Interim Analyses: Promises and Pitfalls

Michael P. McDermott, Ph.D.
University of Rochester
October 26, 2012

Adaptive Designs

- Methodological research on adaptive designs and interest in their application has grown tremendously over the last 20 years
  - Industry, academia, and government
- Why?

Motivation for Adaptive Designs

- Need for greater efficiency in drug development
  - Low success rates (< 50%) in confirmatory (Phase III) studies
  - Traditional paradigm supports solid inferential procedures, but efficiency is relatively low
    - Dose finding problem: How many dosages (and which ones) can/should be studied?
  - Accumulating data can be used to modify the course of the trial
    - Often results in a smaller amount of better information

Motivation for Adaptive Designs

- Efficiency
  - Smaller studies
  - Better use of limited resources
  - Fewer trials to accomplish multiple objectives
  - Savings in time/cost
- Ethical issues
  - Strong accumulating evidence of efficacy or of safety concerns

What is an Adaptive Design?

- Pharmaceutical Research and Manufacturers of America (PhRMA) Working Group (2006):
  - "By *adaptive design*, we refer to a clinical study design that uses accumulating data to decide how to modify aspects of the study as it continues, without undermining the validity and integrity of the trial."
  - "In such trials, changes are made 'by design,' and not on an ad hoc basis; therefore, adaptation is a *design feature* aimed to enhance the trial, not a remedy for inadequate planning."

Gallo P, et al. J Biopharm Statist 2006; 16:275-283

What is an Adaptive Design?

- FDA Draft Guidance (2010):
  - "For the purposes of this guidance, an *adaptive design clinical study* is defined as a study that includes a prospectively planned opportunity for modification of one or more specified aspects of the study design and hypotheses based on analysis of data (usually interim data) from subjects in the study."
  - "The term *prospective* here means that the adaptation was planned (and details specified) before data were examined in an unblinded manner by any personnel involved in planning the revision."

What is an Adaptive Design?

- Validity
  - Correct statistical inference
    - Control of Type I and Type II errors
    - Minimization of bias
  - Consistency between stages of the trial
  - Low operational bias
- Integrity
  - Results are convincing to a broader scientific community
  - Pre-planned adaptations
  - Maintenance of the blind to interim analysis results

Dragalin V. Drug Inf J 2006; 40:425-435

NOT Adaptive Design . . .

- Unplanned adaptations
  - Difficult to remedy the resulting bias
  - May be valid if the adaptations are made without knowledge of treatment group assignment
- Adaptations made on the basis of information external to the study

Some Possible Adaptations

- Dose finding
  - Continual reassessment method
- Dropping/adding treatment arms
  - “Drop the loser” designs
- Randomization
  - Response-adaptive or covariate-adaptive
- Sample size re-estimation
- Early stopping for safety, futility, or efficacy
  - Group sequential monitoring

Some Possible Adaptations

- Hypotheses/objectives
  - Non-inferiority → superiority
- Primary outcome variable
  - Variable, timing, composite components
- Eligibility criteria
  - Enrichment; subgroup for analysis
- Statistical analysis plan
- Biomarker-related adaptations
- "Seamless" Phase II/III designs

Main Concerns with Adaptive Designs

- Control of Type I error probability
- Bias (estimates of treatment effects)
- Interpretation of results
- Logistical issues
- Procedural issues

Group Sequential Designs

- Problem of repeated significance testing
  - Effect on Type I error probability
  - Effect on Type II error probability (or power)
  - Bias in estimation of treatment effect
- Construction of stopping boundaries for efficacy, futility, or harm
  - α-spending functions
  - β-spending functions
    - Can rule out a benefit the size of the hypothesized treatment effect

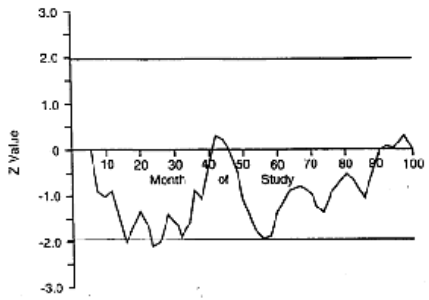


Fig. 15-1 Interim survival analyses comparing mortality in clofibrate- and placebo-treated participants in the Coronary Drug Project. A positive Z value favors placebo.[31] Reprinted by permission of Elsevier Science.

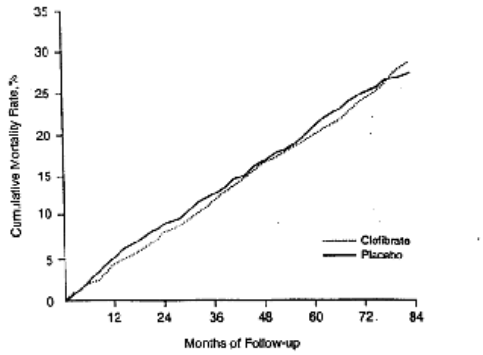
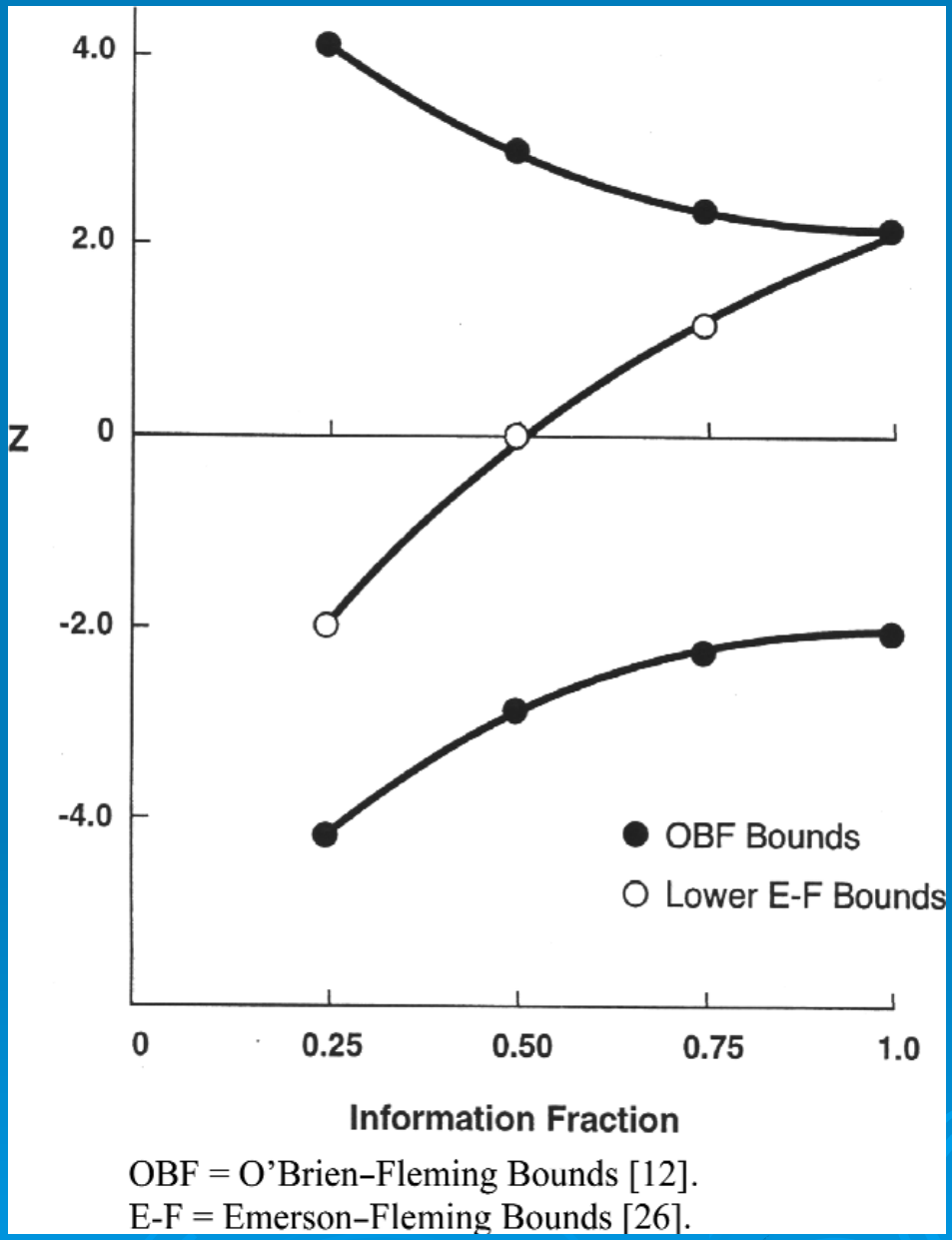


Fig. 15-2 Cumulative mortality curves comparing clofibrate- and placebo-treated participants in the Coronary Drug Project.[31] Reprinted by permission of Elsevier Science.

Friedman LM, Furberg CD, DeMets DL (1998), p. 252



OBF = O'Brien–Fleming Bounds [12].
E-F = Emerson–Fleming Bounds [26].

DeMets DL. Clin Trials 2006; 3:522-529

Group Sequential Designs

- Stopping early for efficacy
  - Patients benefit sooner from the new treatment
  - Less information on secondary outcomes (including safety) and subgroups
- Stopping early for futility
  - Can reallocate resources to other projects
  - May not be able to determine if the treatment is merely ineffective or actually harmful

Sample Size Re-estimation

- Sample size calculations depend on:
  - Significance level and power
  - Magnitude of treatment effect
    - Chosen to be that of minimal clinical importance
  - "Nuisance parameters"
    - SD of outcome variable
    - Event rate in control group
- Sample size depends critically on the specified values

Sample Size Re-estimation

- Internal pilot study
  - After outcome data have accumulated from a specified number of subjects, estimate the "nuisance parameter" and recalculate the required sample size
    - Blinded vs. unblinded methods
    - Risk in decreasing the required sample size
  - Type I error probability is inflated if the final analysis is not adjusted for this interim examination of the data
    - Inflation is negligible unless the size of the internal pilot study is very small
    - Adjustments can be made in the case of small internal pilot studies

Sample Size Re-estimation

- Methods that adjust the sample size based on the estimated treatment effect are controversial and generally discouraged
  - Early estimates of treatment effects may be imprecise
  - Small treatment effects may not be clinically relevant
  - Research has demonstrated that existing group sequential designs are generally more efficient
    - Jennison and Turnbull (2003, 2006)
    - Tsiatis and Mehta (2003)
  - Better to estimate the sample size at the design stage using the treatment effect that is of minimal clinical relevance

Sample Size Re-estimation

- Example: Two group comparison
- Suppose that a treatment effect between $\Delta = 2$ and $\Delta = 4$ is thought to be plausible and clinically relevant
- Fixed sample size design:
  - N = 560 required for $\Delta = 2$
  - N = 140 required for $\Delta = 4$

Sample Size Re-estimation

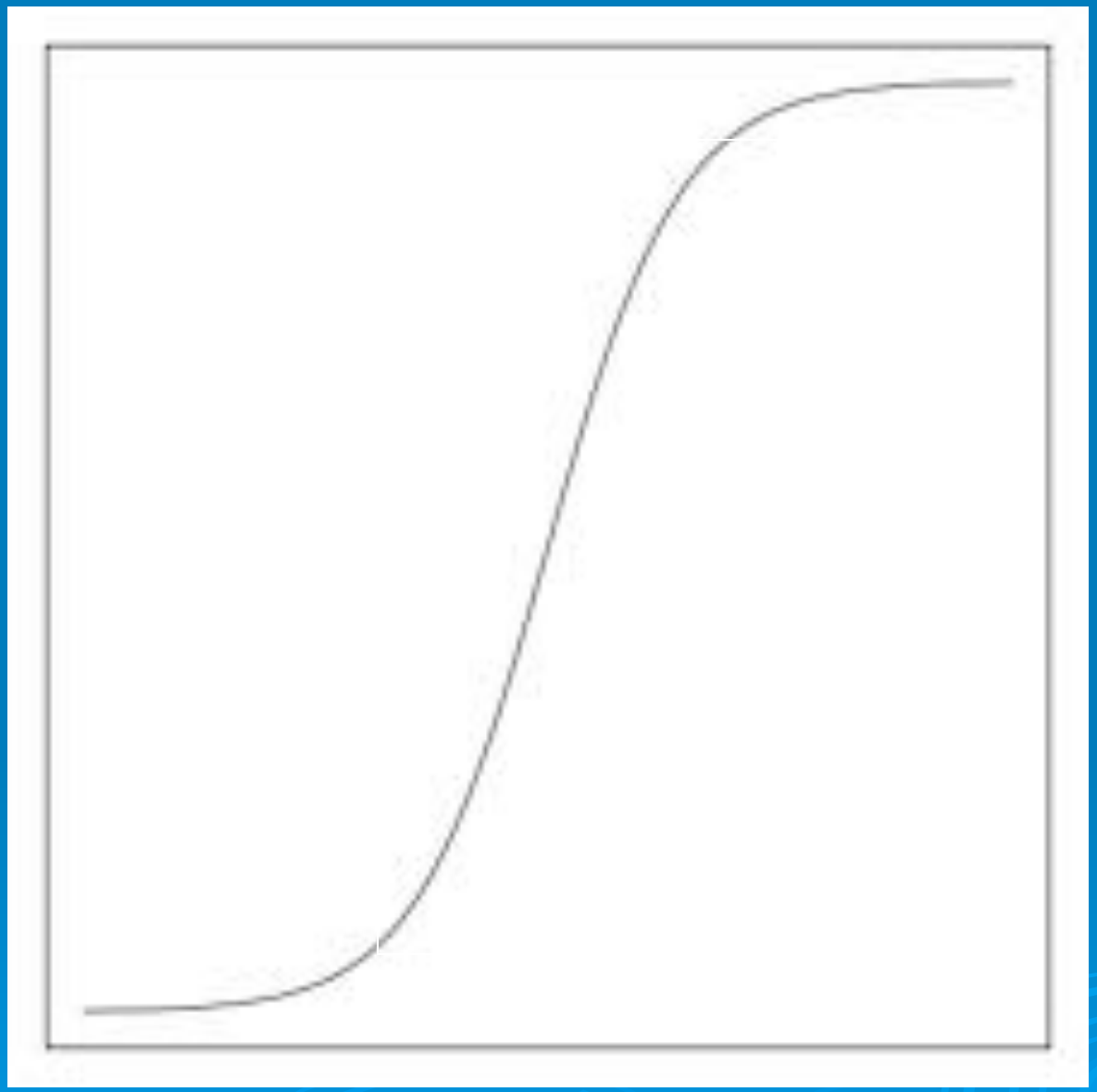
- Possible Designs
  - Group sequential design
    - Design the trial to detect the smaller value $\Delta = 2$
    - If $\Delta > 2$, a stopping boundary will likely be crossed during interim monitoring and the trial will terminate early
  - Sample size re-estimation design
    - Design the trial to detect the larger value $\Delta = 4$
    - If interim results indicate that $\Delta < 4$, increase the sample size to detect the (possibly) smaller effect

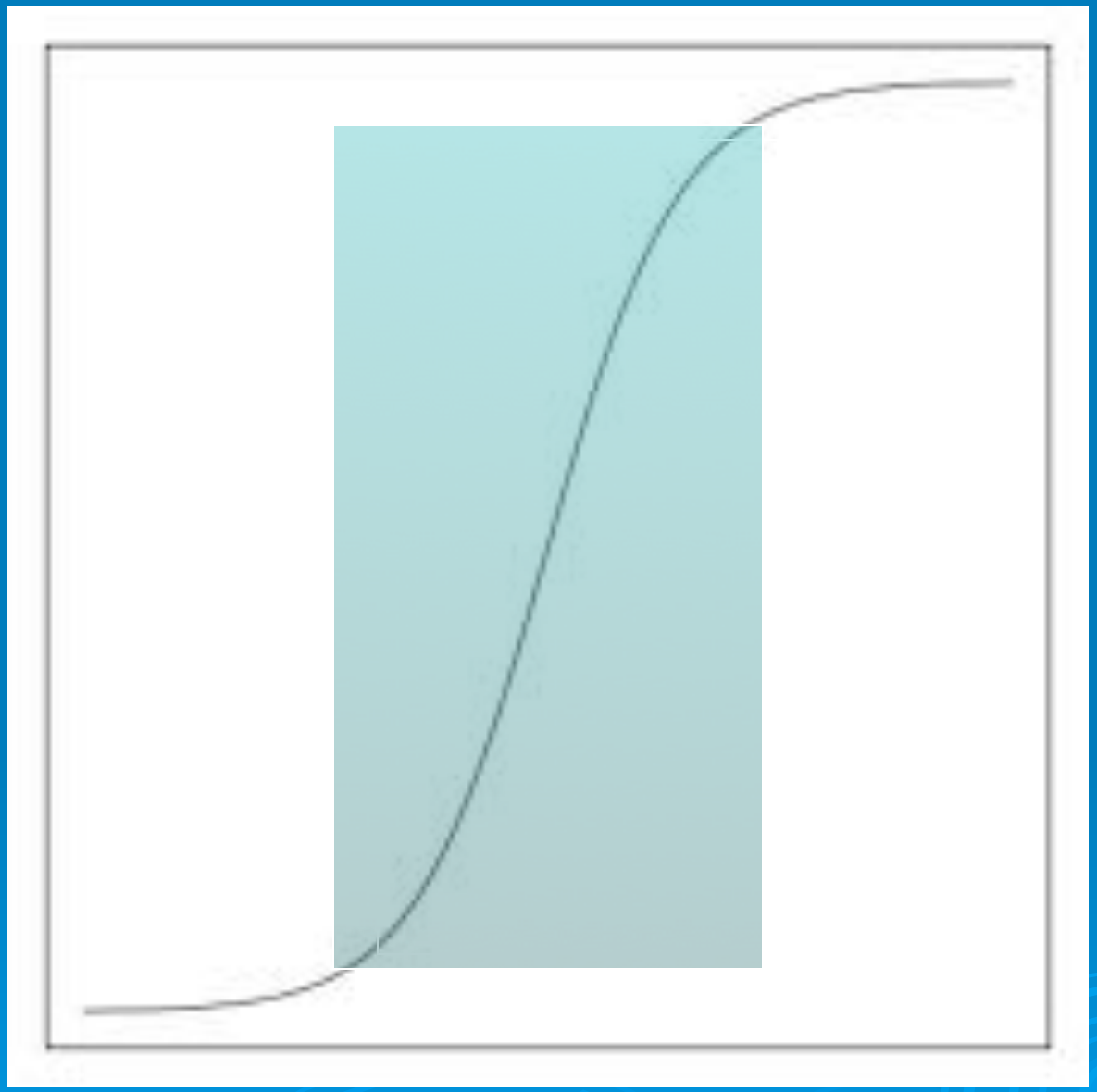
Adaptive Dose Finding

- Traditional approach in Phase II
  - Randomization to a relatively small number of fixed dosages (3-4) and placebo
  - Disadvantages
    - Large “distance” between adjacent dosages
      - Optimal dosage may not be studied
    - Some of the studied dosages may not be useful
      - This may become apparent relatively quickly
    - Accumulating evidence may suggest early stopping for futility or identification of a sufficient dosage to study further

Adaptive Dose Finding

- Example: D-optimal response-adaptive design
  - Choose a rich family of models for the dose-response relationship (e.g., sigmoid E*max* model)
  - Randomize the first cohort of patients to several active dosages and placebo (equal allocation)
  - Estimate the model parameters and, hence, the dose-response curve
  - Determine (from this curve) the allocation scheme that will maximize the amount of information about the overall dose-response relationship (model parameters)

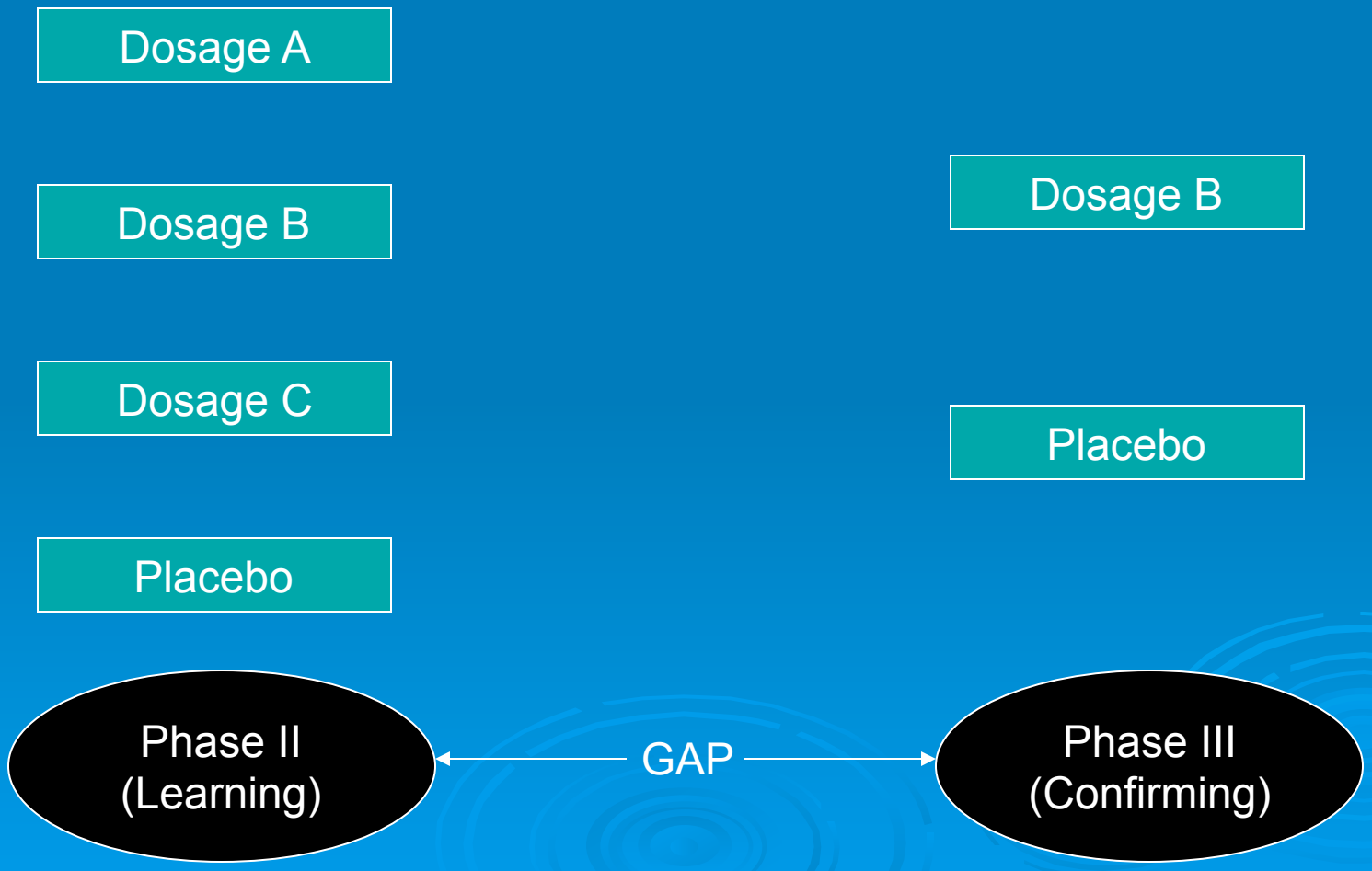

Response
Dosage


Response
Dosage

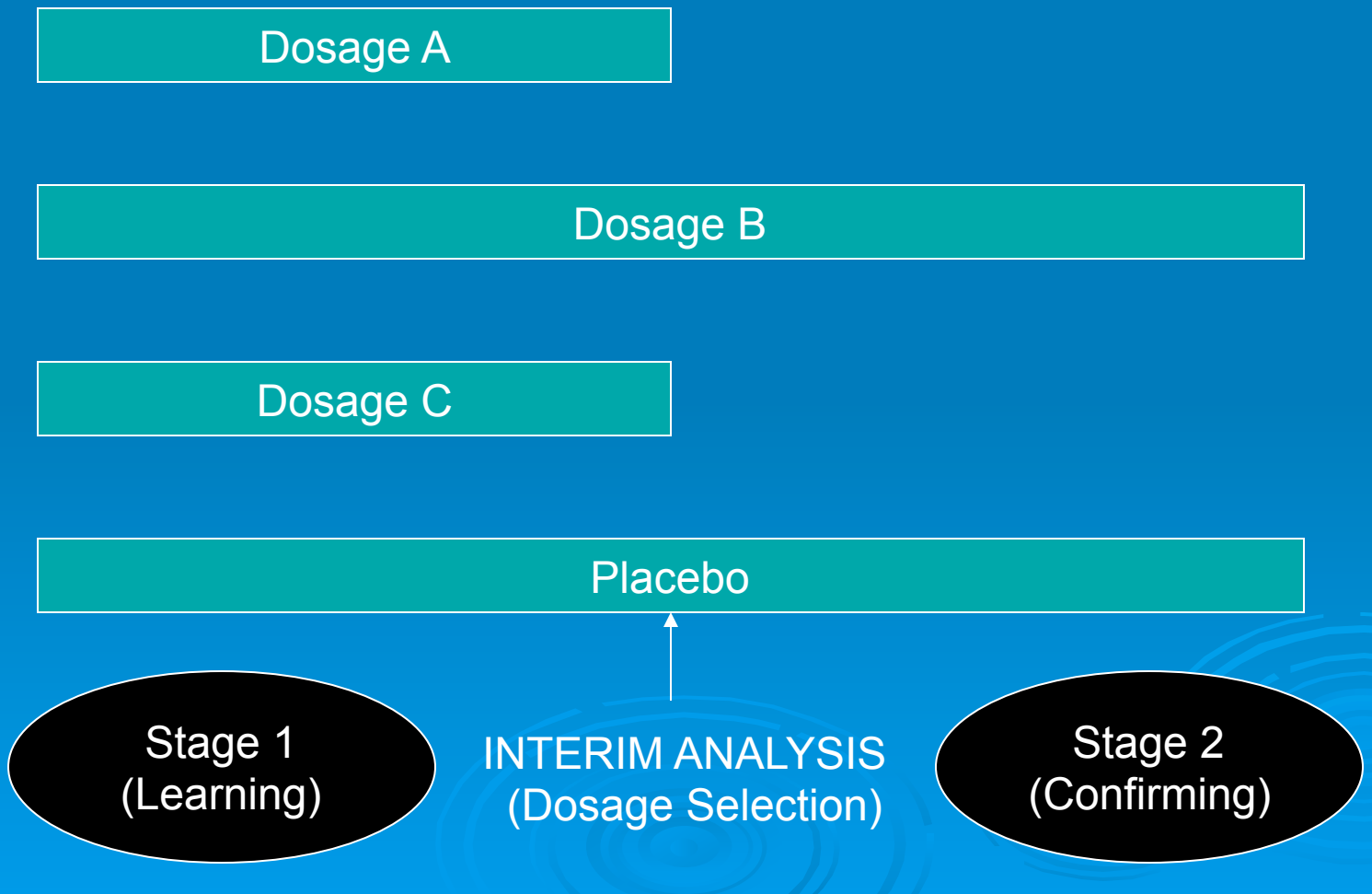
Adaptive Dose Finding

- Example: D-optimal response-adaptive design
  - Randomly assign next cohort of patients to the active dosages and placebo according to the derived optimal allocation scheme
  - Repeat the process until total sample size is reached or a pre-specified futility rule is met
  - Recent methods include a penalty for allocation of patients to ineffective or toxic dosages
    - Balance between individual and collective ethics
    - Padmanabhan et al. (2010)

Seamless Phase II/III Designs



Seamless Phase II/III Designs



Logistical Issues

- Availability of outcomes in the short-term (relative to the overall duration of the trial) to allow efficient adaptation
- Information technology infrastructure capable of rapid data acquisition, quality control, analysis, and reporting
- Integration of data capture, drug supply management, and interactive communication system between sites and the coordination center

Logistical Issues

- Ability to perform extensive simulation studies in the planning phase in order to thoroughly evaluate the operating characteristics of the statistical procedures to be used
  - Type I error probability; power
  - Distributions of sample size, trial duration, amount of study drug required
  - Costs
  - Need to examine many scenarios (accrual rate is an important factor)
- Funding for this activity in academic settings?

Logistical Issues

- Ability to quickly perform complex and sometimes computationally intensive analyses
- Growing availability of commercial software to plan and implement adaptive designs
- Budgetary issues
  - Government vs. industry-sponsored trials
- Grant review
  - Space limitations
  - Expertise of reviewers

Procedural Issues

- Who should review the interim study results and implement the adaptations?
  - DSMB? Separate "Adaptation Committee"?
    - Need relevant expertise
    - Need increased level of commitment from the usual standard
  - Involvement of independent study statistician
    - Intimately familiar with the trial
    - Expertise to report/explain interim results
    - Importance of independence (same applies to data analysts)

Procedural Issues

- Who should review the interim study results and implement the adaptations?
  - Importance of pre-specified decision algorithms
    - Minimize need for subjective judgment
    - Trust issues / loss of control
  - Representation of sponsor?
    - May involve business/funding decisions
    - Need to maintain trial integrity

Procedural Issues

- Operational bias
  - Visibility of adaptations may be used to infer trial results, affecting investigator or subject behavior during the trial
    - Participation, adherence, objectivity of ratings, etc.
    - May cause heterogeneity of results before vs. after adaptation
  - Not a statistical form of bias
    - Cannot adjust for this

Procedural Issues

- Need adequate “firewall” from project personnel
  - Plans for interim analyses
    - Details documented in separate “closed” protocol
  - Results of adaptations
- Need prospectively defined plan for communicating information on results to the appropriate parties (only when necessary)
  - Steering Committee
  - Sponsor
- SAD-PD example

Summary

- Adaptation should be considered an integral design feature
  - Requires careful and extensive planning
- When used appropriately, adaptive designs can be efficient and highly informative
  - Being emphasized in exploratory phases (learning)
- When used inappropriately, adaptive designs can threaten trial validity and integrity

Summary

- There are many logistical and procedural issues that are introduced by the possibility of adaptation
  - Careful evaluation of feasibility
- Trial integrity should be preserved by minimizing access to information on interim analyses and their results
  - Control of operational bias