

Innovative Clinical Trial Design

- Donald A. Berry
- dberry@mdanderson.org

1

**Janet Woodcock
Director CDER FDA**

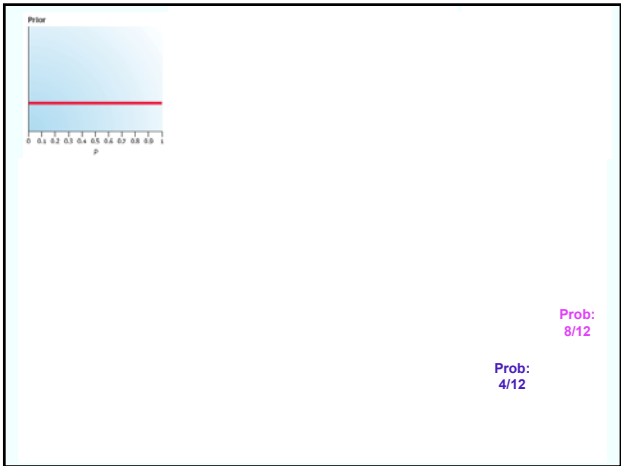
- 2006: “Improved utilization of adaptive and Bayesian methods” could help resolve the low success rate of and expense of phase III clinical trials
- 2013: FDA will need to “turn the clinical trial paradigm on its head” to allow personalized drug therapies to get on the market faster

OUTLINE

- Introduction to Bayes adaptive
- BATTLE trial in lung cancer
- I-SPY 2, brief intro for Jane
- Goldilocks and I-SPY 3
- Basket trials
- Decision analysis & rare disease

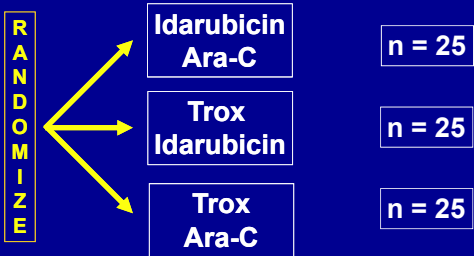
Bayesian Updating

- Paired observations, T vs C
- $P(S) = P(T \text{ wins pair})$
- $H_0: P(S) = 1/2$
- Data: SSFSS FSSSF
SFSSS SS



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

Standard design



* Giles JCO 2003

Example: Troxacitabine in AML*

(endpoint: CR by day 50)

Our design

Adaptive randomization to learn, while effectively treating patients in trial

R

A

N

D

O

M

I

Z

E

Idarubicin
Ara-C

Trox
Idarubicin

Trox
Ara-C

* Giles JCO 2003

7

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:

IA 10/18 = 56%

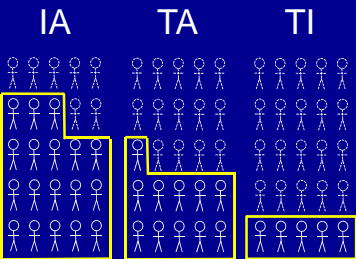
TA 3/11 = 27%

TI 0/5 = 0%

Adaptive Randomization Compared with Balanced Randomization

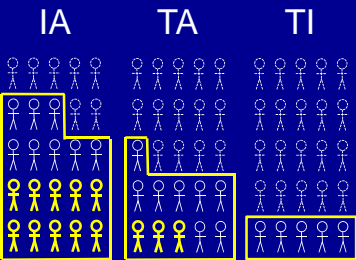
10

Adaptive Randomization



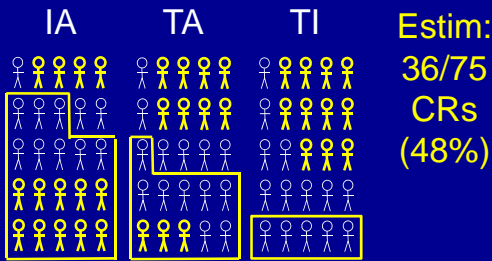
11

Adaptive Randomization: CRs in bold yellow



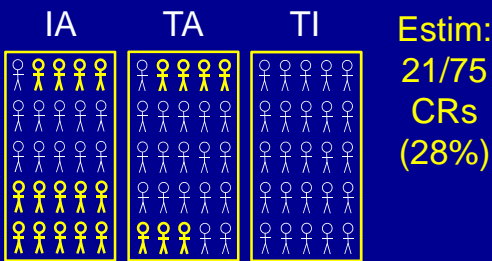
12

Adaptive Randomization:
other 41 patients on IA



13

Balanced Randomization



14

Cure Magazine (2006)

“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.”

15

However!!!!

External impact of the trial?

16

SGX Pharmaceuticals Discontinues Phase II/III Clinical Trial Of Troxatyl In Third-Line Acute Myelogenous Leukemia

Main Category: [Lymphoma / Leukemia / Myeloma](#) 0 tweets
Also Included In: [Clinical Trials / Drug Trials](#)
Article Date: 29 Aug 2006 - 0:00 PDT

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.

SGX Pharmaceuticals (SGX) today that it has discontinued the Phase II/III clinical trial of Troxatyl(TM) as a third-line treatment for patients suffering from acute myelogenous leukemia (AML), based upon the recommendation of the study's independent data and

Current Article Ratings:

Patient / Public: Not yet rated
Health Professional:  3 (1 votes)
Article Opinions: 0 posts

Find other articles on: "[troxatyl](#)"

safety monitoring board (DSMB). 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. The recommendation to discontinue the clinical trial was not made due to safety concerns.

"The response rates observed to date in our Phase II/III trial of Troxatyl are not at a level that we believe would support a New Drug Application as a third-line

JAMA The Journal of the American Medical Association

Home Current Issue All Issues Online First Specialties & Topics CME Multim

June 13, 2012, Vol 307, No. 22 >

Adaptive Clinical Trials
A Partial Remedy for the Therapeutic Misconception?

William J. Meurer, MD, MS; Roger J. Lewis, MD, PhD; Donald A. Berry, PhD
JAMA. 2012;307(22):2377-2378. doi:10.1001/jama.2012.4174.

Article References

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.³ 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

Can advocates influence the way we think about the purpose of clinical trials?

19

Joffe/Weeks, JNCI (2002)

“Many respondents viewed the main societal purpose of clinical trials as benefiting the participants rather than as creating generalizable knowledge to advance future therapy. This view, which was more prevalent among specialists such as pediatric oncologists that enrolled greater proportions of patients in trials, conflicts with established principles [from Belmont Report] of research ethics.”

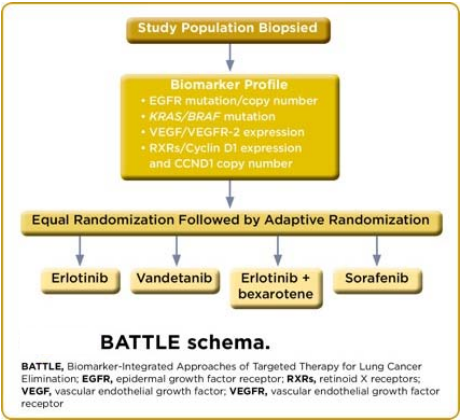
OUTLINE

- Introduction to Bayes adaptive
- **BATTLE** trial in lung cancer
- I-SPY 2, brief intro for Jane
- Goldilocks and I-SPY 3
- Basket trials
- Decision analysis & rare disease

BATTLE Trial in NSCLC

Kim et al. *Cancer Discovery* 2011

22



23

BATTLE results, N (Disease Control Rate)

	EGFR mut	KRAS/ BRAF mut	VEGF/ VEGFR	RXR/Cy cD1	None	Total
Erlotinib	17 (35%)	7 (14%)	25 (40%)	1 (0%)	8 (38%)	58 (34%)
Vandetanib	27 (41%)	3 (0%)	16 (38%)	0 (--)	6 (0%)	52 (33%)
Sorafenib	23 (39%)	14 (79%)	39 (64%)	4 (25%)	18 (61%)	98 (58%)
Total	87 (43%)	27 (48%)	83 (49%)	6 (33%)	41 (46%)	244 (46%)

Is EGFR wt a biomarker signature for sorafenib?

IN THE SPOTLIGHT

The BATTLE Trial: A Bold Step toward Improving the Efficiency of Biomarker-Based Drug Development

Eric H. Rubin, Keaven M. Anderson, and Christine K. Gause

Summary: Successful completion of the Biomarker-Integrated Approaches of Targeted Therapy for Lung Cancer

The precise biomarker hypotheses, as well as the associated type I and type II statistical errors, are not clear. Thus, the study should be considered as generating a hypothesis rather than as confirming a particular biomarker hypothesis.

Advances in basic cancer research have led to a widely used discovery and development approach for drugs designed to inhibit specific cancer pathways. However, clinical trial designs have not kept pace with basic research advances, and use of traditional, histology-based "all-comers" phase I and II trial designs for these drugs has led typically to failure in

Advances in basic cancer research have led to a widely used discovery and development approach for drugs designed to inhibit specific cancer pathways. However, clinical trial designs have not kept pace with basic research advances, and use of traditional, histology-based "all-comers" phase I and II trial designs for these drugs has led typically to failure in

Advances in basic cancer research have led to a widely used discovery and development approach for drugs designed to inhibit specific cancer pathways. However, clinical trial designs have not kept pace with basic research advances, and use of traditional, histology-based "all-comers" phase I and II trial designs for these drugs has led typically to failure in

Authors' Affiliation: Merck Research Laboratories, Kenilworth, Pennsylvania
Corresponding Author: Eric H. Rubin, Merck Research Laboratories, 351 North Sumner Street, North Wales, PA 19454, Phone: 267-305-1717, Fax: 267-305-6042; E-mail: eric_rubin@merck.com
doi:10.1158/2159-8274.CD-11-0036
© 2011 American Association for Cancer Research.

OUTLINE

- Introduction to Bayes adaptive
- BATTLE trial in lung cancer
- I-SPY 2, brief intro for Jane
- Goldilocks and I-SPY 3
- Basket trials
- Decision analysis & rare disease

I-SPY2

<http://www.ispy2.org>


Berry DA. Adaptive Clinical Trials.
Nature Reviews Clinical Oncology (2011)

October 2011

www.fda.gov/innovation

Driving Biomedical Innovation:

Initiatives to Improve Products for Patients



U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
U.S. FOOD AND DRUG ADMINISTRATION

28

To respond to these challenges, FDA will hold a series of scientific meetings with academic investigators, patient groups, drug developers, statistical and methodological experts, and ethicists to achieve a common understanding of steps that can be taken when an investigational drug being studied for a serious disease with no acceptable treatment option shows exceptional promise. CDER will then publish a draft guidance on an expedited development pathway based on the outcome of these meetings.

FDA is also working on two more immediate and related steps toward expedited drug development. First, the Agency is developing a draft guidance on enrichment strategies in clinical drug development. This is a major step forward for speeding progress for targeted

therapies and will lay out many strategies for selecting the patients most likely to benefit from a particular drug. These enrichment strategies are expected to improve the efficiency of clinical trials and serve as a source of expedited drug development.


Second, as a working example of an expedited pathway, CDER will publish a draft guidance on the use of pathologic complete response (pCR)—when no clinical evidence of a disease remains—as a surrogate endpoint for accelerated approval in primary high-risk breast cancer. This guidance will outline a relatively seamless pathway that could be followed from a multi-drug screening trial such as I-SPY 2 to an accelerated approval. This would speed the availability of targeted therapies for breast cancer. [FDA](#)

U.S. Food and Drug Administration / Driving Innovation

29

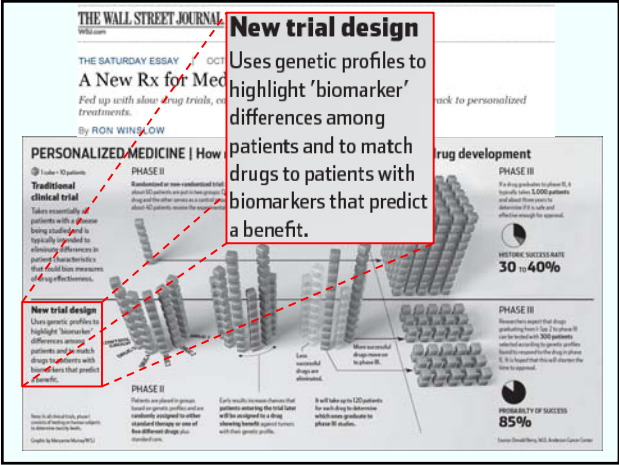
I-SPY 2

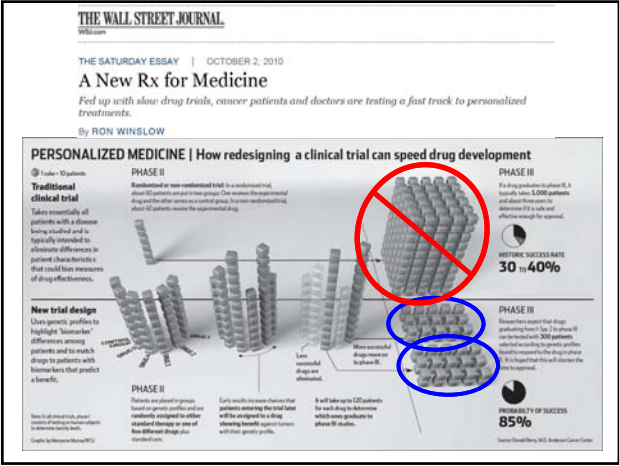
In March 2010, the I-SPY 2 Trial was launched. This is a groundbreaking clinical trial model that will help quickly and efficiently test promising drugs in development for women with high-risk, rapidly growing breast cancers. During the trial, drugs are individually targeted to the biology of each woman's 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.

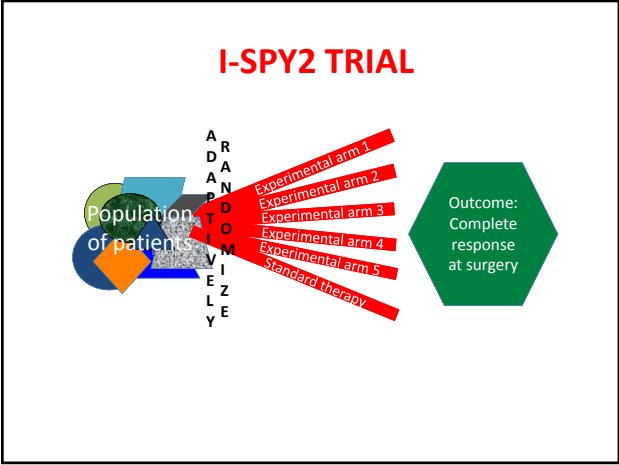


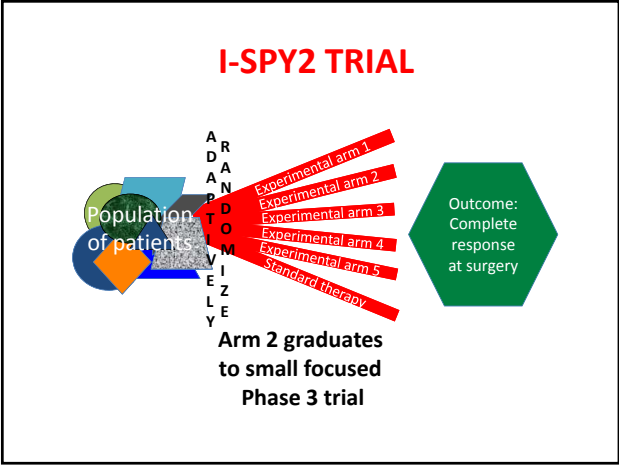
U.S. Food and Drug Administration / Driving Innovation

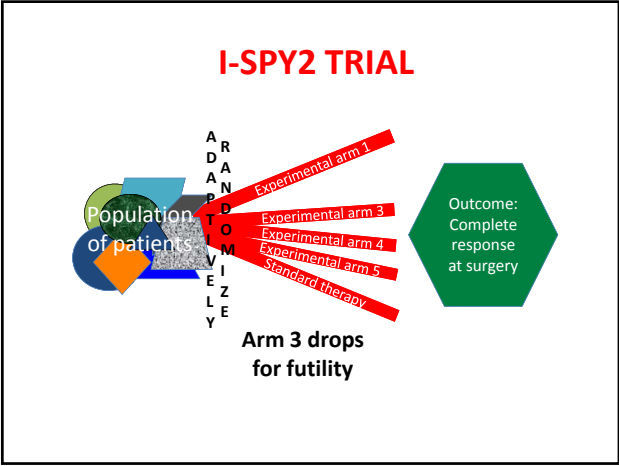
30

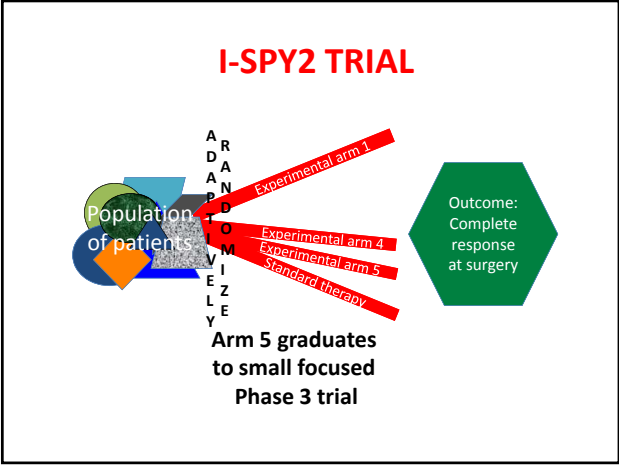


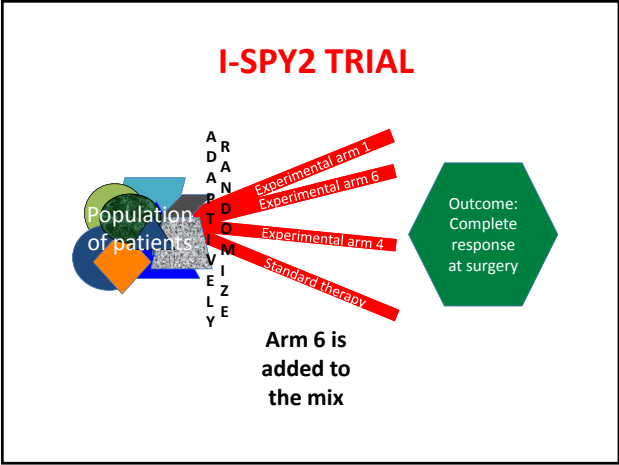


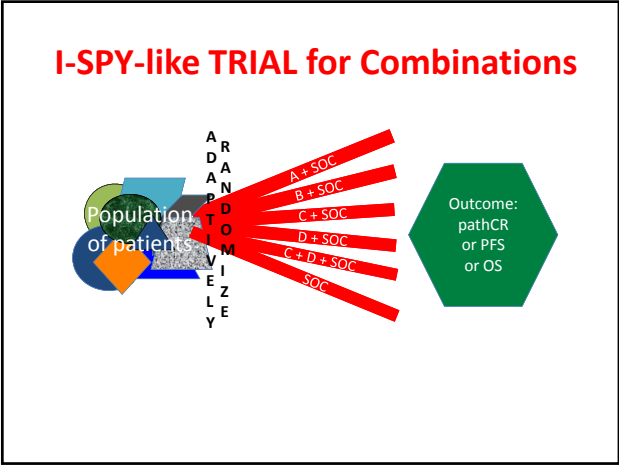


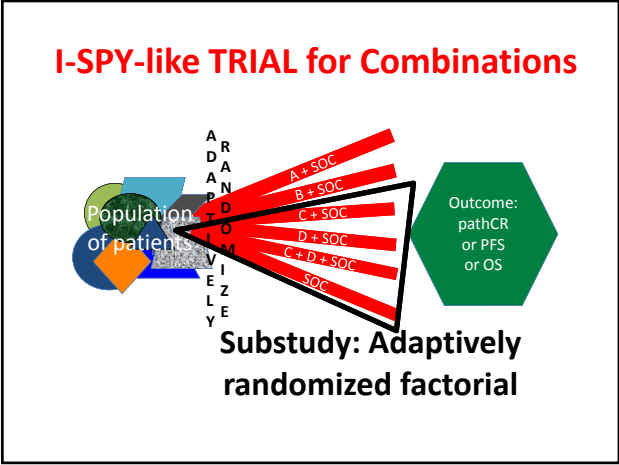




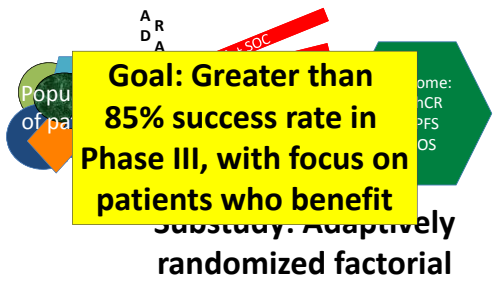








I-SPY-like TRIAL for Combinations



I-SPY2 Adaptive Process

- Primary endpoint: pCR (role of MRI?)
- Match drugs with up to 10 biomarker signatures
- n between 60 and 120 for “graduates”
- Currently:
 - ◆ 19 centers, US & Canada
 - ◆ ~500 pts randomized
 - ◆ First 7 exp drugs:
neratinib, ABT888, AMG386, AMG479, MK2206, pertuzumab, pertuzumab+T-DM1

I-SPY2 Effects & Clones

- 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
- Offspring of I-SPY 2: lymphoma, HIV, melanoma, Alzheimer’s, acute heart failure, scleroderma, SARI/H1N1, ...

**May 2012
Clinical/Medical**

- Introduction to Bayes adaptive
- BATTLE trial in lung cancer
- I-SPY 2, brief intro for Jane
- **Goldilocks and I-SPY 3**
- Basket trials
- Decision analysis & rare disease

45

Assumptions

➤ Focus on triple-negative neoadjuvant breast cancer as an example

➤ Hypothetical phase II results

- Control pCR rate 35%
- Experimental: pCR rates 35%, 40%, ..., 70%

➤ Benefit in EFS due to pCR increase

➤ Relationship between pCR and EFS from CTNeoBC (Cortazar et al. SABCs 2012)

46

San Antonio Breast Cancer Symposium – December 4 – 8, 2012

CTNeoBC

Association of pCR with EFS in Triple Negative Subtype

Triple Negative

HR=0.24, $P < 0.001$
(CI: 0.18 to 0.33)
— pCR (n = 389)
— no pCR (n = 768)

HR = 0.24

(Courtesy of Patricia Cortazar)

pCR=ypT0/is ypN0

* Nominal p-value

This presentation is the intellectual property of the author/presenter. Contact patricia.cortazar@utsa.hhs.gov for permission to reprint and/or distribute.

18

(Bayesian) Distributions of Hazards (CTNeoBC), Assuming Exponential Event Rate

If pCR

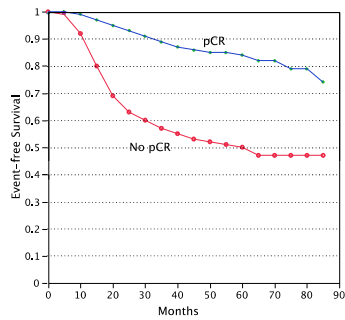
If not pCR

Hazard (per month)

48

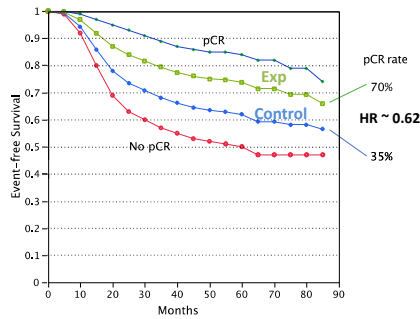
16

Smoothed Version of Cortazar



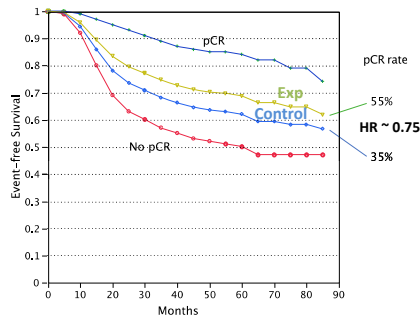
49

EFS for pCR Rates 70% vs 35%



50

EFS for pCR Rates 55% vs 35%



51

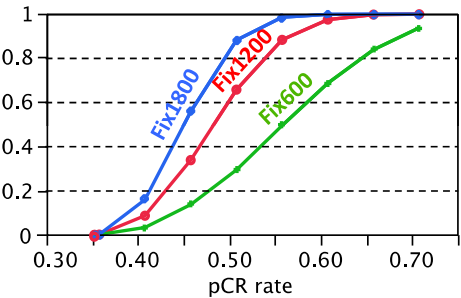
For All Designs Considered

- Accelerated approval if superiority on pCR
 - pCR analysis when all patients have surgery
 - Single final pCR comparison with control
- Full approval if superiority on EFS
 - 3 years minimum follow-up for EFS
 - Single final EFS comparison
 - Type I error rate controlled $\leq 2.5\%$

52

Power Via Simulations: Fixed Design

Same for Group-Sequential



53

One-Look (at pCR) Design

- Maximum sample size N (= 1200)
- When 300th patient has surgery, find predictive probabilities that both pCR and EFS stat sig based on pCR results (*only!*) from I-SPY 3 (assume CTNeoBC)
 - If PP < 5% when get to N stop now for futility
 - If PP > 90% with current n then stop accrual (final n = 300 + ~120 due to delayed surgery)
 - Otherwise choose smallest sample size (multiple of 100) having PP > 90% (or go to N if none)

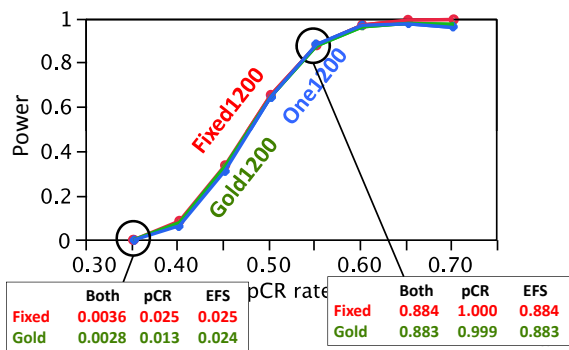
54

Goldilocks Design

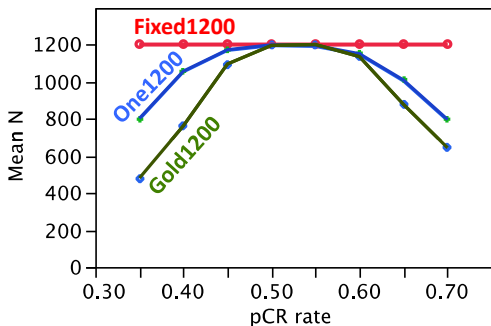
- Maximum sample size N (= 1200)
- When 300th pt has surgery, find PP of both pCR and EFS stat sig based on pCR data (only!) from I-SPY 3 (assume CTNeoBC)
 - If PP < 5% when N pts then stop for futility
 - If PP > 90% with current n then stop accrual (final n greater by ~120)
- Else continue to next 100 surgeries; repeat above until n = N
- In all cases, pCR analysis after 6 mos, EFS analysis after 3 yrs

55

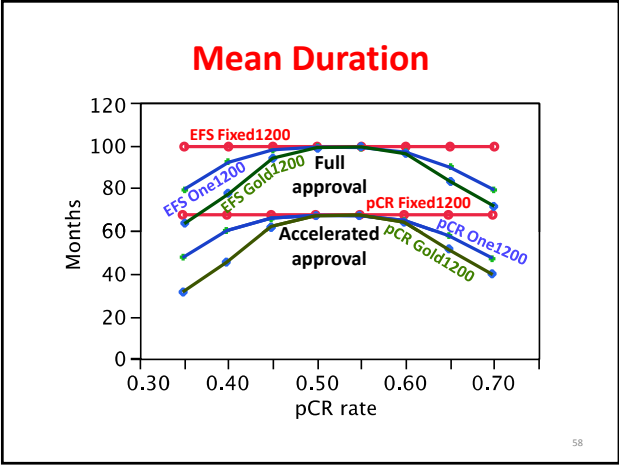
Goldilocks Vs Fixed Vs One IA



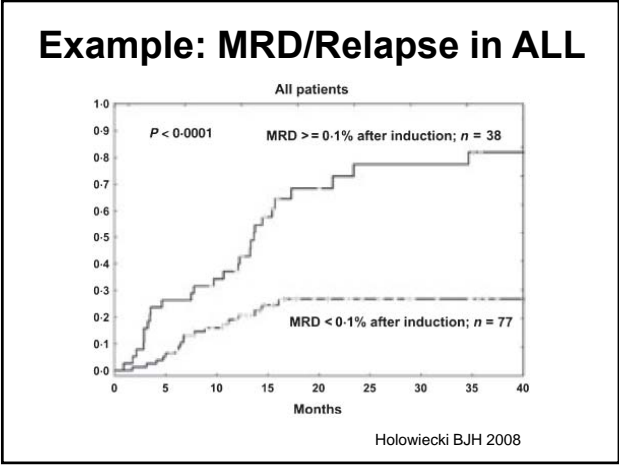
Mean Sample Size



57



**MRD a surrogate marker
in leukemia?**



OUTLINE

- Introduction to Bayes adaptive
- BATTLE trial in lung cancer
- I-SPY 2, brief intro for Jane
- Goldilocks and I-SPY 3
- **Basket trials**
- Decision analysis & rare disease

Basket Trials

- Targeted drug, develop simultaneously across organ-specific cancers
- Restrict to tumors expressing target
- Population sizes small means trial sample sizes must be small

The Approaching Wall

- Ever finer grid of biomarker categories: Within 10 years every cancer patient will have an orphan disease.
- How to develop drugs in this setting?

CLINICAL
TRIALS

ARTICLE

Clinical Trials 2013

Bayesian hierarchical modeling of patient subpopulations: Efficient designs of Phase II oncology clinical trials

Scott M Berry^a, Kristine R Broglio^a, Susan Groshen^b and Donald A Berry^{a,c}

Background In oncology, the treatment paradigm is shifting toward personalized medicine, where the goal is to match patients to the treatments most likely to deliver benefit. Treatment effects in various subpopulations may provide some information about treatment effects in other subpopulations.

Purpose We compare different approaches to Phase II trial design where a new treatment is being investigated in several groups of patients. We compare considering each group in an independent trial to a single trial with hierarchical modeling of the patient groups.

Methods We assume four patient groups with different background response rates and simulate operating characteristics of three trial designs, Simon's Optimal Two-Stage design, a Bayesian adaptive design with frequent interim analyses, and a Bayesian adaptive design with frequent interim analyses and hierarchical modeling across patient groups.

Results Simon's designs are based on 10% Type I and Type II error rates. The inde-

Hierarchical modeling/ Bayes borrowing assumptions

Population of response rates within 10 tumor types:

Response rates p_i have a distribution, one that is imperfectly known, even after observing the R_i/N_i

65

Hierarchical modeling/ Bayes borrowing assumptions

Population of response rates within 10 tumor types:

R_1/N_1 gives info about p_1 which gives info about population of p 's which gives info about p_2 , say. Hence "borrowing."

66

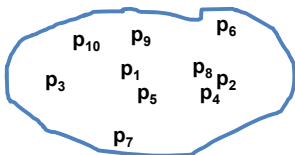
Hierarchical modeling
assumptions/prior

- Distribution of response rates p_i is unknown—itself has a probability distribution
- Expectations regarding p 's can differ by tumor type
- Prior distribution (“hyperprior”) of heterogeneity σ in population of p 's is important in determining borrowing

67

Learn about heterogeneity
parameter σ from trial results

σ large:



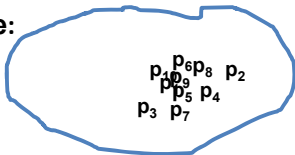
σ small:



68

Many Possibilities ...

σ moderate:



σ moderate:



69

Compare Bayes with S2S
(same type I and II error rates)

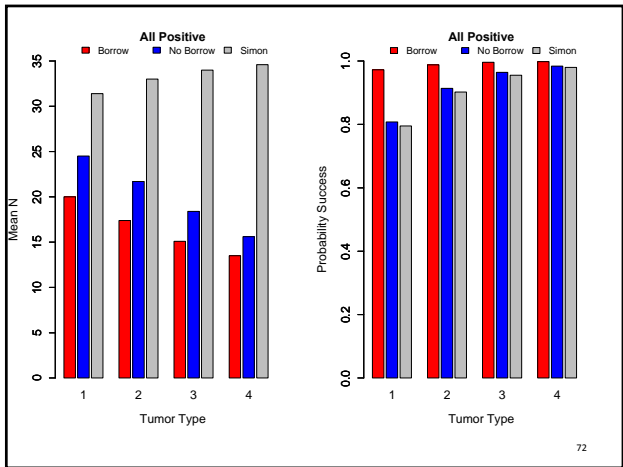
- Intermediate design: Bayes with no borrowing across tumor types
- Criteria:
 - ◆ Total sample size
 - ◆ Probability correct decision

70

One case: 4 tumor types, all positive

Type	p	Bayes Borrow			Bayes No Borrow			S2S		
		Mean N	Prob Success	Prob Futility	Mean N	Prob Success	Prob Futility	Mean N	Prob Success	Prob Futility
1	0.25	10.0	0.972	0.028	24.5	0.808	0.192	31.4	0.795	0.205
2	0.30	17.4	0.988	0.012	21.7	0.914	0.086	33.0	0.902	0.098
3	0.35	17.1	0.996	0.004	18.4	0.964	0.036	34.0	0.955	0.045
4	0.40	13.5	0.998	0.002	15.6	0.984	0.016	34.6	0.980	0.020
Total		66.0			80.2			133.0		

71



72

OUTLINE

- Introduction to Bayes adaptive
- BATTLE trial in lung cancer
- I-SPY 2, brief intro for Jane
- Goldilocks and I-SPY 3
- Basket trials
- Decision analysis & rare disease

Designing a clinical trial
is making a decision

74

Standard Approach to
Choosing Sample Size

- Example: time to event. Want
 - 25% reduction in hazard
 - 5% type I error, two-sided
 - 80% power
 - 2/month accrual
 - 8 mo median for control
 - 12 mos follow-up
- Answer: $n = 650$

75

Standard Approach

- Would be an accident if this design is optimal (or even okay) when goal is delivering good medicine
- Can't be right for both
 - Disease is CHF
 - Disease is a rare pediatric cancer
- Can't be right for both
 - Product is bone marrow transplant
 - Product is a chocolate bar

76

Of course investigators and regulators adjust

- For rare diseases, accept smaller trials
- For highly invasive or toxic products, consider disease severity & require stronger evidence of effectiveness
- Still: How small, how severe, how strong?
- To know whether deliver good medicine, evaluate impact on patient population

77

Decision-Analytic Clinical Trial

- Deliver good medicine to patients
- Which patients?

78

Ethics: Individual vs Collective

- Fundamental conflict
- Learning with time is inevitable in medicine: Better to delay getting any disease!
- But all patients should count equally *a priori*, sacrificing neither learning nor effective treatment of patients in the trial

79

Decision-analytic Clinical Trial

- Goal: Effective overall treatment of patients, both
 - Those in the trial and
 - Those who come after the trial
- Maximize overall benefit in “patient horizon” N: All patients with the disease who may benefit from therapies considered

80

- Goal: maximize expected number of successes in N
- Either one- or two-armed trial
- Suppose $n = 1000$ is right for one trial & $N = 1,000,000$
- Then for other N's use $n =$

Optimal sample size for one trial and first of two trials					
N	1,000,000	100,000	10,000	1000	100
One trial	1000	320	100	32	10
One/two	170	78	36	17	8

Ratio of sample sizes within row is general
Ratio across rows applies for particular prior distribution

Optimal allocations
in a two-armed trial:

82

Optimal allocations to Arms 1 and 2									
N	Prior distributions								
			Succ prop			Succ prop			Succ prop
100	6	5	.63	4	8	.71	9	0	.60
1000	21	20	.65	16	30	.74	29	0	.62
10,000	70	69	.66	56	98	.75	99	0	.62
Large N	$\frac{\sqrt{N}}{2}$	$\frac{\sqrt{N}}{2}$	$\frac{2}{3}$	$\frac{\sqrt{N}}{3}$	\sqrt{N}	$\frac{3}{4}$	\sqrt{N}	0	$\frac{5}{8}$

Of course, optimal sequential
assignments are even better:
Bandit strategies

84

“Bandit Strategies” & Google

85

Berry' s bandit stuff

- Berry DA (1972). A Bernoulli two-armed bandit. *The Annals of Mathematical Statistics* 43:871-897.
- Berry DA (1978). Modified two-armed bandit strategies for certain clinical trials. *Journal of the American Statistical Association* 73:339-346.
- Berry DA, Fristedt B (1979). Bernoulli one-armed bandits—arbitrary discount sequences. *The Annals of Statistics* 7:1086-1105.
- Berry DA, Fristedt B (1980). Two-armed bandits with a goal, I. One arm known. *Advances in Applied Probability* 12:775-798.
- Berry DA, Fristedt B (1980). Two-armed bandits with a goal, II. Dependent arms. *Advances in Applied Probability* 12:958-971.
- Berry DA, Viscont WK (1981). Bernoulli two-armed bandits with geometric termination. *Stochastic Processes and Their Applications* 11:35-45.
- Berry DA, Fristedt B (1983). Maximizing the length of a success run for many-armed bandits. *Stochastic Processes and Their Applications* 15:317-325.
- Berry DA (1983). Bandit Problems with Random Discounting. In *Mathematical Learning Models—Theory and Algorithms* 12-25. New York, Springer-Verlag. (Ed: Herkenrath U, Kalin D, Vogel W.)
- Berry DA (1985). One- and Two-armed Bandit Problems. *Encyclopedia of Statistical Sciences*, Vol. VI, 418-422. New York: John Wiley & Sons. (Ed: Kotz S, Johnson NL)
- Clayton MK, Berry DA (1985). Bayesian nonparametric bandits. *The Annals of Statistics* 13:1523-1534.
- Berry DA, Fristedt B (1985). *Bandit Problems: Sequential Allocation of Experiments*. 275 pp. London: Chapman-Hall.
- Berry DA, Kertz RP (1991). Worth of perfect information in Bernoulli bandits. *Advances in Applied Probability* 23:1-23.
- Berry DA, Eick SG (1995). Adaptive assignment versus balanced randomization in clinical trials: A decision analysis. *Statistics in Medicine* 14:221-246.
- Berry DA, Chen RW, Zame A, Heath DE, Shepp LA (1997). Bandit problems with infinitely many arms. *The Annals of Statistics* 25:2103-2116.
- Giles FJ, Kantarjian HM, Cortes JE, Garcia-Manero G, Verstovsek S, Faderl S, Thomas DA, Ferrajoli A, O'Brien S, Wathen JK, Xiao L-C, Berry DA, Estey EH (2003). Adaptive randomized study of idarubicin and cytarabine versus flouxacinib and cytarabine versus flouxacinib and idarubicin in untreated patients 50 years or older with adverse karyotype acute myeloid leukemia. *Journal of Clinical Oncology* 21:1722-1727.
- Cheng Y, Su F, Berry DA (2003). Choosing sample size for a clinical trial using decision analysis. *Biometrika* 90:923-936.
- Cheng Y, Berry DA (2007). Optimal adaptive randomized designs for clinical trials. *Biometrika* 94:673-689.
- Blawas S, Liu DD, Lee JJ, Berry DA (2009). Bayesian clinical trials at the University of Texas M. D. Anderson Cancer Center. *Clinical Trials* 6:205-216.

86
