluation and Research



Guidance for Industry

Adaptive Design Clinical Trials for Drugs and Biologics

DRAFT GUIDANCE

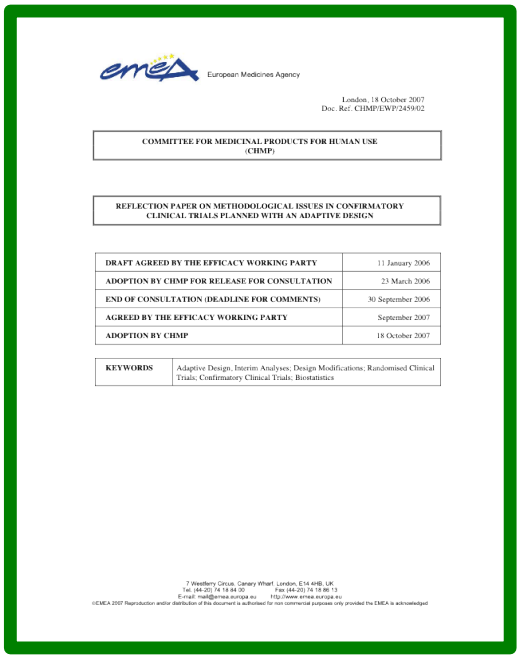
This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document contact Robert O'Neill or Sue-Jane Wang at 301-796-1700, Marc Walton at 301-796-2600 (CDER), or the Office of Communication, Outreach and Development (CBER) at 800-835-4709 or 301-827-1800.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

February 2010
Clinical/Medical



European Medicines Agency

London, 18 October 2007
Doc. Ref. CHMP/EWP/2459/02

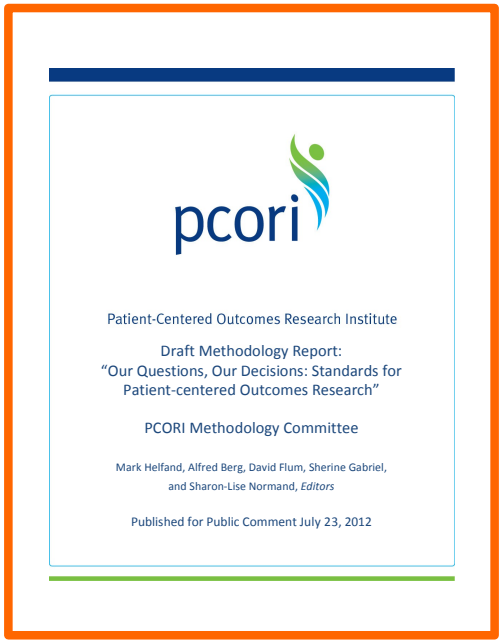
COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE (CHMP)

REFLECTION PAPER ON METHODOLOGICAL ISSUES IN CONFIRMATORY CLINICAL TRIALS PLANNED WITH AN ADAPTIVE DESIGN

DRAFT AGREED BY THE EFFICACY WORKING PARTY	11 January 2006
ADOPTION BY CHMP FOR RELEASE FOR CONSULTATION	23 March 2006
END OF CONSULTATION (DEADLINE FOR COMMENTS)	30 September 2006
AGREED BY THE EFFICACY WORKING PARTY	September 2007
ADOPTION BY CHMP	18 October 2007

KEYWORDS	Adaptive Design, Interim Analyses; Design Modifications; Randomised Clinical Trials; Confirmatory Clinical Trials; Biostatistics

7 Westferry Circus, Canary Wharf, London, E14 4HB, UK
Tel. (44-20) 74 18 84 00 Fax (44-20) 74 18 86 13
E-mail: mail@emea.europa.eu http://www.emea.europa.eu
©EMEA 2007 Reproduction and/or distribution of this document is authorised for non commercial purposes only provided the EMEA is acknowledged



pcori

Patient-Centered Outcomes Research Institute

Draft Methodology Report:
"Our Questions, Our Decisions: Standards for Patient-centered Outcomes Research"

PCORI Methodology Committee

Mark Helfand, Alfred Berg, David Flum, Sherine Gabriel, and Sharon-Lise Normand, *Editors*

Published for Public Comment July 23, 2012

Draft Guidance for Industry and Food and Drug Administration Staff

Adaptive Designs for Medical Device Clinical Studies

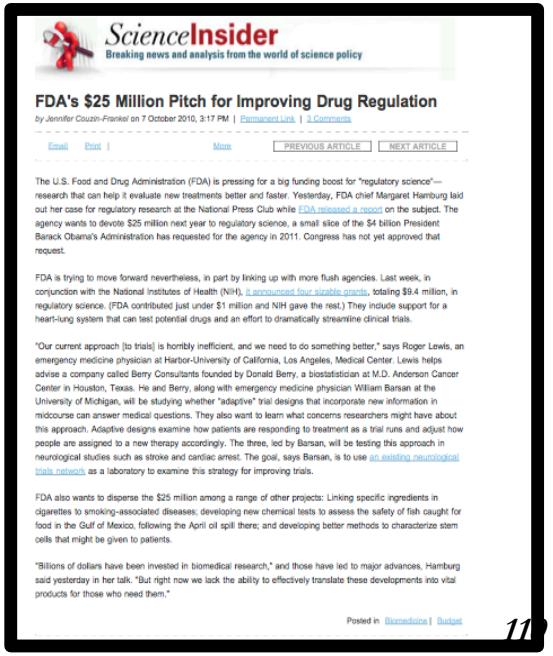
DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.
Document issued on: May 18, 2015

You should submit comments and suggestions regarding this draft document within 90 days of publication in the Federal Register of the notice announcing the availability of the draft guidance. Submit written comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. Submit electronic comments to http://www.regulations.gov. Identify all comments with the docket number listed in the notice of availability that publishes in the Federal Register.

For questions regarding this document that relate to devices regulated by CDRH, contact Dr. Greg Campbell (CDRH) at 301-796-5750 or by e-mail at greg.campbell@fda.hhs.gov. For questions regarding this document that relate to devices regulated by CBER, contact the Office of Communication, Outreach and Development (CBER) at 1-800-835-4709 or240-402-7800.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Devices and Radiological Health
Center for Biologics Evaluation and Research



ScienceInsider
Breaking news and analysis from the world of science policy

FDA's $25 Million Pitch for Improving Drug Regulation

by Jennifer Couzin-Frankel on 7 October 2010, 3:17 PM | Permanent Link | 3 Comments

Email Print | More PREVIOUS ARTICLE NEXT ARTICLE

The U.S. Food and Drug Administration (FDA) is pressing for a big funding boost for "regulatory science"—research that can help it evaluate new treatments better and faster. Yesterday, FDA chief Margaret Hamburg laid out her case for regulatory research at the National Press Club while FDA released a report on the subject. The agency wants to devote $25 million next year to regulatory science, a small slice of the $4 billion President Barack Obama's Administration has requested for the agency in 2011. Congress has not yet approved that request.

FDA is trying to move forward nevertheless, in part by linking up with more flush agencies. Last week, in conjunction with the National Institutes of Health (NIH), it announced four sizable grants, totaling $9.4 million, in regulatory science. (FDA contributed just under $1 million and NIH gave the rest.) They include support for a heart-lung system that can test potential drugs and an effort to dramatically streamline clinical trials.

"Our current approach [to trials] is horribly inefficient, and we need to do something better," says Roger Lewis, an emergency medicine physician at Harbor-University of California, Los Angeles, Medical Center. Lewis helps advise a company called Berry Consultants founded by Donald Berry, a biostatistician at M.D. Anderson Cancer Center in Houston, Texas. He and Berry, along with emergency medicine physician William Barsan at the University of Michigan, will be studying whether "adaptive" trial designs that incorporate new information in midcourse can answer medical questions. They also want to learn what concerns researchers might have about this approach. Adaptive designs examine how patients are responding to treatment as a trial runs and adjust how people are assigned to a new therapy accordingly. The three, led by Barsan, will be testing this approach in neurological studies such as stroke and cardiac arrest. The goal, says Barsan, is to use an existing neurological trials network as a laboratory to examine this strategy for improving trials.

FDA also wants to disperse the $25 million among a range of other projects: Linking specific ingredients in cigarettes to smoking-associated diseases; developing new chemical tests to assess the safety of fish caught for food in the Gulf of Mexico, following the April oil spill there; and developing better methods to characterize stem cells that might be given to patients.

"Billions of dollars have been invested in biomedical research," and those have led to major advances, Hamburg said yesterday in her talk. "But right now we lack the ability to effectively translate these developments into vital products for those who need them."

Posted in Biomedicine | Budget

Online Tools & Resources

- MD Anderson
  - http://biostatistics.mdanderson.org/SoftwareDownload/
  - Lots of good utilities, including "Adaptive Randomization" to help with response adaptive trials
  - Allows 10 arms; minimum number of patients before adapting randomization scheme; maximum number of patients or length of trial
  - Free
- Commercial resources increasingly available
  - Lack of affordable academic options

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