

# Promoting the Use of Complex Innovative Designs in Clinical Trials

March 20, 2018



### WELCOME AND OPENING REMARKS

#### Aloka Chakravarty, PhD

Acting Director, Office of Biostatistics
Office of Translational Science
Center for Drug Evaluation and Research
U.S. Food and Drug Administration





- Session I: General Considerations for Complex Adaptive Clinical Trial Designs to Support the Effectiveness and Safety of Drugs or Biologics
- Session II: General Considerations for Other Innovative Designs Including External/Historical Control Subjects, Bayesian Designs, and Master Protocols
- Session III: Clinical Trial Simulations for Confirmatory Trial Design and Planning
- Session IV: Complex Innovative Design Pilot Program
- Closing Remarks



# Session 1 Complex Adaptive Designs

Gregory Levin
FDA/CDER/OTS/OB/DBII
March 20, 2018

## Outline



- Overview of adaptive clinical trial designs
- Considerations for adaptive designs (including complex adaptive designs)
- Questions for discussion





- A clinical trial design that allows for prospectively planned modifications to one or more aspects of the design based on accumulating data from subjects in the study
- Not within scope of today's discussion:
  - Unplanned changes based on comparative interim results
  - Protocol amendments based on information from sources external to study





- Adaptations based on baseline characteristics
- Adaptations based on pooled outcome data
- Adaptations based on comparative data
  - Group sequential designs
  - Adaptations to the sample size
  - Adaptations to the patient population
  - Adaptations to treatment arm selection
  - Adaptations to patient allocation
  - Adaptations to endpoint selection





- Adaptations to statistical aspects of design (primary estimand typically does not change)
  - Group sequential designs
  - Adaptations to sample size / statistical information, analysis schedule, decision criteria, randomization ratio
- Adaptations to scientific aspects of design (primary estimand does change)
  - Adaptations to patient population, treatment arm selection, endpoint selection





- Advantages in statistical efficiency
  - e.g., a greater chance of detecting a drug effect at a given expected sample size
- Ethical advantages
  - e.g., stop trial if data not consistent with an effective drug (futility) or if persuasive evidence of important effect (efficacy)
- Advantages in understanding of drug effects
  - e.g., improved estimation of dose-response relationship in study with adaptive dose selection





- Methodology challenges in ensuring control of chance of erroneous conclusion, reliability in estimation (e.g., bias, Cl coverage)
- Operational challenges in maintaining confidentiality and trial integrity
- Potential challenges in interpretability due to changes in estimand of interest





- Often include multiple types of adaptations (e.g., to treatment arm and sample size)
- Often include adaptations to scientific aspects of design (i.e., to the estimand)
- Often involve simulations to evaluate operating characteristics
- Examples
  - PREVAIL II (to evaluate ZMapp for Ebola virus disease)
  - I-SPY 2 (to screen breast cancer treatments)





- Control of chance of erroneous conclusions
- Reliability of estimation of treatment effects
- Extent of pre-specification of details of design
- Confidentiality to comparative interim results and preservation of trial integrity
- Documentation
- Additional considerations



## Controlling the chance of erroneous conclusions

- Incorrect conclusions of safety or effectiveness, incorrect conclusions of lack of safety or effectiveness, misleading estimates contributing to evaluation of benefit-risk
- Effectiveness typically evaluated through test of null hypothesis
  - Adaptive designs can inflate type I error probability
  - Consider testing methods with error probability control supported by theory or comprehensive simulation





- Accurate and precise estimates facilitate benefit-risk evaluation and appropriate labeling/reporting to enable evidence-based medicine
- Adaptations induce bias in estimates, incorrect confidence interval coverage
- Some methods exist for adjusting estimates





- Extent of prospective planning can vary
  - Could include anticipated number and timing of interim analyses, type of adaptation, statistical methods for interim and final analysis, algorithm governing adaptation decision
- Possible motivation for pre-specification
  - Facilitates use of appropriate inferential methods for many types of adaptations
  - Increases confidence that adaptations not based on accumulating knowledge in unplanned way
  - Motivates careful planning, reduces desire for sponsor access to comparative interim data, ensures that DMC (if involved in adaptive process) focuses on patient safety and trial integrity

## Maintaining trial integrity



- Recommended in ICH E9 that access to comparative interim results in all trials limited to individuals independent of personnel conducting or managing the trial
  - Knowledge can affect conduct of sponsor, investigators, participants in ways difficult to predict and adjust for
  - Provides confidence that unplanned design changes based on external information not motivated by accumulating data
- Additional considerations with adaptive design
  - Dedicated adaptation committee versus DMC, confidentiality agreements, physical/logistical firewalls, data access plan, steps to minimize knowledge inferred through adaptation





- Can be more comprehensive, may include:
  - Rationale for design, comparison to alternatives
  - Evaluation of important operating characteristics
  - Adaptation plan
  - Monitoring plan
  - Data access plan
  - Simulation report





- Use of simulations in planning
- Bayesian adaptive designs
  - Use of explicit borrowing with informative priors
- Adaptations in time-to-event settings
- Adaptations based on potential surrogate or intermediate endpoints
- Evaluation of secondary endpoints
- Safety considerations
- Adaptations in early-phase exploratory studies





What are the two to three most important principles for ensuring the appropriate and effective use of complex adaptive designs?



## Question #2 for discussion

Discuss the extent to which complex adaptive designs should be pre-specified. For example, discuss the importance of pre-specification of the specific algorithm that will be used to determine adaptive decision-making.



## Question #3 for discussion

Bias in treatment effect estimation is currently less well-studied than Type I error probability control in the context of complex adaptive designs. How important is the evaluation of the properties of point and interval estimates? Should adjusted estimates be included in labeling and reporting of results?

## Session II: General Considerations for Other Innovative Designs

Lisa M. LaVange, PhD
Department of Biostatistics
Gillings School of Global Public Health
University of North Carolina
Chapel Hill, NC

FDA Public Meeting Complex Innovative Designs White Oak Campus March 20, 2018

## Complex Innovative Designs

#### PDUFA VI and 21st Century Cures Act Commitment

- Complex Innovative Design project
  - Focus is complex adaptive and Bayesian designs requiring computer simulation to determine operating characteristics (power, Type I error)
- Milestones
  - Public meeting
  - Guidance
- Pilot program
  - Sponsors gain increased interaction with agency on design, simulations, etc.
  - Agency allowed to publicly discuss design prior to approval

## Complex Innovative Designs

- Adaptive designs that are complex, due to:
  - Adaptations on multiple factors, and/or
  - Requiring simulations to determine operating characteristics
- Other designs incorporating
  - Innovative use of external data
  - Innovative criteria for decision-making
  - Innovative collaborative efforts

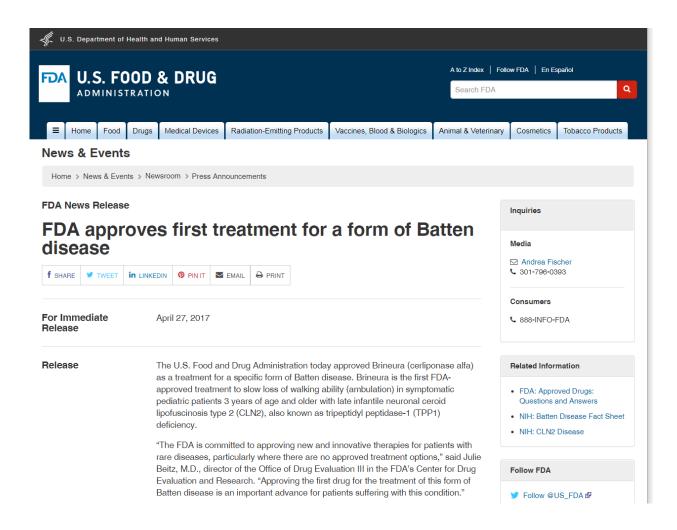
## **Complex Innovative Designs**

- Why the need to innovate?
  - Small populations: leverage other data sources to provide additional power
  - Improve decision-making when reliable prior information is available
  - Optimize product development with coordinated trial structures
  - →Ensure the trial will be able to answer the relevant questions and provide regulators with information needed for decisions
- Why the need for FDA guidance?
  - To better understand CDER/CBER's acceptance of innovative designs
  - To better understand CDER/CBER's expectations for information needed for submissions involving complex trials
  - To ensure consistency of advice about and acceptance of complex trials across therapeutic areas

## Small or Limited Populations

- Rare diseases -- interest in designs that use:
  - External control patients from patient registries or natural history studies
  - External control patients from control arms of earlier phase trials
  - Information on disease progression from natural history studies to improve analytical model
  - Prior information on treatment effects from earlier phase trials
  - Endpoints that maximize power in presence of disease heterogeneity

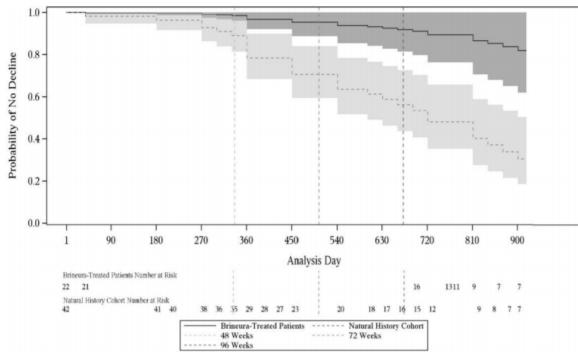
## Rare Disease Example



#### Brineura label

Given the non-randomized study design, a Cox Proportional Hazards Model adjusted for age, initial motor score, and genotype was used to evaluate time to unreversed 2-category decline or unreversed score of 0 in the Motor domain. This model showed a lesser decrease in motor function in the Brineura-treated patients when compared to the natural history cohort (see Figure 7).

Figure 7. Estimated Time to Unreversed (Sustained) 2-Category Decline or Unreversed Score of Zero in Motor Domain for Symptomatic Pediatric Patients in the Brineura Single-Arm Clinical Study with Extension and for Patients in a Natural History Cohort (Based on the Cox Proportional Hazards Model Adjusting for Covariates)



Shading represents 95% confidence intervals.

Follow-up for the natural history cohort begins at 36 months of age or greater and at the first time a Motor plus Language CLN2 score less than 6 was recorded.

The Brineura-treated population is the full population (N=24) minus two patients with baseline Motor plus Language CLN2 score = 6. Covariates: screening age, screening Motor score, genotype: 0 key mutations (yes/no). "Screening age" was defined in the natural history cohort as the age at the first time a Motor plus Language CLN2 score less than 6 was recorded, and no earlier than 36 months of age. The "screening Motor score" of the natual history cohort was defined as the Motor score at the screening age. Decline is defined as an unreversed (sustained) 2-category decline or unreversed score of 0 in the Motor domain of the CLN2 Clinical Rating Scale.

## **Bayesian Applications**

#### Safety monitoring

 Large CV risk studies that leverage control patient data from other sources via Bayesian adaptive designs

#### Oncology

- Early phase dose-finding trial designs, e.g., CRM
- Bayesian adaptive trials that use intermediate or accelerated approval endpoints for decision-making

#### Rare diseases

- Incorporate prior information from early phase trials
- Use information about disease progression in analytical model
- Compute shrinkage estimators of effects in rare subsets of disease
- Incorporate prior information from adult trials to improve efficiency of pediatric trials

## Adaptive Phase 2b/3 Trial Example

#### **CLINICAL STUDY DESIGN**

#### Rationale and Design of an Adaptive Phase 2b/3 Clinical Trial of Selepressin for Adults in Septic Shock

Selepressin Evaluation Programme for Sepsis-induced Shock—Adaptive Clinical Trial

Roger J. Lewis<sup>1,2,3,4</sup>, Derek C. Angus<sup>5,6</sup>, Pierre-François Laterre<sup>7</sup>, Anne Louise Kjølbye<sup>8</sup>, Egbert van der Meulen<sup>8</sup>, Allan Blemings<sup>8</sup>, Todd Graves<sup>4</sup>, James A. Russell<sup>9</sup>, Jan E. Carlsen<sup>10</sup>, Karsten Jacobsen<sup>8</sup>, Donald M. Yealy<sup>11</sup>, Steven M. Opal<sup>12</sup>, Nis A. Windeløv<sup>8</sup>, Bruno François<sup>13</sup>, Anders Perner<sup>14</sup>, Peter Pickkers<sup>15</sup>, and Scott M. Berry<sup>4</sup>

Department of Emergency Medicine, Harbor–UCLA Medical Center, Torrance, California; "Los Angeles Biomedical Research Institute, Torrance, California; "Department of Emergency Medicine, David Geffen School of Medicine at UCLA, Los Angeles, California; "Berry Consultants, LLC, Austin, Texas; "Glinical Research, Investigation, and Systems Modeling of Acute Illness Center, Department of Critical Care Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; "Department of Critical Care Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; "Department of Critical Care Medicine, St. Luc University Hospital, Université Catholique de Louvain, Brussels, Belgium; "Ferring Pharmaceuticals A/S, Copenhagen, Denmark; "Ocnarer for Heart Lung Innovation and the Division of Critical Care Medicine, St. Paul's Hospital, University of British Columbia, Vancouver, British Columbia, Canada; "StraDevo A/S, Kongens Lyngby, Denmark; "Department of Emergency Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; '2"Infectious Disease Division, Alpert Medical School of Brown University, Providence, Rhode Island; "3"Centre d'Investigation Clinique 1435 and Intensive Care Unit, Centre Hospitalier Universitaire, Limoges, France; "4"Department of Intensive Care, Rigshospitalet, Copenhagen, Denmark; and "5"Department of Intensive Care Medicine, Radboud University Medical Center, Nijmegen, the Netherlands

#### **Abstract**

Septic shock carries substantial morbidity and mortality. The failure of many promising therapies during late-phase clinical trials prompted calls for alternative trial designs. We describe an innovative trial evaluating selepressin, a novel selective vasopressin V<sub>1a</sub> receptor agonist, for adults with septic shock. SEPSIS-ACT (Selepressin Evaluation Programme for Sepsis-induced Shock-Adaptive Clinical Trial) is a blinded, randomized, placebocontrolled, two-part, adaptive phase 2b/3 trial, evaluating up to four selepressin dosing strategies. The primary outcome is pressorand ventilator-free days, with a value of zero assigned for death within 30 days. We calculate Bayesian probabilities of final trial success to guide interim decision-making. Part 1 (dose-finding) has an adaptive sample size based on response-adaptive randomization and prespecified rules to determine stopping for futility or selection of the best dosing regimen for Part 2. Part 2 (confirmation) randomizes a minimum of 1,000 patients equally

to the selected dosing regimen or placebo. The final estimate of treatment effect compares all selepressin-treated patients with all placebo-treated patients. The sample size of 1,800 provides 91% power to detect an increase of 1.5 pressor- and ventilator-free days with a reduction in mortality of 1.5%. The trial received a Special Protocol Assessment agreement from the U.S. Food and Drug Administration Center for Drug Evaluation and Research and is underway in Europe and the United States. SEPSIS-ACT is an innovative trial that addresses both optimal dose and confirmation of benefit, accelerating the evaluation of selepressin while mitigating risks to patients and sponsor through use of response-adaptive randomization, a novel registration endpoint, prespecified futility stopping rules, and a large sample size.

Clinical Trial registered with www.clinicaltrials.gov (NCT02508649).

Keywords: septic shock; adaptive clinical trial design; vasopressor treatment

## **Master Protocols**

- Multiple diseases, multiple patient subgroups (biomarker-defined), and/or multiple therapies studied under one, over-arching protocol\*
  - I-SPY 2, Lung-MAP, DIAN-TU, ADAPT
- Areas of innovation:
  - Establish a trial network with infrastructure in place to streamline trial logistics, improve data quality, and facilitate data sharing and new data collection
  - Develop a common protocol for the network that incorporates innovative statistical approaches to study design and data analysis

<sup>\*</sup>Woodcock J, LaVange LM. Master Protocols to Study Multiple Therapies, Multiple Disease, or Both. N Engl J Med 2017; 377:62-70

## Innovative Design Possibilities

- Adaptive randomization and/or adaptive enrichment
- Use of external or historical control data
  - In conjunction with concurrent controls (with 2:1 or higher randomization ratios); potential adaptation to ratio based on similarity between two sources of controls
- Sharing of control groups across protocols within a specific pathway or marker subgroup
- Model-based analysis methods (e.g., hierarchical Bayes) for pooled analysis of multiple disease or tumor types, markers, body sites, etc.

#### The NEW ENGLAND JOURNAL of MEDICINE

#### REVIEW ARTICLE

#### THE CHANGING FACE OF CLINICAL TRIALS

Jeffrey M. Drazen, M.D., David P. Harrington, Ph.D., John J.V. McMurray, M.D., James H. Ware, Ph.D., and Janet Woodcock, M.D., Editors

#### Master Protocols to Study Multiple Therapies, Multiple Diseases, or Both

Janet Woodcock, M.D., and Lisa M. LaVange, Ph.D.

From the Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD. Address reprint requests to Dr. LaVange at the Office of Biostatistics, Office of Translational Sciences, Center for Drug Evaluation and Research, Food and Drug Administration, 10903 New Hampshire Blvd., Silver Spring, MD 20993, or at lisa.lavange@fda.hhs.gov.

N Engl J Med 2017;377:62-70.
DOI: 10.1056/NEJMra1510062
Copyright © 2017 Massachusetts Medical Society.

IGH-QUALITY EVIDENCE IS WHAT WE USE TO GUIDE MEDICAL PRACTICE. The standard approach to generating this evidence — a series of clinical trials, each investigating one or two interventions in a single disease — has become ever more expensive and challenging to execute. As a result, important clinical questions go unanswered. The conduct of "precision medicine" trials to evaluate targeted therapies creates challenges in recruiting patients with rare genetic subtypes of a disease. There is also increasing interest in performing mechanism-based trials in which eligibility is based on criteria other than traditional disease definitions. The common denominator is a need to answer more questions more efficiently and in less time.

A methodologic innovation responsive to this need involves coordinated efforts to evaluate more than one or two treatments in more than one patient type or disease within the same overall trial structure. L4 Such efforts are referred to as master protocols, defined as one overarching protocol designed to answer multiple questions. Master protocols may involve one or more interventions in multiple diseases or a single disease, as defined by current disease classification, with multiple interventions,

## Platform Trial Example

- 2014-2016 Ebola Outbreak
  - Urgent need to identify safe and effective therapies
  - Limited or intermittent drug supply for several potential therapeutic agents
  - Desire to maximize information from potentially limited data
  - Flexible design and analysis needed
- Prevail II \*
  - Shared control arm
  - Ability to simultaneously evaluate multiple therapies
  - Add or remove treatment arms
  - Bayesian stopping rules

<sup>\*</sup>The PREVAIL II Writing Group. A Randomized, Controlled Trial of ZMapp for Ebola Virus Infection. N Engl J Med 2016; 375:1448-56.

## Ebola Platform Trial Example

Design

#### CLINICAL TRIALS

## Statistical considerations for a trial of Ebola virus disease therapeutics

Clinical Trials
2016, Vol. 13(1) 39–48
© The Author(s) 2016
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1740774515620145
ctj.sagepub.com



Michael A Proschan<sup>1</sup>, Lori E Dodd<sup>1</sup> and Dionne Price<sup>2</sup>

#### **Abstract**

The 2014 West African outbreak of Ebola virus ravaged Liberia, Sierra Leone, and Guinea, causing hemorrhagic fever and death. The need to identify effective therapeutics was acute. The usual drug development paradigm of phase I, followed by phase II, and then phase III trials would take too long. These and other factors led to the design of a clinical trial of Ebola virus disease therapeutics that differs from more conventional clinical trial designs. This article describes the Ebola virus disease medical countermeasures trial design and the thinking behind it.

#### **Keywords**

Barnard's test, Bayesian methods, beta-binomial distribution, conditional power, emerging infectious diseases, Fisher's exact test, group-sequential monitoring, non-informative prior

## **Discussion Questions**

1. What types of innovative trial designs would facilitate the advancement of drug development, particularly in areas of unmet medical needs, such as rare diseases, antimicrobial agents, etc.?

## **Discussion Questions**

2. What factors impact the perceived acceptability of innovative designs by sponsors and regulatory agencies?

## **Discussion Questions**

3. Are there additional outreach or research activities or areas for collaboration that might further advance the use and acceptance of these and other innovative designs?



# Clinical trial simulations for confirmatory trial design and planning

John Scott, Ph.D., FDA/CBER/OBE
Public Meeting on Promoting the Use of Complex
Innovative Designs in Clinical Trials
March 20, 2018



#### Overview

- Clinical trials have a variety of important operating characteristics
  - Expected behavior under clinical, operational and statistical assumptions
- These operating characteristics guide trial design and interpretability
- One way of estimating trial operating characteristics is to simulate large numbers of clinical trials and observe their outcomes



## Simulation: When and why

- For many simple and some complicated clinical trial designs, statistical theory provides estimates or bounds on trial operating characteristics
- Simulations may be preferable or necessary for:
  - Complex designs with multiple adaptations
  - Bayesian trial designs
  - Small sample designs (e.g. rare diseases) where asymptotics may be unreliable



# Operating characteristics

- Type I error probability
  - Other clinically relevant error probabilities
- Power
  - Possibly under various alternatives
- Expected sample size
- Estimation properties
- Bayesian alternatives: maximum posterior probability of the null in rejection region, Bayesian average errors, etc.



#### Basic idea

- To estimate Type I error probability:
  - Assume the null hypothesis is true
  - Generate trial data under that hypothesis
  - Apply trial analyses and decision rules to that data
  - Repeat large number of times
  - Proportion of positive decisions is an estimate of Type I error probability



# Complications

- Definition of null space
- Scope of simulations
- Multiple testing / multiple hypotheses
- Application in Bayesian settings with informative priors
- Resource issues



# Definition of null space

- Typically there are many ways for a drug to be ineffective
- Consider a drug for a very aggressive cancer, with a historical oneyear median survival
  - In new trial, drug could be same as control with one-year median survival
  - Drug could also be same as control with a 20-year or (mathematically) a 1000-year median survival
- Do we simulate all possible null configurations or just clinically plausible configurations?
  - How should the line be drawn?



## Scope of simulations

- In addition to treatment effect, typically assumptions needed about many other parameters
  - Clinical parameters (e.g. control rate)
  - Statistical nuisance parameters (e.g. variance)
  - Operational parameters (e.g. accrual rate)
- How many different combinations of parameters need to be explored?
- How fine of a grid?



# Multiple testing

- Simulation focus is typically on primary analysis of primary efficacy endpoint
- Decisions are complicated
  - Multiple primary and secondary endpoints may be involved
  - Safety is involved
- Can simulation encompass all of this?



# Bayesian settings

- Many "Bayesian" clinical trial designs use Bayesian calculations but rely on frequentist operating characteristics
  - When these do not borrow prior information, considerations same as above
  - When they do borrow prior information, the definition of the null space becomes quite murky – borrowing means conditioning on known data
- True Bayesian designs (e.g. designs that follow the likelihood principle and base inference on posterior probability interpretations) raise different issues
  - Simulation still important, but interpretation different



#### Resource issues

- Simulations can be computationally intensive
  - Constant progress via improved algorithms, parallel computing and other innovations
  - Some problems truly infeasible to simulate
    - Combinations of multiple endpoints, adaptive allocation, MCMC inference steps and/or permutation tests....
- Reviewing simulations is resource-intensive for FDA
  - Timelines, workload, skills/training, software



## Simulation reports

- Best practices developing on format
- Features generally include:
  - Overview
  - Trial design
  - Hypothetical trial outcomes
  - Scenarios simulated
  - Results
  - Summary
- May sometimes include simulation code, technical details, statistical derivations



#### FDA review of simulations

- Scope of review varies best practices developing
- Review may include:
  - Verification with applicant's simulation code or off-the-shelf software
  - Verification with reviewer-generated code or other software
  - Exploration of additional scenarios
- No standardization of acceptance criteria for operating characteristics



Regarding the scope of Type I error probability simulations:

- Should all mathematically possible parameter values for which the drug is ineffective be included, or only values that are in some sense clinically plausible?
- How is clinically plausible defined / agreed to?



- How should error rate simulations be conducted when formally borrowing prior information, such as in a Bayesian framework?
- What does Type I error mean in this setting? Should we consider other error rates instead?



- What are some practical suggestions for implementing trial simulations?
  - For example, number of iterations, computational details, documentation details



- How extensively should the parameter space be explored in simulations?
  - Is it important to evaluate every possible combination of nuisance parameter values / ranges?
  - When is a grid of assumptions comprehensive enough?



## Complex Innovative Designs Pilot Program



Dionne L. Price, Ph.D.

Acting Deputy Director, Office of Biostatistics

OTS/CDER/FDA

March 20, 2018



# PDUFA VI Complex Innovative Design (CID)

- Objective: To facilitate the advancement and use of complex adaptive, Bayesian, and other novel clinical trial designs through
  - Development of staff capacity
  - Conducting a pilot program
  - Convening a public workshop
  - Publishing draft guidance
  - Developing or revising relevant MAPPs, SOPPs, and/or review templates



## CID Pilot Program

- Designed for highly innovative trial designs which may require simulations to determine operating characteristics
- Sponsors
  - submit designs (up to two per quarter selected)
  - have the opportunity to engage with regulatory staff on designs via two meetings
- Agency
  - uses the design as a case study for continuing education and information sharing
- Program will be announced in the Federal Register



# CID Pilot Program

- FDA will grant two meetings
  - -Initial and follow-up meeting on the same design
  - Occur within a span of approximately 120 days
  - Led by the statistical review components within CDER or CBER



#### Points to Consider

- Eligibility Criteria
- Timelines
- Submission expectations
- Disclosure
- Communication





#### Discussion Point 1

• The FDA will select two proposals quarterly for entry into the pilot program. The proposals will need to capture sufficient details to facilitate an understanding of the design and analysis. Discuss specific elements of the design and analysis that are important for the initial proposal.



## Discussion Points 2 and 3

 Discuss types of trial designs that should be prioritized for selection into the pilot program.

 Discuss factors that might inhibit or encourage submissions for the program.