Innovating the Clinical Trial: Adaptive and Platform

Scott M. Berry
Jan 15, 2020
@statberry
scott@berryconsultants.com

NISS - Merck Meet-up for Adaptive Design for Drug Development
Austin Bradford Hill

• Credited with designing the first randomized clinical trial in humans

• Traditional looking Table 1 and Table 2
• Streptomycin very effective!
• “typo” (4/55 = 7.3%)

• The clinical trial didn’t change much for the next 60 years...

• The clinical trial is now the ‘science’ being innovated!
Outline

• Some comments on FDA Adaptive Design Guidance and FDA draft Guidance on CID
• Couple quick examples of complex adaptive designs
• Moving to platform trials
• And beyond?
Two Key Guidances

Adaptive Designs for Clinical Trials of Drugs and Biologics
Guidance for Industry

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)

November 2019
Biostatistics

Interacting with the FDA on Complex Innovative Trial Designs for Drugs and Biological Products

Draft Guidance for Industry

Additional copies of this guidance are available from:
Office of Communication, Outreach and Development
Center for Biologics Evaluation and Research
Food and Drug Administration
10901 New Hampshire Avenue, WO71, Room 3128
Silver Spring, MD 20993
Phone: 800-855-4789 or 240-846-8010
E-mail: druginfo@fda.hhs.gov
https://www.fda.gov/announcements-guidance Straßen and Regulatory Information
Guidances for Drugs

or

Office of Communications
Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10901 New Hampshire Ave., Hillblade Bldg., 4th Floor
Silver Spring, MD 20993
Phone: 301-796-2400 or 855-543-1784; Fax: 301-432-6533
E-mail: druginfo@fda.hhs.gov
https://www.fda.gov/drug/announcements-guidance/Regulatory information/guidance-drugs

Berry Consultants
Utilization Innovation
Discussion of Guidances

• Assuming you all have read them.... I’ll react to several important aspects of them
• Existence may be the most important thing
• They are in NO way limiting
• Completely opens this concept of adaptive designs and complex trials as “good”
• They are forward looking and do not restrict tomorrow’s innovation of the trial design
Snippets of AD Guidance (I added color)

• In still other cases, such as many Bayesian adaptive designs (section VI.B.), it may be critical to use simulations (section VI.A.) to evaluate the chance of an erroneous conclusion.

• The choice of scale is relatively unimportant as long as the operating characteristics of the designs are adequately evaluated.
Snippets of AD Guidance (I added color)

• The second type is response-adaptive randomization, an adaptive feature in which the chance of a newly-enrolled subject being assigned to a treatment arm varies over the course of the trial based on accumulating outcome data for subjects previously enrolled....

• Response-adaptive randomization alone does not generally increase the Type I error probability of a trial when used with appropriate statistical analysis techniques.
• In almost all cases, there are an infinite number of scenarios potentially compatible with the null hypothesis. **Identifying which scenarios should be considered when estimating Type I error probability** can be challenging and may rely on a combination of medical and mathematical considerations.

• It is common to perform simulations on a grid of plausible values and argue based on the totality of the evidence from the simulations that maximal Type I error probability likely does not exceed a desired level across the range covered by the grid.

• However, with any approach, the evaluation at the end of the trial should consider whether the statistical inference is appropriate and the conclusions are justified in light of the accumulated information about the nuisance parameters. In the example, if the observed placebo mortality rate was unexpectedly 50 percent, additional simulations would be required. “**POST-SIMULATION**”
In general, the same principles apply to Bayesian adaptive designs as to adaptive designs without Bayesian features. Trial designs that use Bayesian adaptive features may rely on frequentist or Bayesian inferential procedures to support conclusions of drug effectiveness.
Although CID has been referred to as complex adaptive, Bayesian, and other novel clinical trial designs, there is no fixed definition of CID because what is considered innovative or novel can change over time. For the purposes of this guidance, CID includes trial designs that have rarely or never been used to date to provide substantial evidence of effectiveness in new drug applications or biologics license applications.

A common feature of many CID designs is the need for simulations rather than mathematical formulae to estimate trial operating characteristics.

A simulation report if simulations are used to evaluate study operating characteristics.
Diabetes II/III Seamless

• 7 doses + PBO + Active Control
Diabetes II/III Seamless

- 7 doses + PBO + Active Control
  - Interims every 2 weeks
  - RAR based on 4 endpoints
    - HbA1c, Weight Loss, DBP, HR
      with utility function
Diabetes II/III Seamless

- **7 doses + PBO + Active Control**
  - Interims every 2 weeks
  - RAR based on 4 endpoints
    - HbA1c, Weight Loss, DBP, HR with utility function
  - 200-400 make decision:
    - Go to Phase III (pick 1 or 2 doses);
      spawn other phase III
    - Stop futility

- **Futility**
  - Go Part 2
  - Forced @ 400
Diabetes II/III Seamless

• 7 doses + PBO + Active Control
  – Interims every 2 weeks
  – RAR based on 4 endpoints
    • HbA1c, Weight Loss, DBP, HR with utility function
  – 200-400 make decision:
    • Go to Phase III (pick 1 or 2 doses); open more phase III
    • Stop futility
  – Phase III part powered by phase II

• Futility
• Go Part 2
• Forced @ 400
Diabetes II/III Seamless

- 7 doses + PBO + Active Control
  - Interims every 2 weeks
  - RAR based on 4 endpoints
    - HbA1c, Weight Loss, DBP, HR with utility function
  - 200-400 make decision:
    - Go to Phase III (pick 1 or 2 doses); open more phase III
    - Stop futility
- Phase III part powered by phase II
- Final Analysis includes all on arms continue (0.020 ANCOVA)
Modeling

- Bayesian repeated measures & dose-response models for four endpoints
- Single utility function connecting 4 endpoints on one scale
- Predictive probability of statistical success
Diabetes II/III Seamless

• Trial ran (for 3,467,321\textsuperscript{st} time!)
• Shifted at 200 -- very successful!
  – Ran \textit{exactly} as planned, spawned other phase III
  – Selected 0.75mg and 1.5mg doses

---

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

Zachary Skrivanek, Ph.D.,\textsuperscript{1} Scott Berry, Ph.D.,\textsuperscript{2} Don Berry, Ph.D.,\textsuperscript{2,3} Jenny Chien, Ph.D.,\textsuperscript{1} Mary Jane Geiger, M.D., Ph.D.,\textsuperscript{1} James H. Anderson, Jr., M.D.,\textsuperscript{4} and Brenda Gaydos, Ph.D.\textsuperscript{5}
An embattled Eli Lilly (LLY) won a major battle today, gaining the FDA's approval to market dulaglutide for diabetes.

“Projecting $1.3 billion in 2020”

With Novo Nordisk (NVO) already digging in to defend its position around Victoza, the once-weekly treatment has been widely billed as a likely blockbuster. The Phase III program has long represented Eli Lilly's best shot at opening a major new market, though many analysts are skeptical.

Peak sales projections for dulaglutide are all over the map. Cowen has pegged the potential at $700 million, with Bernstein's Tim Anderson now projecting $1.3 billion in 2020. That's not enough to make up for the patent losses, but it would go a long way to providing some credibility for an R&D group that is drawing an increasing amount of critical scrutiny.
### Established Pharma Products

<table>
<thead>
<tr>
<th>Products</th>
<th>2018</th>
<th>2017</th>
<th>% Change</th>
<th>2018</th>
<th>2017</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humalog®</td>
<td>$770.4</td>
<td>$762.2</td>
<td>(2)%</td>
<td>$2,996.5</td>
<td>$2,865.2</td>
<td>5%</td>
</tr>
<tr>
<td>Alimta</td>
<td>556.9</td>
<td>525.2</td>
<td>6%</td>
<td>2,132.9</td>
<td>2,062.5</td>
<td>3%</td>
</tr>
<tr>
<td>Cialis</td>
<td>350.7</td>
<td>597.4</td>
<td>(41)%</td>
<td>1,851.8</td>
<td>2,323.1</td>
<td>(20)%</td>
</tr>
<tr>
<td>Forteo</td>
<td>437.1</td>
<td>513.2</td>
<td>(15)%</td>
<td>1,575.6</td>
<td>1,749.0</td>
<td>(10)%</td>
</tr>
<tr>
<td>Humulin®</td>
<td>337.4</td>
<td>362.6</td>
<td>(7)%</td>
<td>1,331.4</td>
<td>1,335.4</td>
<td>(0)%</td>
</tr>
<tr>
<td>Cymbalta®</td>
<td>184.5</td>
<td>192.8</td>
<td>(4)%</td>
<td>708.0</td>
<td>757.2</td>
<td>(6)%</td>
</tr>
<tr>
<td>Erbitux®</td>
<td>159.8</td>
<td>168.9</td>
<td>(5)%</td>
<td>635.3</td>
<td>645.9</td>
<td>(2)%</td>
</tr>
<tr>
<td>Trajenta®</td>
<td>156.2</td>
<td>129.7</td>
<td>20%</td>
<td>574.7</td>
<td>537.9</td>
<td>7%</td>
</tr>
<tr>
<td>Zyprexa®</td>
<td>110.8</td>
<td>152.2</td>
<td>(27)%</td>
<td>471.3</td>
<td>581.2</td>
<td>(19)%</td>
</tr>
<tr>
<td>Strattera</td>
<td>107.2</td>
<td>98.3</td>
<td>9%</td>
<td>450.8</td>
<td>618.2</td>
<td>(27)%</td>
</tr>
</tbody>
</table>

### Select Products Launched Since 2014

<table>
<thead>
<tr>
<th>Product</th>
<th>2018</th>
<th>2017</th>
<th>% Change</th>
<th>2018</th>
<th>2017</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trulicity</td>
<td>924.7</td>
<td>649.0</td>
<td>42%</td>
<td>3,199.1</td>
<td>2,029.8</td>
<td>58%</td>
</tr>
<tr>
<td>Taltz</td>
<td>307.0</td>
<td>172.5</td>
<td>78%</td>
<td>937.5</td>
<td>559.2</td>
<td>68%</td>
</tr>
<tr>
<td>Cyramza®</td>
<td>220.6</td>
<td>204.8</td>
<td>8%</td>
<td>821.4</td>
<td>758.3</td>
<td>8%</td>
</tr>
<tr>
<td>Basaglar</td>
<td>232.2</td>
<td>153.8</td>
<td>51%</td>
<td>801.2</td>
<td>432.1</td>
<td>85%</td>
</tr>
<tr>
<td>Jardiance®</td>
<td>193.2</td>
<td>143.2</td>
<td>35%</td>
<td>658.3</td>
<td>447.5</td>
<td>47%</td>
</tr>
<tr>
<td>Lartruvo</td>
<td>83.5</td>
<td>59.0</td>
<td>41%</td>
<td>304.7</td>
<td>203.0</td>
<td>50%</td>
</tr>
<tr>
<td>Verzenio</td>
<td>83.1</td>
<td>21.0</td>
<td>NM</td>
<td>255.0</td>
<td>21.0</td>
<td>NM</td>
</tr>
<tr>
<td>Olumiant</td>
<td>70.1</td>
<td>23.0</td>
<td>NM</td>
<td>202.5</td>
<td>45.9</td>
<td>NM</td>
</tr>
<tr>
<td>Engalitx</td>
<td>4.9</td>
<td>—</td>
<td>NM</td>
<td>4.9</td>
<td>—</td>
<td>NM</td>
</tr>
<tr>
<td>Subtotal</td>
<td>2,119.4</td>
<td>1,426.3</td>
<td>49%</td>
<td>7,184.7</td>
<td>4,496.7</td>
<td>60%</td>
</tr>
</tbody>
</table>

### Animal Health

- 2018: $816.5
- 2017: $790.9
- % Change: 3%

### Total Revenue

<table>
<thead>
<tr>
<th>Total Revenue</th>
<th>2018</th>
<th>2017</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>6,438.6</td>
<td>6,160.7</td>
<td>5%</td>
<td>24,555.7</td>
</tr>
</tbody>
</table>

**$3.2 Billion in 2018**
• Endovascular Thrombectomy for ischemic stroke (approved ≤ 8 hours)
• New trial enrolling 6-24 hours since last seen well
• “Clinical Mismatch”
Adaptive Enrichment Design

- Interims at 150, 200, 250, 300, 350, 400, ... max of 500
- At 150, ..., 400 can “enrich” to smaller entry criterion
  - Infarct size of 0-30; 0-35; 0-40; 0-45
  - Restrict final analysis to the ‘restricted group’
  - Adjust CV for ‘cherry picking’
- Could Stop for Expected Success (at 200+ interims)
- Could Stop for Futility
DAWN Result

• At the 150-interim there was no enrichment
  – no futility
  – No expected success possible

• At 200-interim PP > 0.9999; no enrichment; stop for expected success

• Followed for 90 days; success at full data primary analysis (posterior probability superiority greater than 0.986)
Thrombectomy 6 to 24 Hours after Stroke with a Mismatch between Deficit and Infarct


This article was published on November 11, 2017, at NEJM.org.
DOI: 10.1056/NEJMoa1706442
Copyright © 2017 Massachusetts Medical Society.
RESULTS

A total of 206 patients were enrolled; 107 were assigned to the thrombectomy group and 99 to the control group. At 31 months, enrollment in the trial was stopped because of the results of a prespecified interim analysis. The mean score on the utility-weighted modified Rankin scale at 90 days was 5.5 in the thrombectomy group as compared with 3.4 in the control group (adjusted difference [Bayesian analysis], 2.0 points; 95% credible interval, 1.1 to 3.0; posterior probability of superiority, >0.999), and the rate of functional independence at 90 days was 49% in the thrombectomy group as compared with 13% in the control group (adjusted difference, 33 percentage points; 95% credible interval, 24 to 44; posterior probability of superiority, >0.999). The rate of symptomatic intracranial hemorrhage did not differ significantly between the two groups (6% in the thrombectomy group and 3% in the control group, \(P=0.50\)), nor did 90-day mortality (19% and 18%, respectively; \(P=1.00\)).

<table>
<thead>
<tr>
<th>Table 2. Efficacy Outcomes.*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome</strong></td>
</tr>
<tr>
<td>Primary end points</td>
</tr>
<tr>
<td>Score on utility-weighted modified Rankin scale at 90 days§</td>
</tr>
<tr>
<td>Functional independence at 90 days — no. (%)¶</td>
</tr>
</tbody>
</table>
A special case of adaptive treatment arm selection occurs in the context of an adaptive *platform trial* designed to compare more than one experimental treatment against an appropriate control for a disease (e.g., Woodcock and LaVange 2017).
Randomize to exp arm or ctrl

Update patient outcome data

Update longitudinal model: CA19-9 & imaging

Calculate PP Stage 1 arm > ctrl in each signature

Determine randomization prob within each subtype

New patient accrues; assess subtype

Add stage 1 arms accrual permitting

Continue in Stage 1

Go Stage 2;

Graduate (175)

Decision rules for Stage 1 arms

Monthly Interims

Precision Promise

Stop

futility

Stop

N=100

N=100

Go Stage 2;
Adaptation By Arm: Stage 1

• Potential Enrichment biomarker for an arm
• Adaptive Randomization across subtypes and arms
• Sample size between 50 and 100
• Stop enrolling (phase 2 trial) up until n=100
Stage 1 -> Stage 2 Transition

• Graduates at n=100 if predictive probability success at end of Stage 2 ≥ 0.65 in at least one signature (collection of an arms subtypes)
Stage 1

- Fixed randomization within graduating signature
- N=+75 in Stage 2
- Final analysis 1 year after last patient enrolled and on all Stage 1 and Stage 2 patients
- Comparison to all relevant shared controls (before and during era of an arm)
Adaptations for Patients

• Each patient is a member of a trial subtype
  – 1st line; 2nd line; potential biomarker group
  – Randomized 30% to control(s); 70% to experimental
  – The 70% of experimental broken up by RAR for arms by subtype and 40% Stage 2 arms (if applicable)

• Upon progression a patient is re-randomized as part of the platform (potentially a second experimental arm)
Statistical/Scientific Aspects (CID)

- Primary endpoint is overall survival
- Modeling partitions effect of 1\textsuperscript{st}/2\textsuperscript{nd} line therapies
- Common controls; effects of time modeled within the trial
- Single model estimate of effects of all arms
- RAR by subtype
- Potential graduation by signatures (model potential differential effects across subtypes)
- Simulation of type I error; interpretation of type I error
The Future?

• Embedded platform trials?
  – Merging clinical care and learning

• The beauty is that the future of innovation in trial design is awesome and restricted only by our imaginations…

• Tomorrow is here…