
An Overview of Bayesian Adaptive Clinical Trial Design

Roger J. Lewis, MD, PhD
Department of Emergency Medicine
Harbor-UCLA Medical Center
David Geffen School of Medicine at UCLA
Los Angeles Biomedical Research Institute
Berry Consultants, LLC

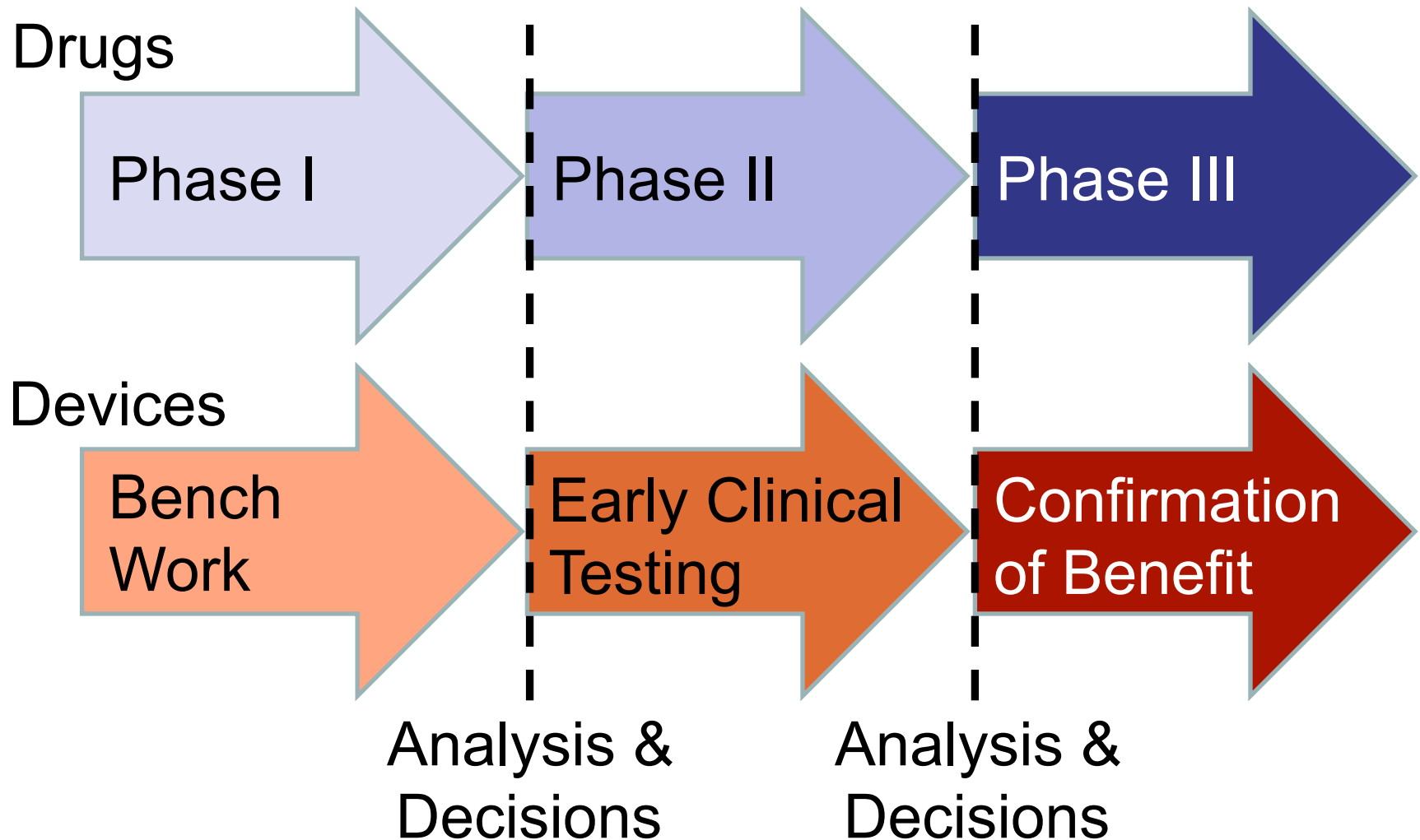
Financial Disclosures

- Berry Consultants, LLC
 - Multiple clients
- Support from
 - National Institutes of Health/NINDS
 - Food and Drug Administration
- FDA/CDRH Entrepreneurs-in-Residence (EIR) Program
- Applied Proteomics, Inc.
- AspenBio Pharma
- Octapharma USA
- Octapharma AG

Motivation for Adaptive Trials

- When designing a trial there is substantial uncertainty regarding how best to treat subjects in the experimental arm (e.g., uncertainty in optimal dose, best duration, target population)
- This creates uncertainty in the optimal design
- Traditionally, all key trial parameters are defined and held constant during execution
- This leads to increased risk of negative or failed trials, even if a treatment is inherently effective

Traditional Restriction of Flexibility in Drug and Device Development



Motivation for Adaptive Trials

- Once patients are enrolled and their outcomes known, information accumulates that reduces uncertainty regarding optimal treatment approaches
- Adaptive clinical trials are designed to take advantage of this accumulating information, by allowing modification to key trial parameters in response to accumulating information and according to predefined rules



JAMA 2006;296:1955-1957.

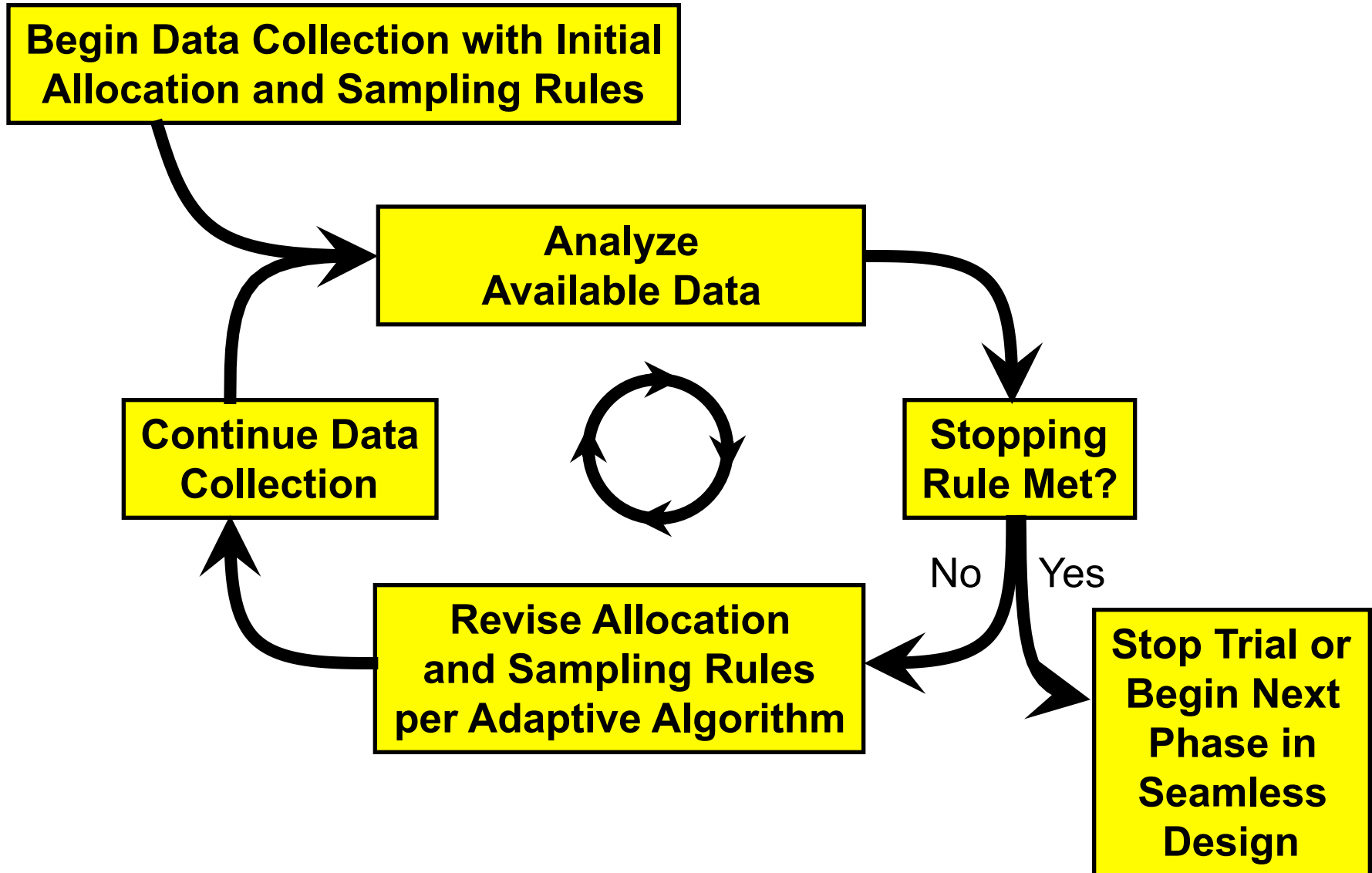
Characteristics of Adaptive Trials

- Clarity of goals
 - E.g., proof of concept vs. identification of dose to carry forward vs. confirmation of benefit
 - A statistically significant p value is not a goal
- Frequent “looks” at the data and data-driven modification of the trial
- Adaptive “by design”
- Extensive use of simulation to adjust characteristics of trial design

Adaptation: Definition

- Making planned, well-defined changes in key clinical trial design parameters, during trial execution based on data from that trial, to achieve goals of validity, scientific efficiency, and safety
 - Planned: Possible adaptations defined *a priori*
 - Well-defined: Criteria for adapting defined
 - Key parameters: *Not* minor inclusion or exclusion criteria, routine amendments, etc.
 - Validity: Reliable statistical inference

The Adaptive Process



Why Do Adaptive Clinical Trials?

- To avoid getting the wrong answer!
 - Drawing an incorrect qualitative conclusion
- To avoid taking too long to draw the right conclusion
 - Time, human subjects, and resources

Avoiding Anticipated Regret

- A substantial fraction of all confirmatory trials fail despite promising “learn phase” results
- Investigators can anticipate the design decisions they would wish to “take over” after the trial fails
- Areas of “anticipated regret” are promising targets for adaptations

The Maginot Line

(adapted from Wikipedia)

- The Maginot Line was a line of concrete fortifications which France constructed along its borders with Germany after World War I
- Military experts extolled the Maginot Line as genius
- In World War II the German army bypassed the Maginot Line, conquering France in days
- “Generals always fight the last war, especially if they have won it”

Avoid Building Statistical Maginot Lines

- Tempting to design protections from minor threats (e.g., covariate imbalance) with familiar solutions rather than address major threats (e.g., limited power, changing endpoints from phase II to phase III)
- Traditional and “well accepted” approaches are often not very good (e.g., traditional dose-finding approaches)

One Truly Understands Only What One Can Simulate

- The relative importance and likelihood of threats to trial validity depends on the ***specifics*** of the situation
- Simulation is the best way to quantify the different threats to validity and inform rational trial design

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

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, etc.
- Logistically practical
- Able to secure buy-in of stakeholders

Some (Bayesian) Adaptive Strategies

- Frequent interim analyses
- Explicit longitudinal modeling of the relationship between proximate endpoints and the primary endpoint of the trial
- Response-adaptive randomization to efficiently address one or more trial goals
- Explicit decision rules based on predictive probabilities at each interim analysis
- Dose-response modeling
- Enrichment designs
- Extensive simulations of trial performance

Frequent Interim Analyses

- **Frequent interim analyses** based on Markov-chain Monte Carlo (MCMC) estimates of Bayesian posterior probability distributions, with multiple imputation and estimation of unknown trial parameters and patient outcomes.
- Typically quantify
 - Evidence of treatment efficacy
 - Trial futility/predictive probability of success
 - Safety and rates of adverse events

Longitudinal Modeling

- Explicit **longitudinal modeling** of the relationship between proximate endpoints and the primary (generally longer term) endpoint of the trial to better inform interim decision making, based on the data accumulating within the trial and without assuming any particular relationship at the beginning of the trial.
- Used to learn about, and utilize, the relationship between proximate and final endpoints
- Frequently misunderstood as “making assumptions” or using “biomarkers”

Response-adaptive Randomization

- **Response-adaptive randomization** to improve important trial characteristics
- May be used to address one or more of:
 - To improve subject outcomes by preferentially randomizing patients to the better performing arm
 - To improve the efficiency of estimation by preferentially assigning patients to doses in a manner that increases statistical efficiency
 - To improve the efficiency in addressing multiple hypotheses by randomizing patients in a way that emphasizes sequential goals
 - Includes arm dropping

Decision Rules/Predictive Probabilities

- Explicit **decision rules based on predictive probabilities** at each interim analysis to define when to stop for futility, early success, etc.
- Examples
 - May define success or futility based on the predictive probability of success if trial is stopped and all patients followed to completion
 - May define success or futility based on the predictive probability of success of a **subsequent** phase III trial
 - May combine probabilities logically: probability that the active agent is **both** superior to a control arm and non-inferior to an active comparator
 - Design “transitions”: e.g., phase II to phase III

Dose-response Modeling

- **Dose-response modeling**, when applicable, so that information from all patients informs the estimate of the treatment effect at all doses—this improves the reliability of interim decision making and improves accuracy in the updating of interim randomization proportions.
- **Examples**
 - Logistic dose-response model: assumes monotonicity
 - Normal dynamic linear model (NDLM): borrows information from adjacent doses but doesn't assume a particular shape of the relationship

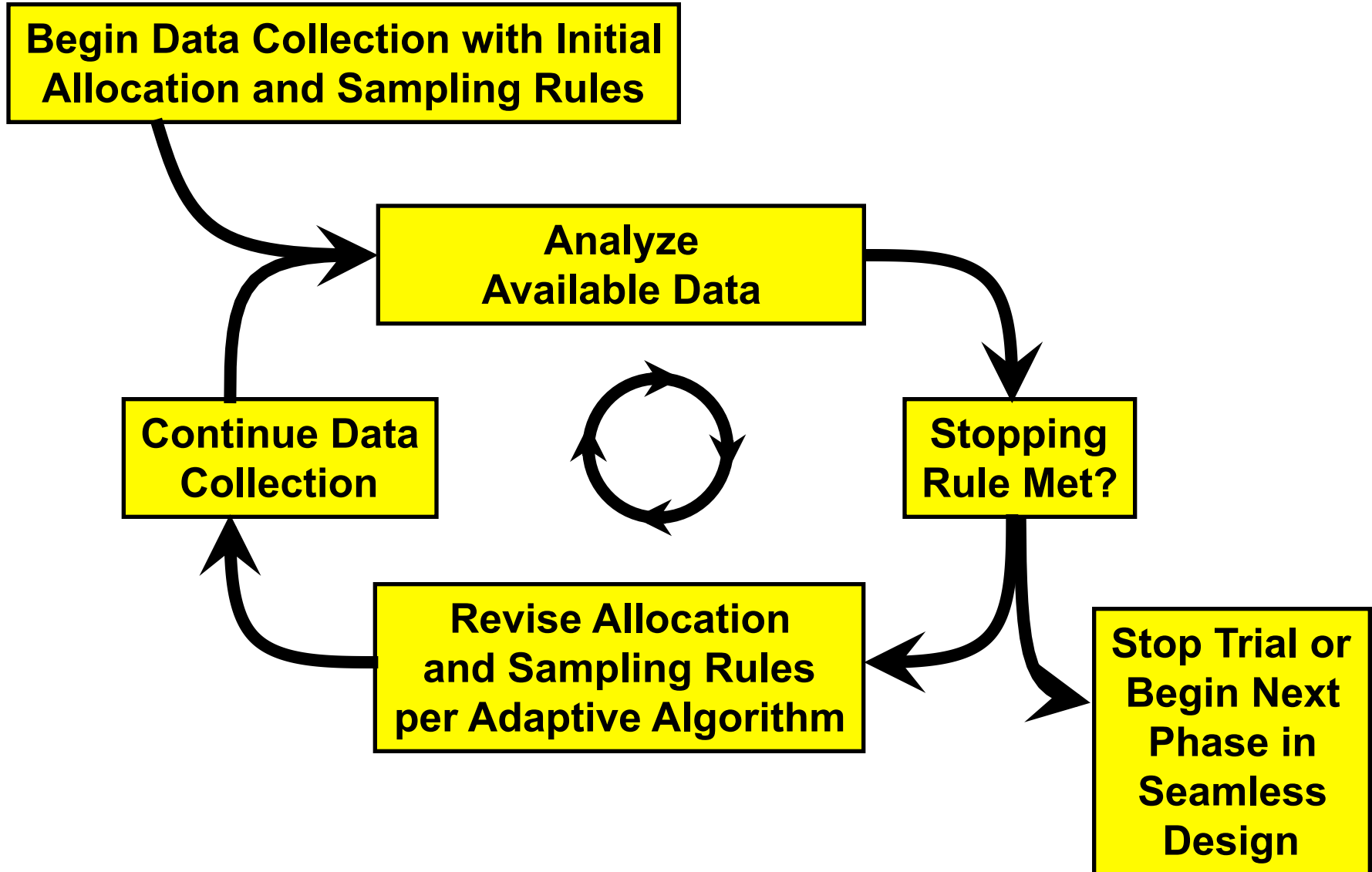
Enrichment Designs

- In an **enrichment design**, the inclusion and exclusion criteria may be altered over time to focus on a population most likely to benefit from the experimental population.
- Final analysis population may include all subjects or only those meeting final enrollment criteria.
- Examples
 - Eliminating enrollment of patients with a long time between symptom onset and seeking care
 - Eliminating enrollment of patients with higher risk of complications

Extensive Simulations

- **Extensive simulations of trial performance** to ensure that the type I error rate, power and accuracy in estimation of treatment effect(s), the rates of adverse events, or dose finding are well defined and acceptable, across a very wide range of possible true treatment effect sizes, dose-response relationships, and population characteristics.
- Often end up exploring and understanding the performance characteristics across a range of null hypotheses much broader than with traditional approaches

The Adaptive Process





HIRAM S. DUDSON
1930 - 1993

Member,
Placebo Group



D. Reilly

L-Carnitine and Sepsis

- Clinical setting
 - Adult patients with severe sepsis or shock
 - Phase II, dose-finding trial of L-carnitine to improve end organ function and survival
- Goals
 - Identify most promising dose
 - Determine if L-carnitine should be evaluated in a confirmatory, phase III trial
 - Enroll more patients to doses most likely to be beneficial, based on accumulating information

L-Carnitine and Sepsis

- More Background
 - L-carnitine is believed to work through reducing multi-organ system failure
 - Multi-organ system failure quantified by SOFA score
 - Baseline SOFA is key predictor of mortality
 - Reduction in SOFA over 48 hours is desired proximate treatment effect
 - Reduction in 28-day mortality would be registration endpoint

Adaptive Trial Structure

- Outcome measures
 - Proximate: Δ SOFA score
 - Definitive: Survival to 28 days
- Structure of trial
 - 4 arms (0 g, 6 g, 12 g, and 18 g) with dose-response model
 - Maximum sample size of 250 subjects
 - Interim analyses at 40 subjects, then every 12
 - Subjects randomized according to probability that the dose results in the best (negative) Δ SOFA
 - May be stopped early for futility or success, based on probability that best dose improves SOFA and would be successful in phase III

Operating Characteristics of Proposed Trial Design: Results of Monte Carlo Simulations

	No Effect (Null)		Mild Effect		Strong Effect	
<u>Assumed Treatment Effects for Simulations</u>						
	Δ SOFA	Mortality	Δ SOFA A	Mortality	Δ SOFA	Mortality
Outcome: Control	0	40%	0	40%	0	40%
Outcome: 6 g	0	40%	0	40%	-1	34%
Outcome: 12 g	0	40%	-1	34%	-2	28%
Outcome: 18 g	0	40%	-2	28%	-4	19%

Operating Characteristics of Proposed Trial Design: Results of Monte Carlo Simulations						
	No Effect (Null)		Mild Effect		Strong Effect	
<u>Assumed Treatment Effects for Simulations</u>						
	Δ SOFA	Mortality	Δ SOFA A	Mortality	Δ SOFA	Mortality
Outcome: Control	0	40%	0	40%	0	40%
Outcome: 6 g	0	40%	0	40%	-1	34%
Outcome: 12 g	0	40%	-1	34%	-2	28%
Outcome: 18 g	0	40%	-2	28%	-4	19%
<u>Trial Performance</u>						
Probability of Positive Trial	0.043 (type I error)		0.911 (power)		0.999	
Probability of Stopping Early	For futility: 0.431 For success: 0.023		For futility: 0.001 For success: 0.679		For futility: 0.000 For success: 0.981	
Average Req'd Sample Size	198.0		172.4		119.5	
Probability of Selecting 18 g	0.35		0.99		1.00	

Operating Characteristics of Proposed Trial Design: Results of Monte Carlo Simulations

	No Effect (Null)	Mild Effect	Strong Effect			
<u>Assumed Treatment Effects for Simulations</u>						
	Δ SOFA	Mortality	Δ SOFA A	Mortality	Δ SOFA	Mortality
Outcome: Control	0	40%	0	40%	0	40%
Outcome: 6 g	0	40%	0	40%	-1	34%
Outcome: 12 g	0	40%	-1	34%	-2	28%
Outcome: 18 g	0	40%	-2	28%	-4	19%
<u>Trial Performance</u>						
Probability of Positive Trial	0.043 (type I error)		0.911 (power)		0.999	
Probability of Stopping Early	For futility: 0.431 For success: 0.023		For futility: 0.001 For success: 0.679		For futility: 0.000 For success: 0.981	
Average Req'd Sample Size	198.0		172.4		119.5	
Probability of Selecting 18 g	0.35		0.99		1.00	
<u>Average Allocation of Subjects Between Treatment Arms – n per arm (%)</u>						
Control	62.7 (32%)		54.1 (31%)		36.5 (31%)	
6 g	47.0 (24%)		13.8 (8%)		10.5 (9%)	
12 g	38.7 (20%)		21.5 (12%)		12.5 (10%)	
18 g	49.6 (25%)		83.0 (48%)		60.0 (50%)	

Trial Status

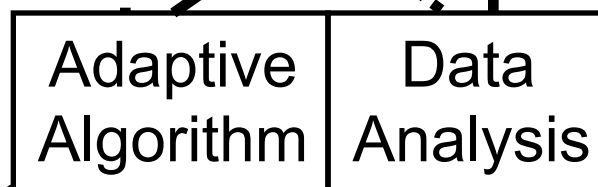
- Recently funded by US National Institutes of Health/National Institute of General Medical Sciences
- Led by Alan E. Jones, MD at the University of Mississippi, Department of Emergency Medicine
- Currently beginning trial implementation

Components of an Adaptive Trial

Management



Adaptive Machinery



Logistics



Clinical



Conclusions

- Not all trials need (or should have) adaptive designs
- When used appropriately, adaptive designs may:
 - Improve efficiency and reduce cost
 - Maximize the information obtained
 - Minimize risk to subjects and sponsor
- Design decisions should be based on objective performance rather than habit
- An adaptive design will not save a poorly planned trial or ineffective treatment

