Adaptive Clinical Trials: Statistical, Ethical & Regulatory Considerations

Donald A. Berry*
dberry@mdanderson.org
Financial Disclosure

Co-owner of Berry Consultants, LLC.

Designs adaptive clinical trials for

- Medical device companies
- Pharmaceutical companies
- NIH cooperative groups
OVERALL SUCCESS AT PHASE II AND III

<table>
<thead>
<tr>
<th>Category</th>
<th>Phase II</th>
<th>Phase III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infect. Dis.</td>
<td>41%</td>
<td>55%</td>
</tr>
<tr>
<td>Endocrine</td>
<td>36%</td>
<td>60%</td>
</tr>
<tr>
<td>Autoimmune</td>
<td>30%</td>
<td>63%</td>
</tr>
<tr>
<td>Respiratory</td>
<td>24%</td>
<td>61%</td>
</tr>
<tr>
<td>Neurology</td>
<td>32%</td>
<td>55%</td>
</tr>
<tr>
<td>CV</td>
<td>28%</td>
<td>46%</td>
</tr>
<tr>
<td>Oncology</td>
<td>29%</td>
<td>34%</td>
</tr>
</tbody>
</table>

Last 20 Phase III Trials in AD: 0/20

16 February 2011
“Improved utilization of adaptive and Bayesian methods” could help resolve low success rate of and expense of phase 3 clinical trials
Example: Troxacitabine in AML*
(endpoint: CR by day 50)

Standard design

Randomize

- Idarubicin
- Trox
- Trox

Ara-C
Idarubicin
Ara-C

n = 25
n = 25
n = 25

* Giles JCO 2003
Example: Troxacitabine in AML* (endpoint: CR by day 50)

Our design

Adaptive randomization to learn, while effectively treating patients in trial

* Giles JCO 2003
Adaptive Randomization

- Assign with higher probability to better performing therapies
- TI dropped after 24th patient
- Trial stopped after 34 patients
Summary of AML trial results

CR by 50 days:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>IA</td>
<td>10/18 = 56%</td>
<td></td>
</tr>
<tr>
<td>TA</td>
<td>3/11 = 27%</td>
<td></td>
</tr>
<tr>
<td>TI</td>
<td>0/5 = 0%</td>
<td></td>
</tr>
</tbody>
</table>
Adaptive Randomization
Compared with
Balanced Randomization
Adaptive Randomization

IA

TA

TI
Adaptive Randomization: CRs in bold yellow
Adaptive Randomization: other 41 patients on IA

<table>
<thead>
<tr>
<th>IA</th>
<th>TA</th>
<th>TI</th>
</tr>
</thead>
</table>

Estim: 36/75 CRs (48%)
**Balanced Randomization**

<table>
<thead>
<tr>
<th>IA</th>
<th>TA</th>
<th>TI</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /></td>
<td><img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /></td>
<td><img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /> <img src="image1.png" alt="Male" /> <img src="image2.png" alt="Female" /></td>
</tr>
</tbody>
</table>

Estim: 21/75 CRs (28%)
“I see no rationale to further delay moving to these designs,” says Dr. Giles, who is currently involved in eight Bayesian-based leukemia studies. “They are more ethical, more patient-friendly, more conserving of resources, more statistically desirable.”
External impact of the trial?
The DSMB found that the study response rates were unlikely to provide evidence of a treatment benefit as a third-line treatment for patients with AML.
Adaptive Clinical Trials
A Partial Remedy for the Therapeutic Misconception?

William J. Meurer, MD, MS; Roger J. Lewis, MD, PhD; Donald A. Berry, PhD


There is a common “therapeutic misconception” among patients considering participation in clinical trials.\(^1\) Some trial participants and family members believe that the goal of a clinical trial is to improve their outcomes—a misperception often reinforced by media advertising of clinical research.\(^2\) Clinical trials have primarily scientific aims and rarely attempt to collectively improve the outcomes of their participants. The overarching goal of most clinical trials is to evaluate the effect of a treatment on disease outcomes.\(^3\) Comparisons are usually made with placebo for conditions having no established treatments and with standard care for conditions having effective treatments. Any benefit to an individual trial participant is a chance effect of randomization and the true, but unknown, relative effects of the treatments. Available evidence is conflicting regarding whether patients receive some benefit from simply participating in a clinical trial.\(^3\) Thus, even though serving as a research participant is essentially an altruistic activity, many patients are not aware of this component of participation.
I-SPY2

http://clinicaltrials.gov/ct2/show/NCT01042379?term=I-SPY2&rank=1

http://www.ispy2.org/

“uncovered a consensus that the two most important areas for improving medical product development are biomarker development and streamlining clinical trials.”

http://www.fda.gov/ScienceResearch/SpecialTopicsCriticalPathInitiative/default.htm
For example, in 2010, the Biomarkers Consortium—a public-private partnership that includes the NIH, the FDA, patient groups, and pharmaceutical and biotech—initiated a groundbreaking trial in breast cancer to predict drug responsiveness based on the presence or absence of genetic and biological markers, ... I-SPY 2 (ClinicalTrials.gov NCT01042379).
Driving Biomedical Innovation:
Initiatives to Improve Products for Patients
I-SPY 2

In March 2010, the I-SPY 2 Trial was launched. This is a groundbreaking clinical trial model that will help quickly and efficiently test promising drugs in development for women with high-risk, rapidly growing breast cancers. During the trial, drugs are individually targeted to the biology of each woman’s tumor. By applying an innovative trial design, researchers then use data from one set of patients’ treatments to decide treatment for future women who join the trial. This will help the researchers learn more quickly which investigational drugs will be most beneficial for women with certain biomarkers.
New trial design
Uses genetic profiles to highlight ‘biomarker’ differences among patients and to match drugs to patients with biomarkers that predict a benefit.

PHASE II
Randomized or non-randomized trials: about 60 patients are put in two groups: one group receives the experimental drug and the other serves as a control group. About 40 patients receive the experimental drug.

PHASE III
If a drug graduates to phase III, it typically takes 3,000 patients and about three years to determine if it is safe and effective enough for approval.

HISTORIC SUCCESS RATE
30 TO 40%

PHASE III
Researchers expect that drugs graduating from I-Spy 2 to phase III can be tested with 300 patients selected according to genetic profiles found to respond to the drug in phase II. It is hoped that this will shorten the time to approval.

PROBABILITY OF SUCCESS
85%

Source: Donald Berry, M.D. Anderson Cancer Center
A New Rx for Medicine

Fed up with slow drug trials, cancer patients and doctors are testing a fast track to personalized treatments.

By RON WINSLOW

PERSONALIZED MEDICINE | How redesigning a clinical trial can speed drug development

Traditional clinical trial
Takes essentially all patients with a disease being studied and is typically intended to eliminate differences in patient characteristics that could bias measures of drug effectiveness.

New trial design
Uses genetic profiles to highlight 'biomarker' differences among patients and to match drugs to patients with biomarkers that predict a benefit.

Phase I
Patients are placed in groups based on genetic profiles and are randomly assigned to either standard therapy or one of five different drugs plus standard care.

Early results increase chances that patients entering the trial later will be assigned to a drug showing benefit against tumors with their genetic profile.

It will take up to 120 patients for each drug to determine which ones graduate to phase III studies.

Note: In all clinical trials, phase I consists of testing on human subjects to determine toxicity levels.

Graphic by Marilyn Manley/WSJ

Phase II
Randomized or non-randomized trial: In a randomized trial, about 60 patients are put in two groups: One receives the experimental drug and the other serves as a control group. In a non-randomized trial, about 40 patients receive the experimental drug.

Less successful drugs are eliminated.

More successful drugs move on to phase III.

Phase III
If a drug graduates to phase III, it typically takes 3,000 patients and about three years to determine if it is safe and effective enough for approval.

HISTORIC SUCCESS RATE
30 TO 40%

Researchers expect that drugs graduating from I-Spy 2 to phase III can be tested with 300 patients selected according to genetic profiles found to respond to the drug in phase I. It is hoped that this will shorten the time to approval.

PROBABILITY OF SUCCESS
85%

Source: Donald Berry, M.D., Anderson Cancer Center
I-SPY 2 Adaptive Trial Design

Screening

Consent #1
Screening

MRI
Biopsy
Blood Draw
MUGA/ECHO
CT/PET

MRI
Biopsy
Blood Draw

MRI
Biopsy
Blood Draw

MRI
Blood Draw

MRI
Blood Draw

MRI
Blood Draw

Paclitaxel *
(12 weekly cycles)

Paclitaxel* +
Investigational Agent A
(12 weekly cycles)

Paclitaxel* +
Investigational Agent B
(12 weekly cycles)

AC
(4 cycles)

AC
(4 cycles)

AC
(4 cycles)

* HER2 positive participants will also receive trastuzumab. An investigational agent may be used instead of trastuzumab.
I-SPY2 TRIAL

Population of patients

ADAPTIVELY

RANDOMIZED

Experimental arm 1
Experimental arm 2
Experimental arm 3
Experimental arm 4
Experimental arm 5
Standard therapy

Outcome: Complete response at surgery
I-SPY2 TRIAL

Outcome: Complete response at surgery

Arm 2 graduates to small focused Phase 3 trial
I-SPY2 TRIAL

Outcome: Complete response at surgery

Population of patients

Arm 3 drops for futility

Experimental arm 1
Experimental arm 2
Experimental arm 3
Experimental arm 4
Experimental arm 5
Standard therapy

Complete response at surgery
I-SPY2 TRIAL

Outcome: Complete response at surgery

Arm 5 graduates to small focused Phase 3 trial

Population of patients

Adaptively randomized

Experimental arm 1

Experimental arm 4

Experimental arm 5

Standard therapy
I-SPY2 TRIAL

Outcome: Complete response at surgery

Arm 6 is added to the mix

Experimental arm 1
Experimental arm 4
Experimental arm 6
Standard therapy

Population of patients

Randomly adaptively
Goal: Greater than 85% success rate in Phase 3, with focus on patients who benefit
I-SPY-like TRIAL for Combinations

Population of patients

Adaptively randomize

Outcome: pathCR or PFS or OS

Trials:

- A + SOC
- B + SOC
- C + SOC
- D + SOC
- C + D + SOC

SOC
I-SPY-like TRIAL for Combinations

Substudy: Adaptively randomized factorial

Population of patients

Outcome: pathCR or PFS or OS

A + SOC

B + SOC

C + SOC

D + SOC

C + D + SOC

SOC
Simulations

- For operating characteristics:
  - Type I error rate
  - Power (many versions)
  - Sample size distribution

- Requires prospective design
  - (computationally intensive)

- Longitudinal info matters

- Accrual rate matters

- # arms matters
Experimental Drugs

- Sample size for each drug, 20 to 120 (minimum 60 if “success”)
- Up to 8 exp drugs at a time
- Patient enters trial, identify subtype
- Find (Bayesian) prob each drug >> control; based on all current results
- Covariate modeling (across subtypes)
- Assign in proportion to current prob drug >> control, by subtype
Dropping, Graduating Drugs

- For each possible biomarker signature $S$, calculate $\text{prob drug} >> \text{control in } S$
- If Bayesian pred $\text{prob 300-pt Phase III success} < 10\%$ for all $S$, drop drug
- If $> 85\%$ for some $S$ then drug graduates
- At graduation we provide predictive probability Phase III success for each $S$, including $S$ on drug’s diploma
I-SPY2 Adaptive Process

- Neoadjuvant breast cancer
- Considers 10 biomarker signatures
- Never-ending screening process
- Sponsored by FNIH: NCI, FDA, industry, academia
- Coordinated with FDA (CDER, CBER, & CDRH)—Regulatory pathway
- Status: 20 centers, ~380 pts so far, first 7 exp arms: neratinib, ABT888, AMG386, AMG 479, MK-2206, pertuzumab, pertuzumab+T-DM1
Effects of I-SPY Approaches

- Match drugs with biomarker signatures
- Savings from common control
- Better therapies move thru faster
- Successful drug/biomarker pairs graduate to small, focused, more successful Phase 3 based on Bayesian predictive probabilities
- I-SPY 2 offspring: melanoma, lymphoma, colorectal cancer, Alzheimer’s, HIV, acute heart failure, SARI/H1N1, …
Explaining Adaptive Designs to PIs, IRBs, Reviewers, ...

Example of Seamless Phase II/III
Treatment Arms Considered

Standard Rx: S

600 mg

Drug X 50 mg
Drug X 100 mg

No S

Drug X 50 mg
Drug X 100 mg

600 mg of S
No Drug X

Control

Experimental Arms
Trial Characteristics

- Primary endpoint: Time to event
- Minimum follow up: 9 months
- Type I error rate: 0.025
- 90% power at targeted HR 0.75
- No more than 750 patients
Design

• First 180 patients, 2:1:1:1:1
• Find Bayesian predictive probabilities of success
• If one is > 90% then select best “dose” and 1:1; accrual continues
• If all are < 10% then stop for futility
• Otherwise 60 more patients with adaptive randomization (1/3:p_1:p_2:p_3:p_4)
Design

● Repeat after 240 patients accrued

● After 300 accrued can
  ■ Stop accrual for futility
  ■ Stop accrual for win
  ■ Continue accrual for next 60, at 1:1

● Possible sample sizes: 180, 240, 300, 360, 420, ..., 750

● Minimum follow-up: 9 months
Flow Diagram

Burn-in
60:30:30:30:30
Flow Diagram

- Burn-in
  60:30:30:30:30

- Futility
- Adaptively randomize 60

- Dose found

- Failed trial

- 1:1 vs S for next 60
An Example Trial

One of 10,000s simulated to calculate operating characteristics
# Analysis 1 (n=180)

| Arm  | N  | #Prog | Expose | Med  | Pr(Best) | Pr(SS|N)  | Pr(SS|Cap) | Rand |
|------|----|-------|--------|------|----------|---------|-----------|------|
| S    | 60 | 41    | 755.5  | 12.77| 0.012    | --      |           | 0.33 |
| S+50 | 30 | 18    | 443.9  | 17.09| 0.260    | 0.091   | 0.622     | 0.18 |
| S+100| 30 | 18    | 393.9  | 15.17| 0.136    | 0.026   | 0.467     | 0.09 |
| 50   | 30 | 19    | 429.0  | 15.65| 0.149    | 0.023   | 0.542     | 0.10 |
| 100  | 30 | 16    | 434.0  | 18.80| 0.443    | 0.244   | 0.758     | 0.30 |
## Analysis 2 (n=240)

| Arm  | N   | #Prog | Expose | Med  | Pr(Best) | Pr(SS|N) | Pr(SS|Cap) | Rand |
|------|-----|-------|--------|------|----------|-------|----------|------|
| S    | 76  | 59    | 1038.2 | 12.20| 0.001    | --    |          | 0.33 |
| S+50 | 44  | 28    | 684.5  | 16.94| 0.270    | 0.222 | 0.741    | 0.19 |
| S+100| 34  | 29    | 554.1  | 13.24| 0.045    | 0.000 | 0.327    | 0.09 |
| 50   | 38  | 28    | 624.3  | 15.45| 0.124    | 0.030 | 0.592    | 0.09 |
| 100  | 48  | 27    | 738.9  | 18.97| 0.560    | 0.560 | 0.885    | 0.39 |
Analysis 3 (n=300)

| Arm  | N  | #Prog | Expose | Med  | Pr(Best) | Pr(SS|N) | Pr(SS|Cap) | Rand |
|------|----|-------|--------|------|----------|--------|----------|------|
| S    | 99 | 77    | 1354.5 | 12.19| 0.004    | --     | 0.50     | 0.50 |
| S+50 | 53 | 43    | 948.8  | 15.29| 0.121    | 0.030  | 0.601    | 0    |
| S+100| 37 | 33    | 622.8  | 13.08| 0.033    | 0.000  | 0.273    | 0    |
| 50   | 42 | 35    | 785.7  | 15.56| 0.146    | 0.030  | 0.614    | 0    |
| 100  | 69 | 42    | 1150.6 | 18.99| 0.696    | 0.796  | 0.937    | 0.50 |
### Analysis 4 (n=360)

| Arm  | N  | #Prog | Expose  | Med  | Pr(SS|N) | Pr(SS|Cap) | Rand |
|------|----|-------|---------|------|--------|----------|------|
| S    | 129| 106   | 1787.2  | 11.69| --     | --       | 0.50 |
| S+50 | 53 | 48    | 1085.6  | 15.68| --     | --       | 0    |
| S+100| 37 | 37    | 650.9   | 12.19| --     | --       | 0    |
| 50   | 42 | 38    | 880.5   | 16.06| --     | --       | 0    |
| 100  | 99 | 62    | 1664.7  | 18.61| 0.974  | 0.978    | 0.50 |
Analysis 4

Prob Survival

Weeks

S

X(100)
Accrual stopped at $N = 360$

9 months later:

Final analysis of $X(100)$ vs $S$

- Posterior probability $X(100)$ better than $S$ is $0.998 > 0.9875$ (superiority)
- Also, logrank $p = 0.002 < 0.0125$
### Operating Characteristics for accrual of 7 patients/week

#### Assumed Median TTE in Weeks

<table>
<thead>
<tr>
<th>Scenario</th>
<th>S</th>
<th>S+X(50)</th>
<th>S+X(100)</th>
<th>X(50)</th>
<th>X(100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>16</td>
<td>16</td>
<td>14</td>
<td>14</td>
</tr>
</tbody>
</table>

#### Probability of stopping for:

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Success</th>
<th>Futility</th>
<th>Cap</th>
<th>Mean Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>333</td>
</tr>
<tr>
<td>2</td>
<td>0.02</td>
<td>0.92</td>
<td>0.06</td>
<td>476</td>
</tr>
<tr>
<td>3</td>
<td>0.88</td>
<td>0.05</td>
<td>0.11</td>
<td>506</td>
</tr>
<tr>
<td>4</td>
<td>0.78</td>
<td>0.12</td>
<td>0.09</td>
<td>497</td>
</tr>
</tbody>
</table>