Multi-armed Bayesian Bandits Change the Paradigm of Clinical Trials

Donald Berry
don@berryconsultants.net

THE UNIVERSITY OF TEXAS
MD Anderson Cancer Center
Making Cancer History®

Berry Consultants
Statistical Innovation

Outline

- **Bandits and me**
- **My efforts trying to inject bandit thinking into clinical research**
- **Recent successes: Multi-armed Bayesian Bandits Are Changing the Paradigm of Clinical Trials**

The Annals of Mathematical Statistics
1972, Vol. 43, No. 3, 871-897

A BERNOULLI TWO-ARMED BANDIT[1]

BY DONALD A. BERRY

University of Minnesota

One of two independent Bernoulli processes (arms) with unknown expectations ρ and λ is selected and observed at each of n stages. The selection problem is sequential in that the process which is selected at a particular stage is a function of the results of previous selections as well as of prior information about ρ and λ. The variables ρ and λ are assumed to be independent under the (prior) probability distribution. The objective is to maximize the expected number of successes from the n selections. Sufficient conditions for the optimality of selecting one or the other of the arms are given and illustrated for example distributions. The stay-on-a-winner rule is proved.

1. Introduction and statement of the problem. Let $\mathscr{R}$ and $\mathscr{L}$ denote independent Bernoulli processes with parameters—probabilities of success—ρ and λ respectively. Call $\mathscr{R}$ the *right arm* and $\mathscr{L}$ the *left arm*. An observation on either arm is called a *pull*. A right pull or a left pull is made at each of n stages and the result of the pull at each stage is known before a right or left pull is made at the next stage. The parameters ρ and λ associated with $\mathscr{R}$ and $\mathscr{L}$ are

Application to Clinical Trials?

Berry's bandits

Berry DA (1972). A Bernoulli two-armed bandit. *Ann Math Statist* 43:871-897.

Berry DA (1978). Modified two-armed bandit strategies for certain clinical trials. *J Amer Statist Assoc* 73:339-345.

Berry DA, Fristedt B (1979). Bernoulli one-armed bandits—arbitrary discount sequences. *Ann Statist* 7:1086-1105.

Berry DA, Fristedt B (1980). Two-armed bandits with a goal, I. One arm known. *Adv Appl Probab* 12:775-798.

Berry DA, Fristedt B (1980). Two-armed bandits with a goal, II. Dependent arms. *Adv Appl Probab* 12:958-971.

Berry DA, Viscusi WK (1981). Bernoulli two-armed bandits with geometric termination. *Stoch Proc Applic* 11:35-45.

Berry DA, Fristedt B (1983). Maximizing the length of a success run for many-armed bandits. *Stoch Proc Applic* 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. *Ann Statist* 13:1523-1534.

Berry DA, Fristedt B (1985). *Bandit Problems: Sequential Allocation of Experiments*. London: Chapman-Hall.

Berry DA, Kertz RP (1991). Worth of perfect information in Bernoulli bandits. *Adv Appl Probab* 23:1-23.

Berry DA, Eick SG (1995). Adaptive assignment versus balanced randomization in clinical trials: A decision analysis. *Statist Med* 14:231-246.

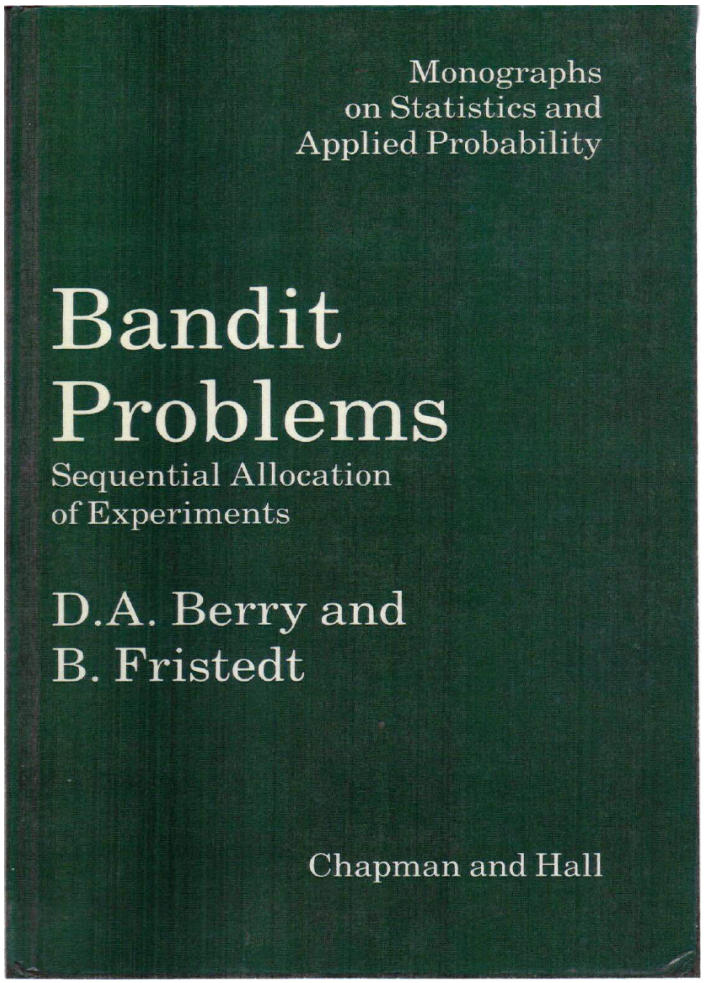
Berry DA, Chen RW, Zame A, Heath DE, Shepp LA (1997). Bandit problems with infinitely many arms. *Ann Statist* 25:2103-2116.

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.

Dissertations in Decision Making, esp. Bandits

1. John S. Croucher (1974)—Game Theoretic Models with Applications
2. Dorothee P. Aeppli (1980)—The Concept of Information and Ferguson's Distribution
3. Larry M. Pearson (1980)—Treatment Allocation for Clinical Trials in Stages
4. Roy F. Mensch (1981)—A Search Problem with Directional Information
5. Ronald R. Christensen (1983)—Searching for the Lowest Price when the Unknown Distribution of Prices is Modeled with a Dirichlet Process
6. Jeffrey A. Witmer (1983)—Bayesian Multistage Decision Problems
7. Murray K. Clayton (1983)—Bayes Sequential Sampling for Choosing the Better of Two Populations
8. Stephen G. Eick (1985)—Sequential Experimentation with Delayed Responses
9. Chih-Hsiang Ho (1986)—One-sided Sequential Stopping Boundaries for Clinical Trials: Classical and Bayesian Approaches
10. Steven N. MacEachern (1988)—Sequential Bayesian Bioassay Design
11. John S. Andersen (1988)—Allocating Experiments in Stages
11. Kumarasiri Samaranayake (1988)—Bernoulli k-Armed Bandits with Dependent Arms
12. Manas K. Chattopadhyay (1991)—Dirichlet Bandit Problems
13. Shipei Weng (1991)—Sequential Allocation to Maximize the Probability of Achieving a Number of Successes
14. Yi Cheng (1992)—Group Sequential Strategies in Two-Armed Bandit Problems
15. Kris Gillingham (1993)—Bayesian Hierarchical Models for Metaanalysis of Dichotomous Response Studies
16. Zhengning Lin (1993)—Statistical Methods for Combining Historical Controls with Clinical Trial Data
17. Ram Gopalan (1994)—Bayesian Multiple Comparisons
18. Chengchang Li (1994)—Metaanalysis of Survival Data
19. Jiang Qian (1994)—A Bayesian Weibull Survival Model
20. Fusheng Su (1996)—Limit Theorems on Deviation Probabilities with Applications to Two-Armed Clinical Trials
21. Lurdes Inoue (1999)—Bayesian Design and Analysis of Clinical Experiments
22. Heidi Ashih (2000)—Bayesian Models of Tumor Growth in Breast Cancer
23. Shu Han (2005)—Modeling Auxiliary Information in Clinical Trials
24. Lin Yang (2011)—Efficient Phase I Cancer Trials


Monographs
on Statistics and
Applied Probability
Bandit
Problems
Sequential Allocation
of Experiments
D.A. Berry and
B. Fristedt
Chapman and Hall

1985

Modified Two-Armed Bandit Strategies for Certain Clinical Trials

DONALD A. BERRY*

A procedure which maximizes the expected number of successes in a clinical trial involving two treatments can usually be found only by backward induction. Not only is it difficult to find an optimal procedure but, once found, it is difficult to describe and cumbersome to communicate. A procedure is suggested which depends on the information present concerning the treatments. This procedure is easy to calculate and approximates an optimal procedure quite well. It is applicable to trials for which the number of patients is unknown as well as those of known duration.

KEY WORDS: Clinical trials; Two-armed bandits; Sequential Bayesian decisions; Feldman's strategy.

1. INTRODUCTION

Suppose that two treatments are available for use in a clinical trial. Further suppose that the response to treatment is either positive (a success) or negative (a failure) and that the patients can be treated one at a time, with each patient's response available before the next patient is to be treated. The number of patients in the trial is N as a function of (p_1, p_2). Another, the Bayesian approach, asks that current information concerning (p_1, p_2) be quantified in the form of a probability distribution. The Bayesian approach will be used in most of this article. A distinct advantage of this approach is that accumulating information can be handled in a unified way: Bayes' theorem is used to modify the probability distribution on (p_1, p_2). The effectiveness of a procedure can then be averaged over (p_1, p_2) and, possibly, a procedure can be found that maximizes the expected number of successes.

If one of the p_i is known, say p_1, and p_2 has probability measure ν, then the procedure selection problem is called a "one-armed bandit" (*cf.* Bradt, Karlin, and Johnson 1956). An intuitively appealing characteristic of optimal procedures for the one-armed bandit is that the initial patients constitute an information-gathering stage, which may be empty or may exhaust the trial, during which

© Journal of the American Statistical Association
June 1978, Volume 73, Number 362
Theory and Methods Section

Advice from biostatisticians

- **"Berry, you don't understand clinical trials."**
- **"You're wet behind the ears. You should learn about real clinical trials."**
- **"Prior distributions are biased."**
- **"Nonrandomization is a non-starter."**

I took the advice

- **I came to understand clinical trials**
- **I learned how to design stupid clinical trials … and I designed some! For U.S. National Cancer Institute's cooperative groups.**

If you can’t beat ‘em, join ‘em!

For example …

ORIGINAL CONTRIBUTION

Estrogen-Receptor Status and Outcomes of Modern Chemotherapy for Patients With Node-Positive Breast Cancer

Donald A. Berry, PhD
Constance Cirrincione, MS
I. Craig Henderson, MD
Marc L. Citron, MD
Daniel R. Budman, MD
Lori J. Goldstein, MD
Silvana Martino, DO
Edith A. Perez, MD
Hyman B. Muss, MD
Larry Norton, MD
Clifford Hudis, MD
Eric P. Winer, MD

Context Breast cancer estrogen-receptor (ER) status is useful in predicting benefit from endocrine therapy. It may also help predict which patients benefit from advances in adjuvant chemotherapy.

Objective To compare differences in benefits from adjuvant chemotherapy achieved by patients with ER-negative vs ER-positive tumors.

Design, Setting, and Patients Trial data from the Cancer and Leukemia Group B and US Breast Cancer Intergroup analyzed; patient outcomes by ER status compared using hazards over time and multivariate models. Randomized trials comparing (1): 3 regimens of cyclophosphamide, doxorubicin, and fluorouracil (January 1985 to April 1991); (2) 3 doses of doxorubicin concurrent with cyclophosphamide, with or without subsequent paclitaxel (May 1994 to April 1997); (3) sequential doxorubicin, paclitaxel, and cyclophosphamide with concurrent doxorubicin and cyclophosphamide followed by paclitaxel, and also 3-week vs 2-week cycles (September 1997 to March 1999). A total of 6644 node-positive breast cancer patients received adjuvant treatment.

Main Outcome Measures Disease-free and overall survival.

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Effect of Screening and Adjuvant Therapy on Mortality from Breast Cancer

Donald A. Berry, Ph.D., Kathleen A. Cronin, Ph.D., Sylvia K. Plevritis, Ph.D., Dennis G. Fryback, Ph.D., Lauren Clarke, M.S., Marvin Zelen, Ph.D., Jeanne S. Mandelblatt, Ph.D., Andrei Y. Yakovlev, Ph.D., J. Dik F. Habbema, Ph.D., and Eric J. Feuer, Ph.D., for the Cancer Intervention and Surveillance Modeling Network (CISNET) Collaborators*

ABSTRACT

BACKGROUND

We used modeling techniques to assess the relative and absolute contributions of screening mammography and adjuvant treatment to the reduction in breast-cancer mortality in the United States from 1975 to 2000.

THOMSON REUTERS

HIGHLY CITED RESEARCHER

PRESENTED TO

Donald Berry

In recognition of ranking among the top 1% of researchers
for most cited documents, in their specific field.

Clinical Medicine!

THOMSON REUTERS™

Signed

Basil Moftah, President IP & Science, Thomson Reuters

2196

NEWS

The New Math of Clinical Trials

Other fields have adopted statistical methods that integrate previous experience, but the stakes ratchet up when it comes to medical research

HOUSTON, TEXAS—If statistics were a religion, Donald Berry would be among its most dogged proselytizers. Head of biostatistics at the M. D. Anderson Cancer Center here, he's dropped all hobbies except reading bridge columns in the newspaper. He sends out e-mail missives at 3:00 in the morning. The running joke in the department is that Berry, his curly gray hair perpetually tousled, never sleeps. Admittedly, sleep doesn't come easily to a man on a mission.

Berry, 63, adheres to a branch of statistics named after an 18th century minister, Thomas Bayes, whose followers advocate incorporating prior knowledge into experiments and sometimes altering them as they run to take into account accumulating results. Although Bayesian designs are now widely used in everything from astrophysics to ecology (*Science*, 19 November 1999, p. 1460), they've been slower to catch on in medical research, particularly clinical trials. That's where Berry comes in.

A Bayesian since the 1960s, Berry for years was unable to implement his unorthodox approach. Then, in 1999, he was offered a golden opportunity: Come to M. D. Anderson, one of the largest cancer centers in the United States, with a reputation for being the "Wild West" of oncology research, and transform how it designs and runs many of the 800 clinical trials being conducted at any given time.

Berry's perch at Anderson has fueled his resolve to spread the Bayesian word. He crisscrosses the country speaking with cancer advocates, drug companies, and the Food and Drug Administration (FDA); the latter is beginning to consider Bayesian trials in new drug applications and is planning a May meeting on the subject.

His critics, however, hope his ideas won't take hold. Berry's skepticism that mammograms help younger women left him accused of risking lives; his approach to clinical trials has prompted worries about bad drugs slipping through the system. Bayesian drug studies risk "saying [a treatment is] positive too often," says biostatistician Stephanie Green of the Fred Hutchinson Cancer Research Center in Seattle, Washington. But critics and supporters alike have a grudging admiration for Berry's persistence. "He isn't swayed by the status quo, by people in power in his field," says Fran Visco, head of the National Breast Cancer Coalition in Washington, D.C. "You have to respect him for that," she adds, "whether you agree with him or not."

Maverick beginnings

Berry stumbled into statistics after an erratic college career. He skipped classes regularly and took a 3-year break, in 1960, to volunteer for the army. By his senior year, he and his wife had four sons (two more children, both girls, would follow), and Berry had little idea what to do with his life. A professor suggested statistics; Berry took the advice and enrolled in graduate school at Yale University. After completing his dissertation in 1971, he moved to the University of Minnesota.

From the start, Berry was drawn to the Bayesian school of thought, then widely viewed as an oddity within the field. The Bayesian approach calls for incorporating "priors"—knowledge gained from previous work—into a new experiment. "The Bayesian notion is one of synthesis … [and] learning as you go," says Berry. He found these qualities immensely appealing, in part because they reflect real-life behavior, including the way doctors practice medicine.

Bucking tradition. Donald Berry's support for Bayesian designs is changing the face of clinical trials, especially at his home base of M. D. Anderson Cancer Center.

CREDIT: BARRY G. SMITH

But learning as you go collides with the decades-old clinical trials paradigm. To guard against bias—from doctors, drug companies, and even patients—each phase of a traditional clinical trial is run from start to finish without interference from interested parties. Outside scientists monitor the data regularly; although a trial can be stopped early if patients appear unduly harmed or helped by the new treatment, the protocol itself can't normally be changed.

A Bayesian approach demands more than sporadic monitoring-board meetings, however. Bayesian trials often unveil data while a study is ongoing. What's more, researchers can use these early results to reallocate patients to different treatment groups depending on how the first batch of patients, or even a single patient, fares. Berry also favors other approaches foreign to clinical trials, such as answering questions about multiple drugs in a single experiment, a method known as factorial design. Factorial designs include a treatment arm for every drug combination possible, leading to unwieldy experiments whose results can be tough to interpret.

Some doctors agree with Berry that the standard approach to clinical trials is problematic. Elihu Estey, who oversees the treatment of acute leukemias at Anderson, points out that the typical paradigm assigns patients to different study arms with equal probability, even in the face of mounting evidence that one arm offers a better shot at survival. "The patients themselves, if they knew the way the trials are conducted, wouldn't be too thrilled," he says.

A big break for Berry came in 1990, when he was invited to join Cancer and Leukemia Group B (CALGB). It's one of the country's 10 cooperative groups: multi-institutional collaborations on large-scale cancer clinical trials. Berry would be the lead statistician for CALGB's breast cancer studies. He was not greeted warmly.

"I objected rather strenuously," recalls I. Craig Henderson, a breast oncologist at the University of California, San Francisco, who had heard that Bayesians were "loosey-

CLINICAL TRIALS *ARTICLE* *Clinical Trials* 2009; **6**: 205–216

Bayesian clinical trials at the University of Texas M. D. Anderson Cancer Center

Swati Biswas[a], *Diane D Liu*[b], *J Jack Lee*[b] *and Donald A Berry*[b]

Background The Bayesian approach is being used increasingly in medical research. In particular, it has become a standard in designing clinical trials at the University of Texas M. D. Anderson Cancer Center.

CLINICAL TRIALS *EDITORIAL* *Clinical Trials* 2009; **6**: 203–204

Bayesian clinical trials: no more excuses

it is possible for a biostatistics group in an academic cancer center to apply Bayesian methods on a broad scale and have them accepted by participating investigators, patients, sponsors, and regulatory bodies.

Biswas *et al*. [em for Bayesians. The their abstract ve randomization and approach is being increasingly used in medical research.' What follows is a recounting of MD Anderson's extensive role in that phenomenon. the third case is a Phase I/II study. These case studies give us a feel for what it takes to implement these methods into clinical research practice.

CLINICAL TRIALS ARTICLE

Clinical Trials 2009; **6**: 205–216

Bayesian clinical trials at the University of Texas M. D. Anderson Cancer Center

Swati Biswas[a], *Diane D Liu*[b], *J Jack Lee*[b] *and Donald A Berry*[b]

Background The Bayesian approach is being used increasingly in medical research. In particular, it has become a standard in designing clinical trials at the University of Texas M. D. Anderson Cancer Center.

CLINICAL TRIALS EDITORIAL

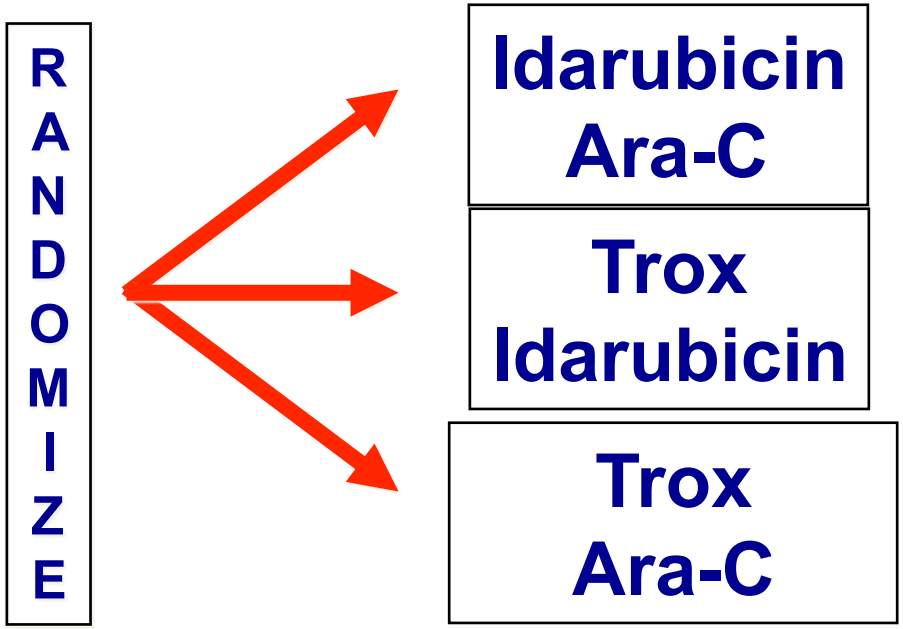
Clinical Trials 2009; **6**: 203–204

Bayesian clinical trials: no more excuses

While there are certainly some at other centers, the bulk of applied Bayesian clinical trial design in this country is largely confined to a single zip code. Why is this the case?

Biswas (for Bayesians. The their ab randomization and approach is being increasingly used in medical research.' What follows is a recounting of MD Anderson's extensive role in that phenomenon. the third case is a Phase I/II study. These case studies give us a feel for what it takes to implement these methods into clinical research practice.

Example: Troxacitabine in AML*
(Endpoint: Complete remission by day 50)



Standard design

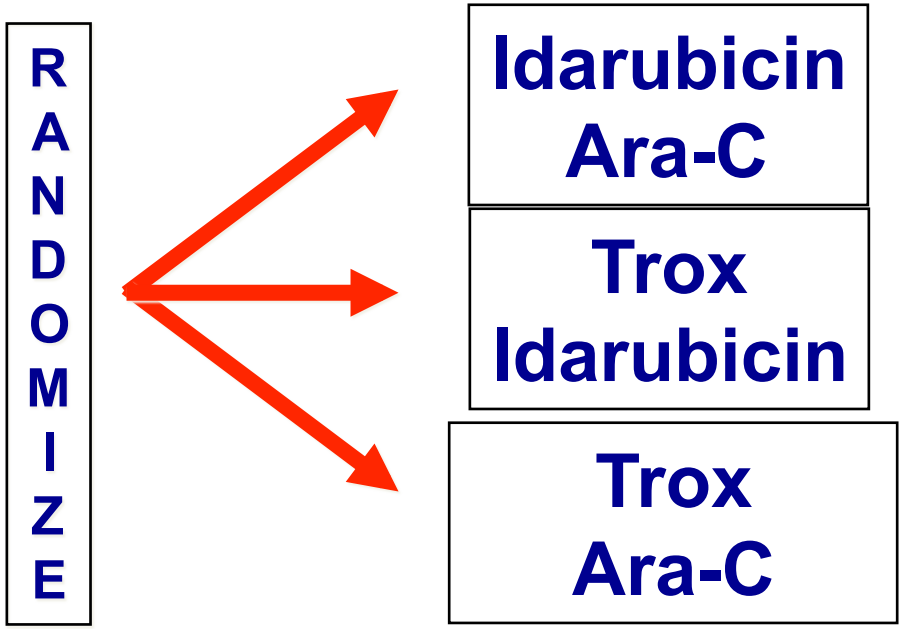
n = 25

n = 25

n = 25

*** Giles J Clin Oncol 2003**

Example: Troxacitabine in AML* (Endpoint: Complete remission day 50)



Our design

Adaptive randomization to learn, while effectively treating patients in trial

* Giles J Clin Oncol 2003

Adaptive Randomization

- **Assign better performing therapies with higher probability (modification of Thompson 1933)**
- **TI dropped after 24th patient**
- **Trial stopped after 34 patients**

Summary of AML trial results

Complete remission rate:

IA 10/18 = 56%

TA 3/11 = 27%

TI 0/5 = 0%

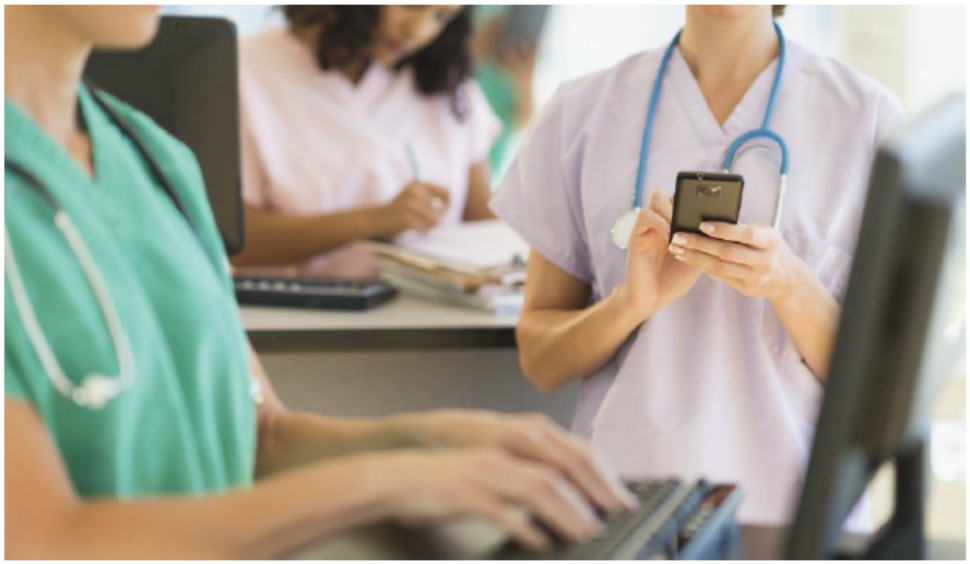
Cure magazine (2006)

“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.”

FORTUNE INSIDER CLINICAL TRIALS

How to take clinical research to the next level

COMMENTARY by Donald Berry OCTOBER 26, 2015, 3:58 PM EDT



Randomized clinical trials have changed little in 70 years, and it's time to revamp the approach by merging clinical research with clinical practice.

Current use of Bayesian adaptive designs

- MDACC (> 500 trials)
- Device companies (> 50 PMAs)*
- Drug companies (all major companies)**
- Berry Consultants (> 500 trials)

*http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071121.pdf

**http://www.fda.gov/downloads/DrugsGuidanceComplianceRegulatoryInformation/Guidances/UCM201790.pdf

Platform Trial

The NEW ENGLAND JOURNAL of MEDICINE

REVIEW ARTICLE

We illustrate the concept using the Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and Molecular Analysis 2 (I-SPY 2) …

Ph.D.,

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.

HIGH-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 ef-

Platform Trials Running or Under Development

- Many cancers
- Alzheimer's
- Ebola
- Antibacterials
- Community acquired pneumonia
- Pandemic flu

Especially in Europe!

NATURE MEDICINE | NEWS

Multicompany trials adapt to disciplines beyond cancer

11

14

When the I-SPY 2 trial launched in 2010, oncologists heralded it as the future of cancer research. ~~Five~~ pharmaceutical companies put aside their differences to participate in the landmark phase 2 breast cancer trial, which adaptively and efficiently randomized patients to one of ~~seven~~ experimental therapies. Now, even as I-SPY 2 propels its first two drugs into phase 3 trials, researchers in other areas of medicine are catching on to the benefits of this collaborative approach. On 11 December, Europe's Innovative Medicines Initiative (IMI) announced a €53 million call for proposals for a similarly designed trial in Alzheimer's disease. Already, at least 12 drug companies are keen to participate.

Who are EPAD?


Clinical research organization
QUINTILES
Patient organization
Alzheimer Europe
Small-medium enterprises
SYNAPSE
RESEARCH MANAGEMENT PARTNERS
Berry Consultants
IXICO
aracLon BIOTECH
EPAD
European Prevention of Alzheimer's Dementia Consortium
Academics
UNIVERSITY OF CAMBRIDGE
THE UNIVERSITY of EDINBURGH
VUmc
Hôpitaux de Toulouse
Karolinska Institutet
Fraunhofer
HÔPITAUX UNIVERSITAIRES PITIÉ SALPÊTRIÈRE - CHARLES FOIX
UNIVERSITÉ DE GENÈVE
University of Leicester
pasqual maragall foundation
MRC Medical Research Council
CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD
UNIVERSITY OF OXFORD
ErasmusMC
Inserm
Radboud Universiteit
EFPIA companies
Lilly
AMGEN
Takeda
biogen idec
Eisai
Pfizer
astellas
Boehringer Ingelheim
NOVARTIS
ucb
SANOFI
janssen
AC IMMUNE

Prototype Bayesian "Platform" Trial: I-SPY 2

http://www.ispy2.org

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

THE WALL STREET JOURNAL.
WSJ.com

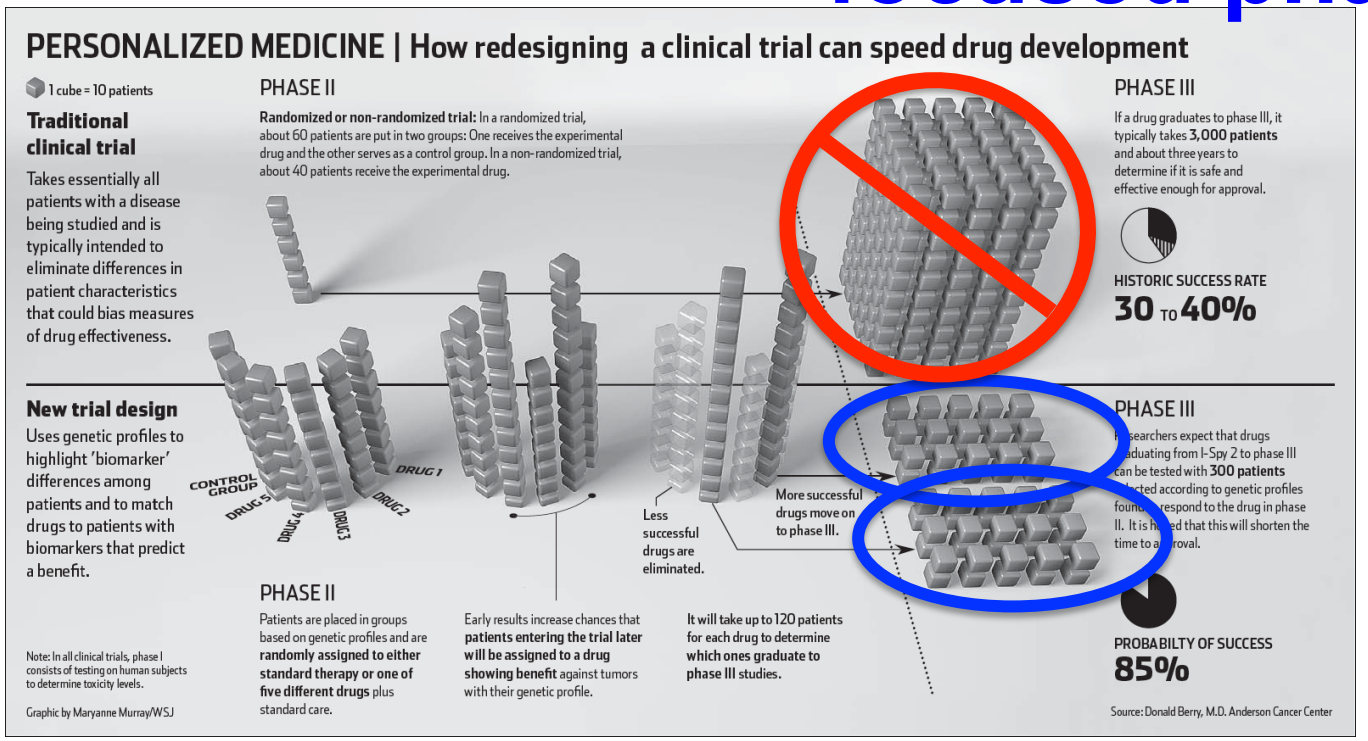
THE SATURDAY ESSAY | OCTOBER 2, 2010

A New Rx for Medicine

Goal: smaller, focused phase 3

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

By RON WINSLOW



The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 JULY 7, 2016 VOL. 375 NO. 1

Adaptive Randomization of Neratinib in Early Breast Cancer

J.W. Park, M.C. Liu, D. Yee, C. Yau, L.J. van 't Veer, W.F. Symmans, M. Paoloni, J. Perlmutter, N.M. Hylton, M. Hogarth, A. DeMichele, M.B. Buxton, A.J. Chien, A.M. Wallace, J.C. Boughey, T.C. Haddad, S.Y. Chui, K.A. Kemmer, H.G. Kaplan, C. Isaacs, R. Nanda, D. Tripathy, K.S. Albain, K.K. Edmiston, A.D. Elias, D.W. Northfelt, L. Pusztai, S.L. Moulder, J.E. Lang, R.K. Viscusi, D.M. Euhus, B.B. Haley, Q.J. Khan, W.C. Wood, M. Melisko, R. Schwab, T. Helsten, J. Lyandres, S.E. Davis, G.L. Hirst, A. Sanil, L.J. Esserman, and D.A. Berry, for the I-SPY 2 Investigators*

EDITORIAL

I-SPY 2 — Toward More Rapid Progress in Breast Cancer Treatment

Lisa A. Carey, M.D., and Eric P. Winer, M.D.

Clinical trials of systemic therapy for operable with human epidermal growth factor rec

ORIGINAL ARTICLE

Adaptive Randomization of Veliparib–Carboplatin Treatment in Breast Cancer

H.S. Rugo, O.I. Olopade, A. DeMichele, C. Yau, L.J. van 't Veer, M.B. Buxtor
M. Hogarth, N.M. Hylton, M. Paoloni, J. Perlmutter, W.F. Symmans, D. Yee
A.J. Chien, A.M. Wallace, H.G. Kaplan, J.C. Boughey, T.C. Haddad, K.S. Albai
M.C. Liu, C. Isaacs, Q.J. Khan, J.E. Lang, R.K. Viscusi, L. Pusztai, S.L. Moulde
S.Y. Chui, K.A. Kemmer, A.D. Elias, K.K. Edmiston, D.M. Euhus, B.B. Haley
R. Nanda, D.W. Northfelt, D. Tripathy, W.C. Wood, C. Ewing, R. Schwab, J. Lyandı
S.E. Davis, G.L. Hirst, A. Sanil, D.A. Berry, and L.J. Esserman, for the I-SPY 2 Investigato

ABSTRACT

PERSPECTIVE

STATISTICS IN MEDICINE

I-SPY 2 — A Glimpse of the Future of Phase 2 Drug Development?

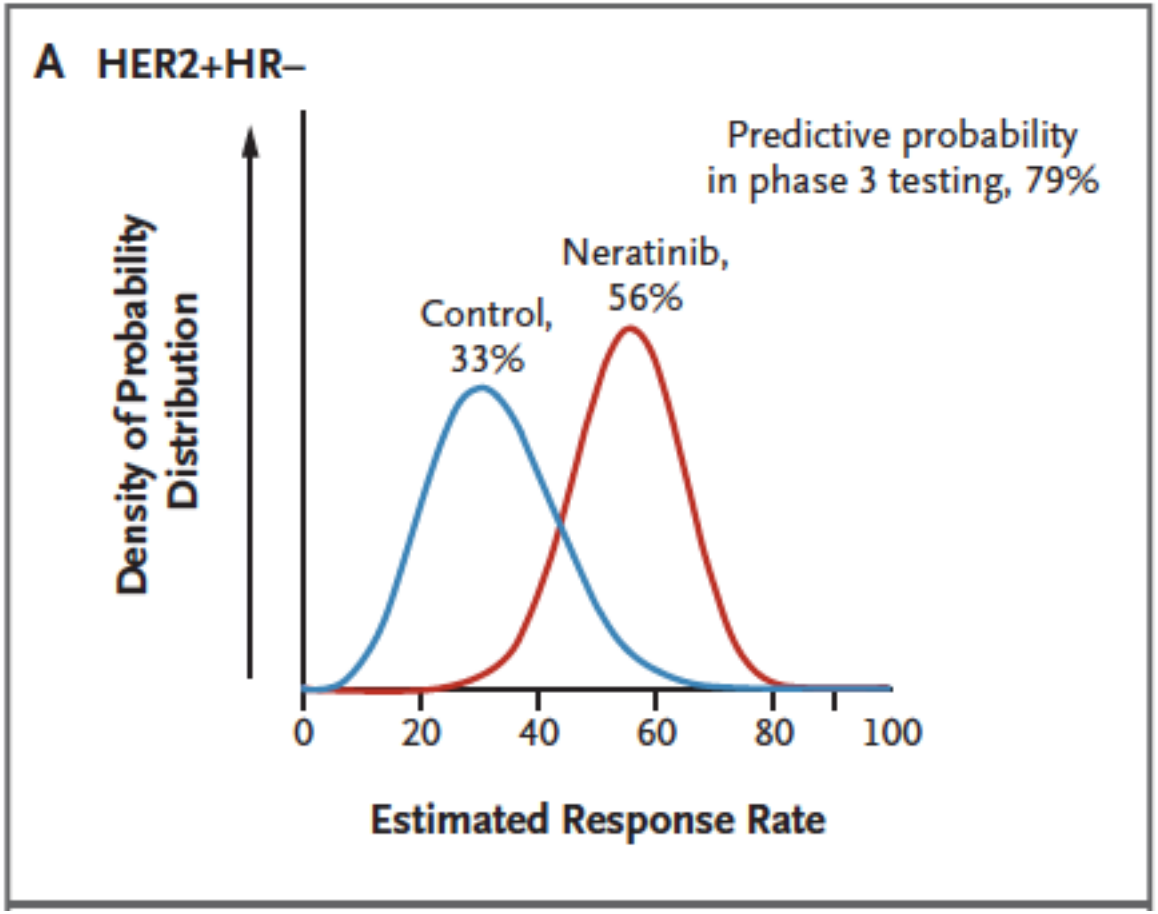
David Harrington, Ph.D., and Giovanni Parmigiani, Ph.D.

The articles by Rugo et al. (pages 23–34) and Park et al. (pages 11–22) in this issue of the *Journal* report results from the I-SPY (Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and Molecular Analysis) 2 platform, a promising adaptive strategy for matching targeted therapies for breast cancer with the patients

ing in larger, phase 3 trials. The value of I-SPY 2, however, may well go beyond the clinical results described in the current articles. Adaptive multigroup trials such as I-SPY 2 have the potential to answer several questions simultaneously and more efficiently than traditionally designed trials. Which of several promising therapies appear best suited for

good responses and, portant, may be usef ing patients to avoid when meaningful be likely. The challenge in identifying success therapies in cancer a tial. Targeted therapi to hit their target, th have the predicted e they do, and they ma

Neratinib's "graduation signature"



Ten biomarker signatures

- **Graduate drugs/signatures from trial:**
  - **Based on effectiveness**
  - **Based on prevalence**
- **Biomarker signatures (2^8 combinations of subtypes): B_1, B_2, …, B_{256}**

	MP-		MP+	
	HR+	HR-	HR+	HR-
HER2+	16%	7%	4%	10%
HER2-	23%	6%	6%	28%

Randomization to neratinib partway through its tour

Neratinib's eventual signature

	MP–		MP+	
	HR+	HR–	HR+	HR–
HER2+	+	++	+	++
HER2–	0	0	+	+

- Would be a nightmare in a 2-armed trial
- Easy in a platform trial: no amendment, only DSMB knew

I-SPY2 Adaptive Process

- Neoadjuvant breast cancer; PIs Esserman/Berry
- Primary endpoint: pCR (Longitudinal model of MRI volume)
- 10 biomarker signatures
- Adapti ing screening process
- Operat ation
- First sponsor: FNIH (NCI, FDA, industry)
- Coordinated wi CDRH)—
  Regulatory pathway via pCR
- Current status: 1 s, 3 0 pts randomized, first 14
  exp drugs: neratinib, veliparib, AMG AMG479,
  MK2206, pertuzumab, pertuzumab+T-DM1, ganetespib,
  pembrolizumab, P 397, talazoparib, patritumab, plus …

With different biomarker signatures and sample sizes

Graduated to phase 3

93

115

72

52

44

69

Adaptive Platform Effects

- Match drugs with biomarker signatures
- Savings from common control
- Better therapies move thru faster
- Drug/biomarker pairs graduate to small, focused, more successful Phase 3 based on Bayesian predictive probabilities

Another example of benefits of having many arms: Dropping control arm, and the “time machine”

JASA 1999 pp 661-685

Bridging Different Eras in Sports

Scott M. BERRY, C. Shane REESE, and Patrick D. LARKEY

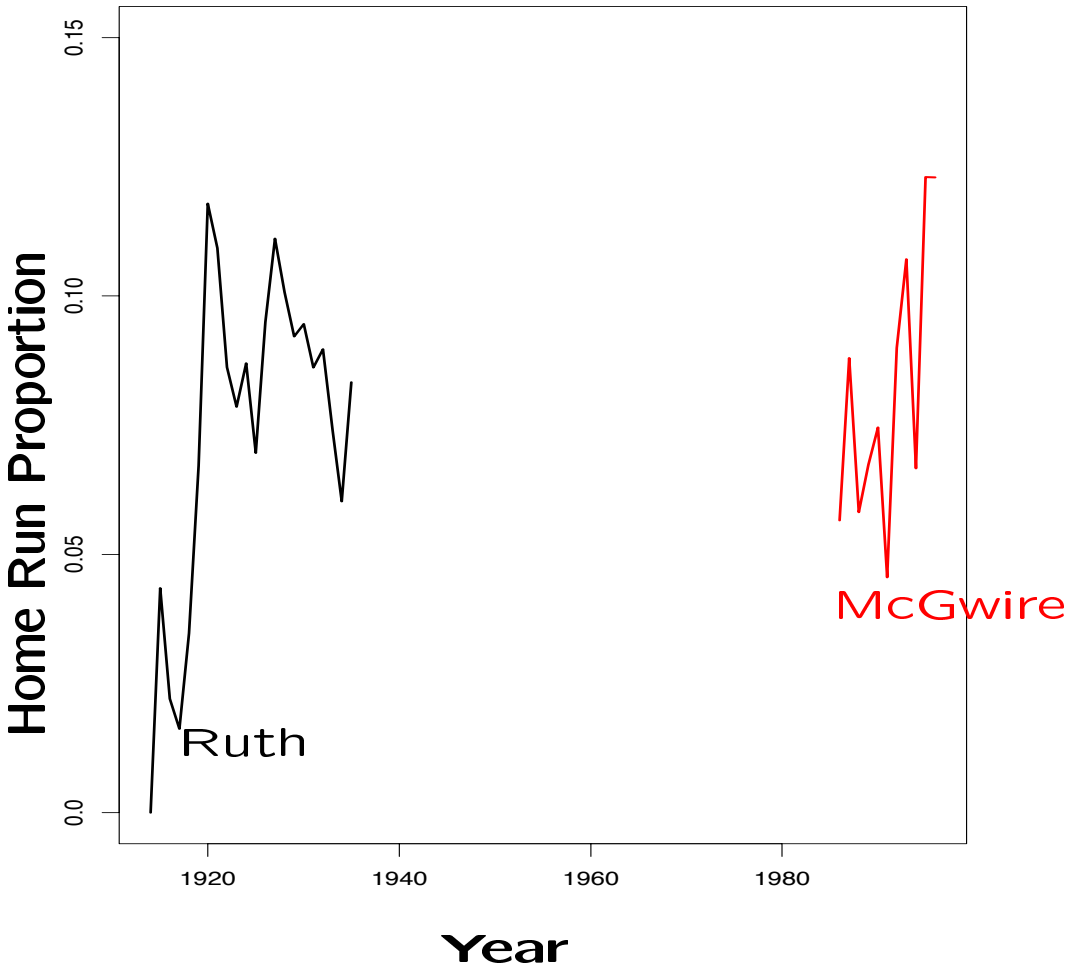
This article addresses the problem of comparing abilities of players from different eras in professional sports. We study National Hockey League players, professional golfers, and Major League Baseball players from the perspectives of home run hitting and hitting for average. Within each sport, the careers of the players overlap to some extent. This network of overlaps, or bridges, is used to compare players whose careers took place in different eras. The goal is not to judge players relative to their contemporaries, but rather to compare all players directly. Hence the model that we use is a statistical time machine. We use additive models to estimate the innate ability of players, the effects of aging on performance, and the relative difficulty of each year within a sport. We measure each of these effects separated from the others. We use hierarchical models to model the distribution of players and specify separate distributions for each decade, thus allowing the "talent pool" within each sport to change. We study the changing talent pool in each sport and address Gould's conjecture about the way in which populations change. Nonparametric aging functions allow us to estimate the league-wide average aging function. Hierarchical random curves allow for individuals to age differently from the average of athletes in that sport. We characterize players by their career profile rather than a one-number summary of their career.

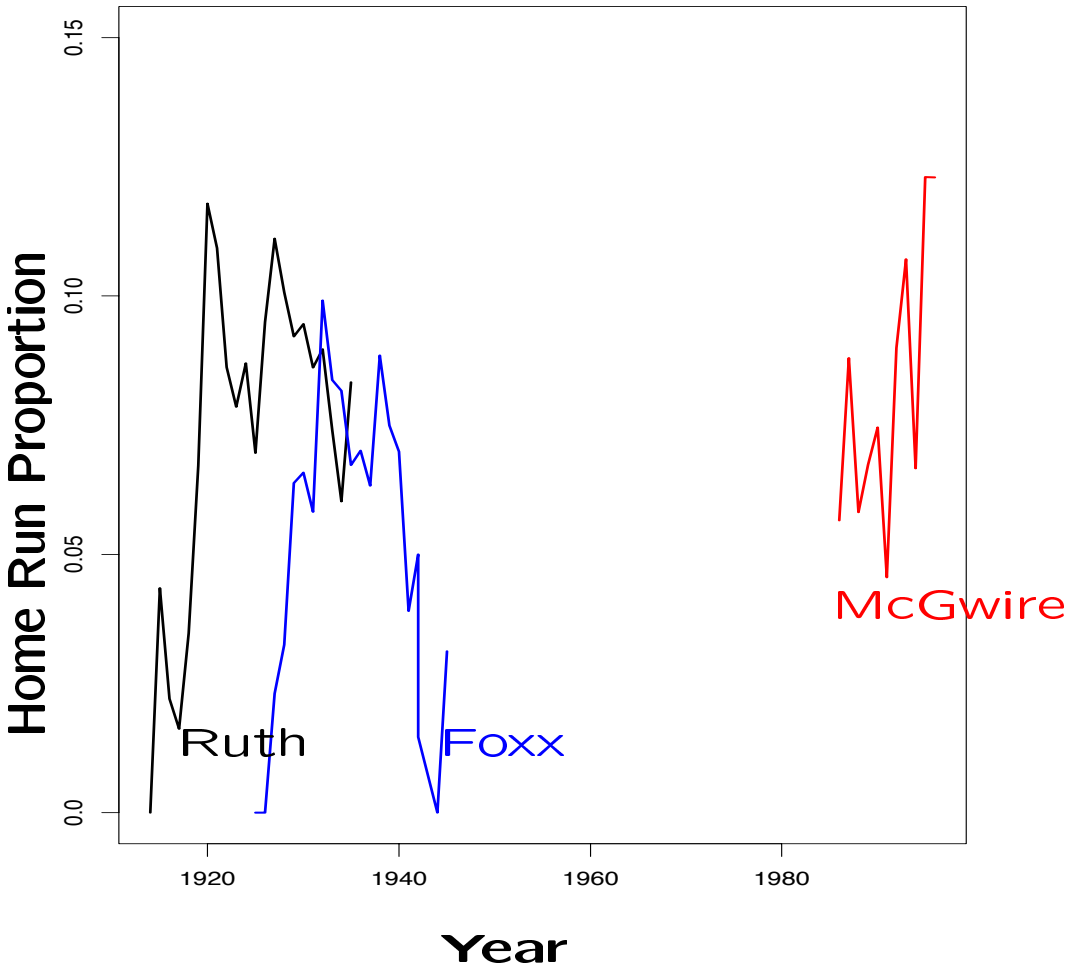
KEY WORDS: Aging function; Bridge model; Hierarchical model; Population dynamics; Random curve.

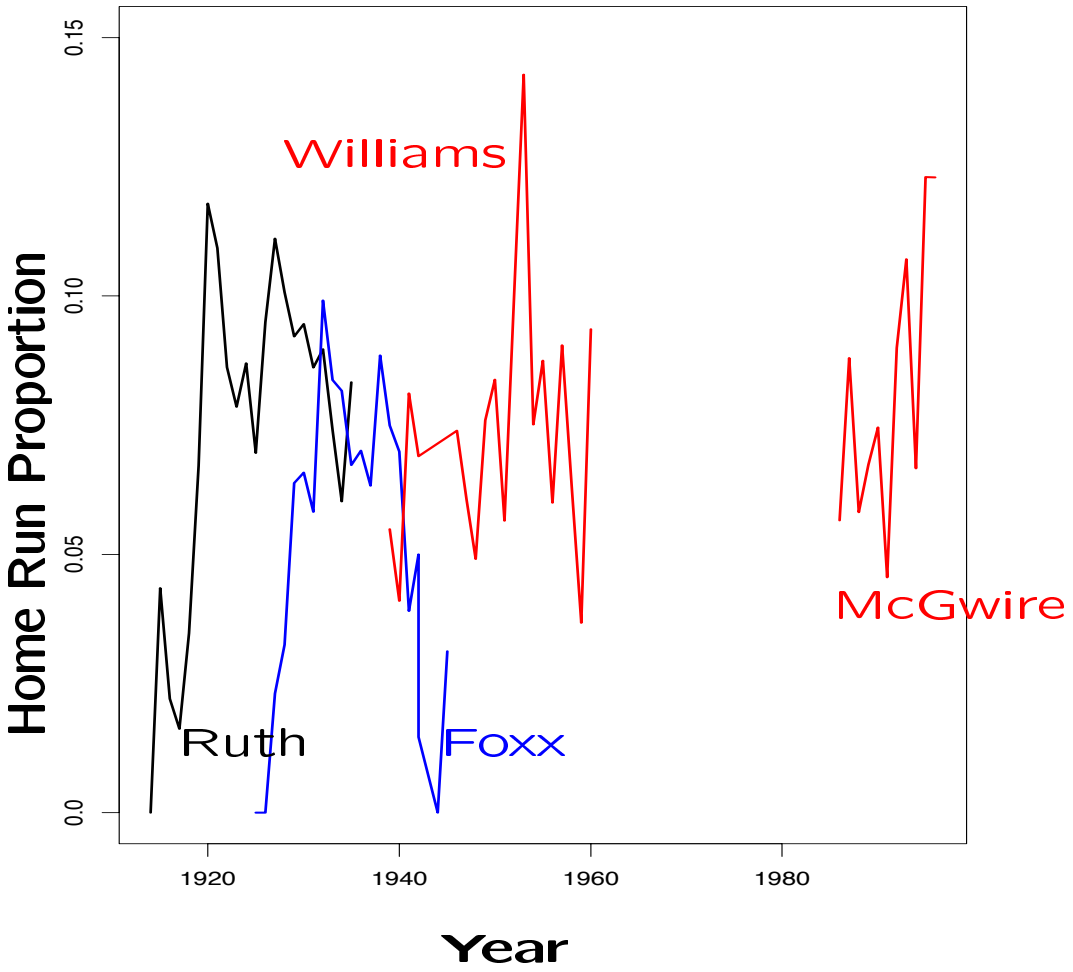
1. INTRODUCTION

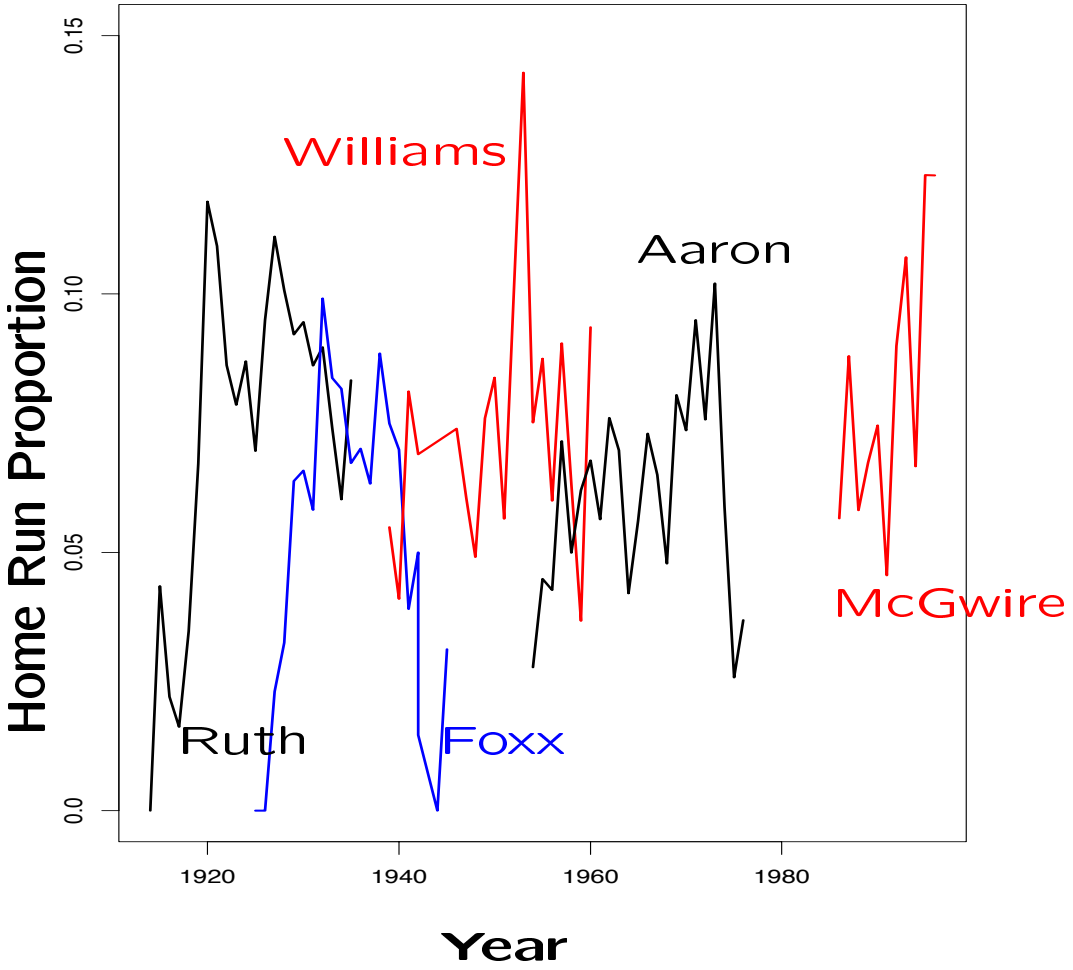
This article compares the performances of athletes from different eras in three sports: baseball, hockey, and golf. A goal is to construct a statistical time machine in which we estimate how an athlete from one era would perform in another era. For examples, we estimate how many home runs Babe Ruth would hit in modern baseball, how many points Wayne Gretzky would have scored in the tight-checking National Hockey League (NHL) of the 1950s, and how well Ben Hogan would do with the titanium drivers and extra-

An additional difficulty in modeling the effects of age on performance is that age does not have the same effect on all players. To handle such heterogeneity, we use random effects for each player's aging function, which allows for modeling players that deviate from the "standard" aging pattern. A desirable effect of using random curves is that each player is characterized by a career profile, rather than by a one-number summary. Player A may be better than player B when they are both 23 years old, and player A may be worse than player B when they are both 33 years


Home Run Proportion
0.0
0.05
0.10
0.15
1920
1940
1960
1980
Year
Ruth
McGwire


Home Run Proportion
0.0
0.05
0.10
0.15
1920
1940
1960
1980
Year
Ruth
Foxx
McGwire


Home Run Proportion
0.0
0.05
0.10
0.15
1920
1940
1960
1980
Year
Williams
Ruth
Foxx
McGwire


Home Run Proportion
0.0
0.05
0.10
0.15
1920
1940
1960
1980
Year
Williams
Aaron
McGwire
Ruth
Foxx

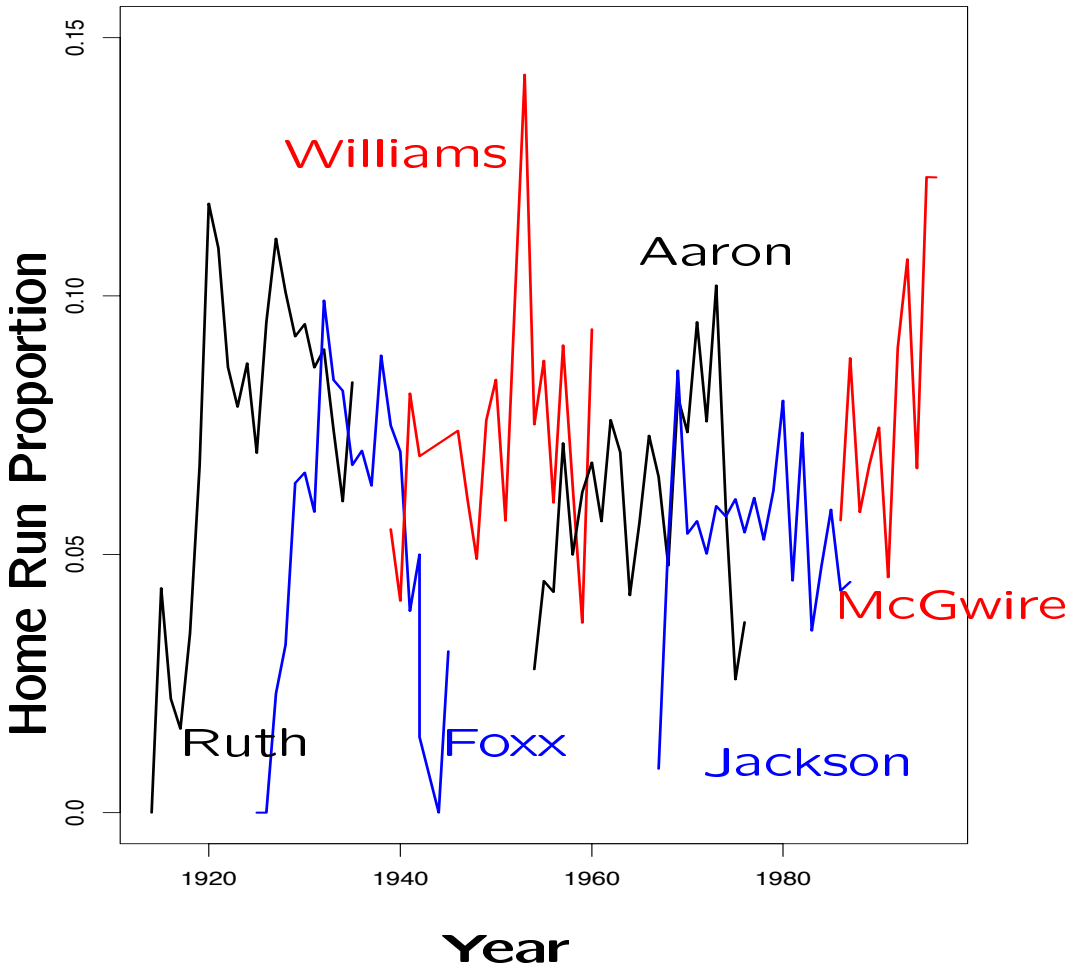

Home Run Proportion
0.0
0.05
0.10
0.15
1920
1940
1960
1980
Year
Williams
Aaron
McGwire
Ruth
Foxx
Jackson

Figure 11. A profile of some of the best players in the home run study. The estimated number of home runs, conditional on 500 at bats, for each age of the player, if that year were 1996.

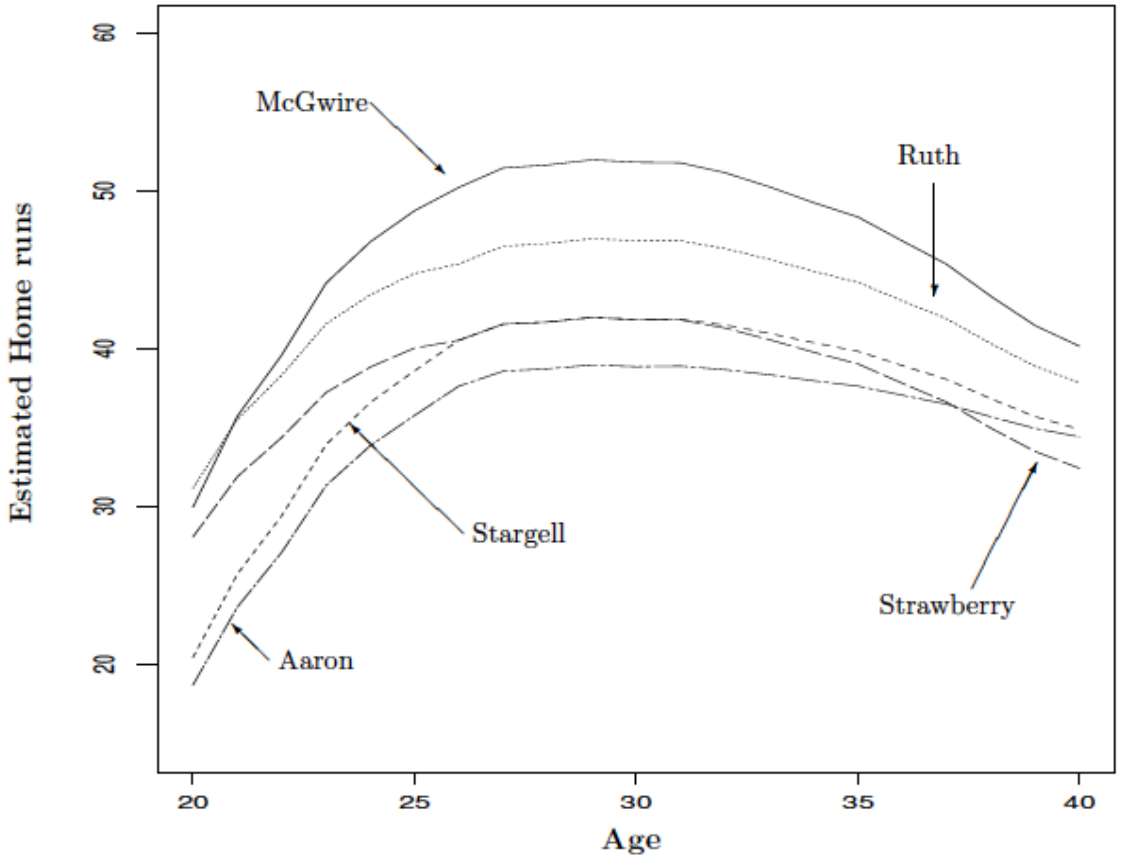
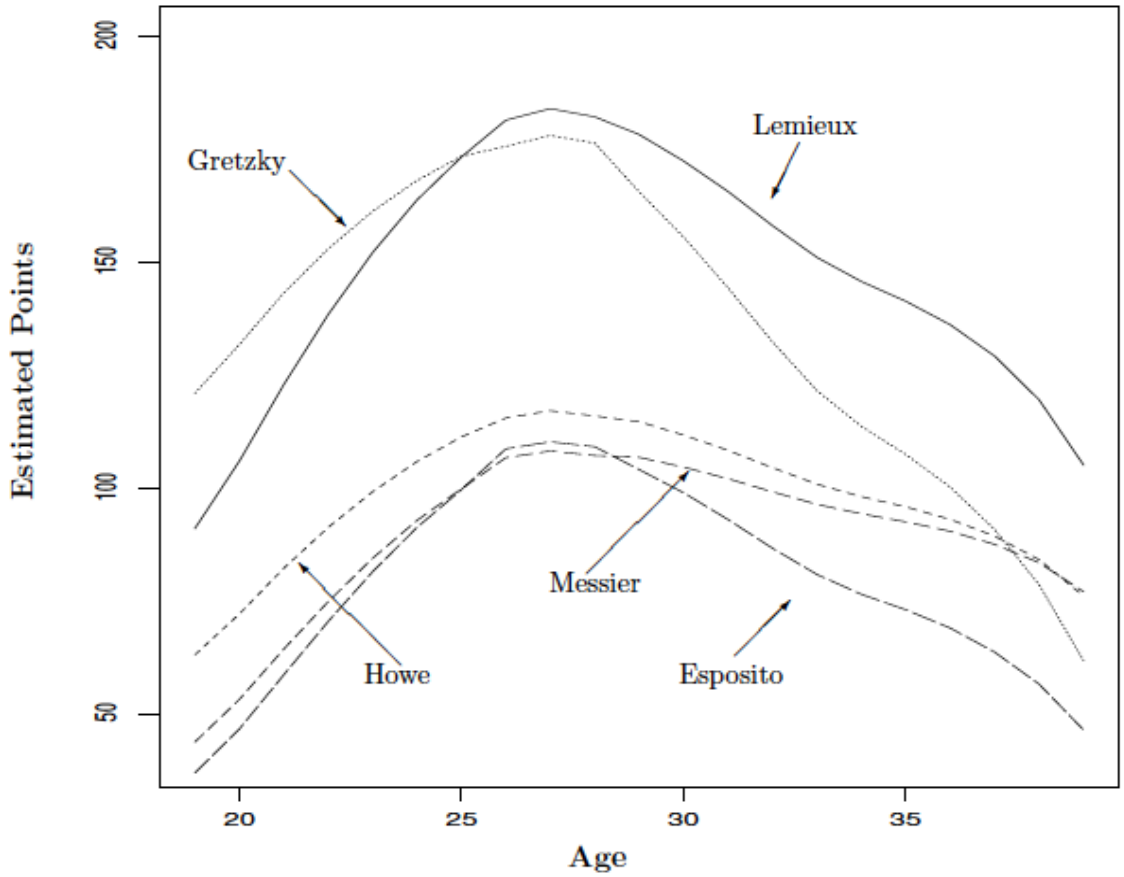
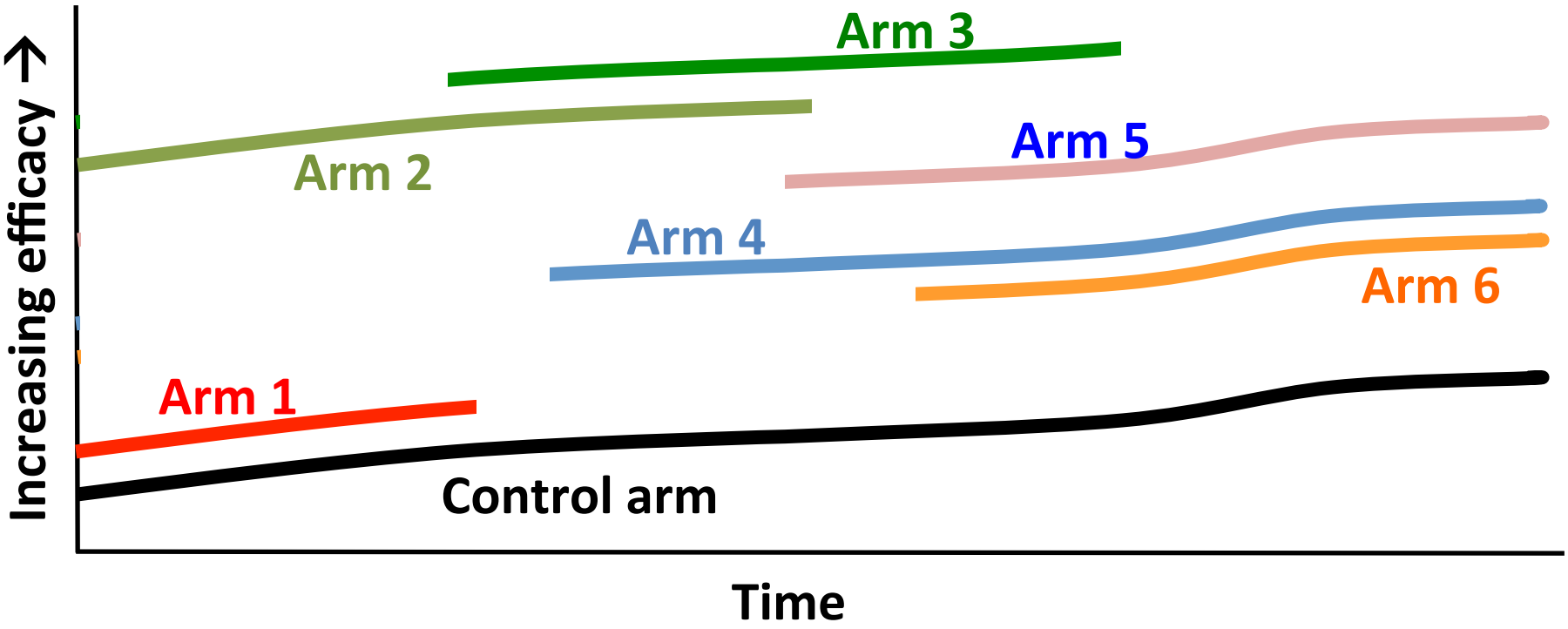


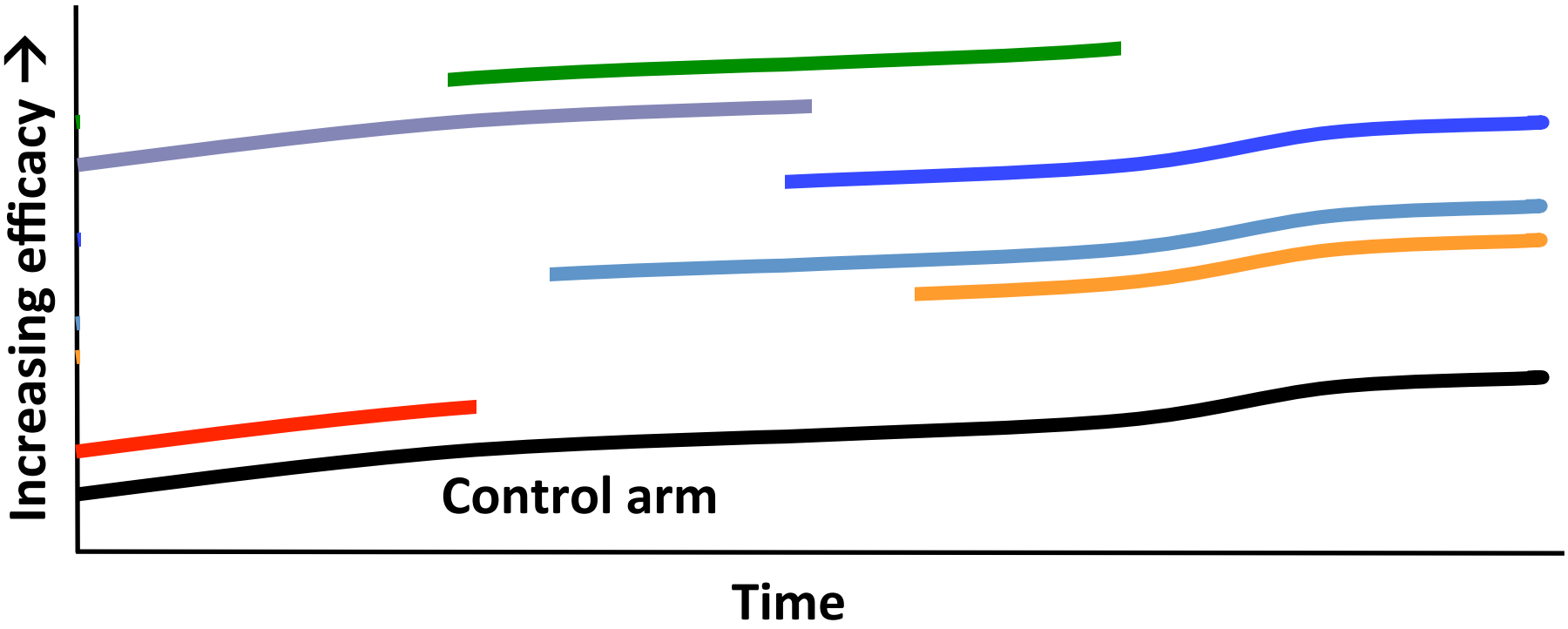
Figure 8. A profile of some of the best players in the hockey study. The estimated mean number of points for each age of the player, if that season were 1996.



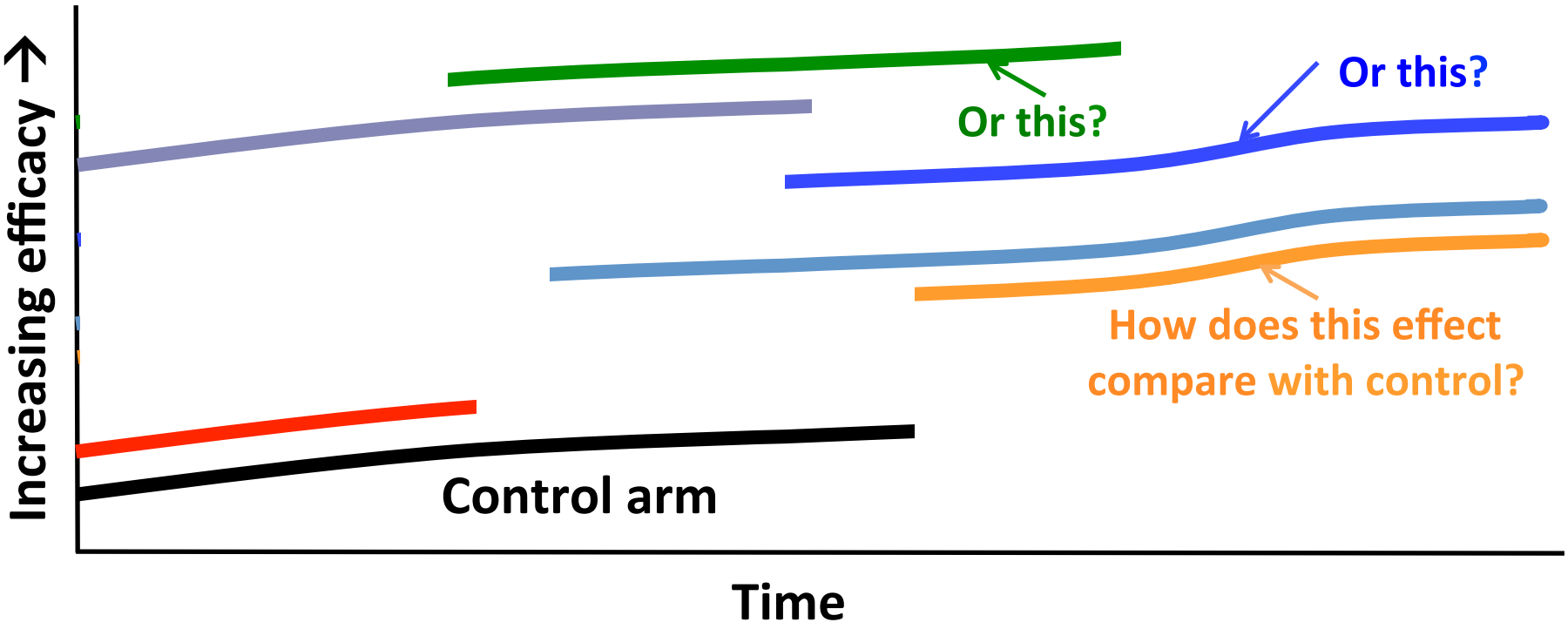
The Time Machine in Platform Trials



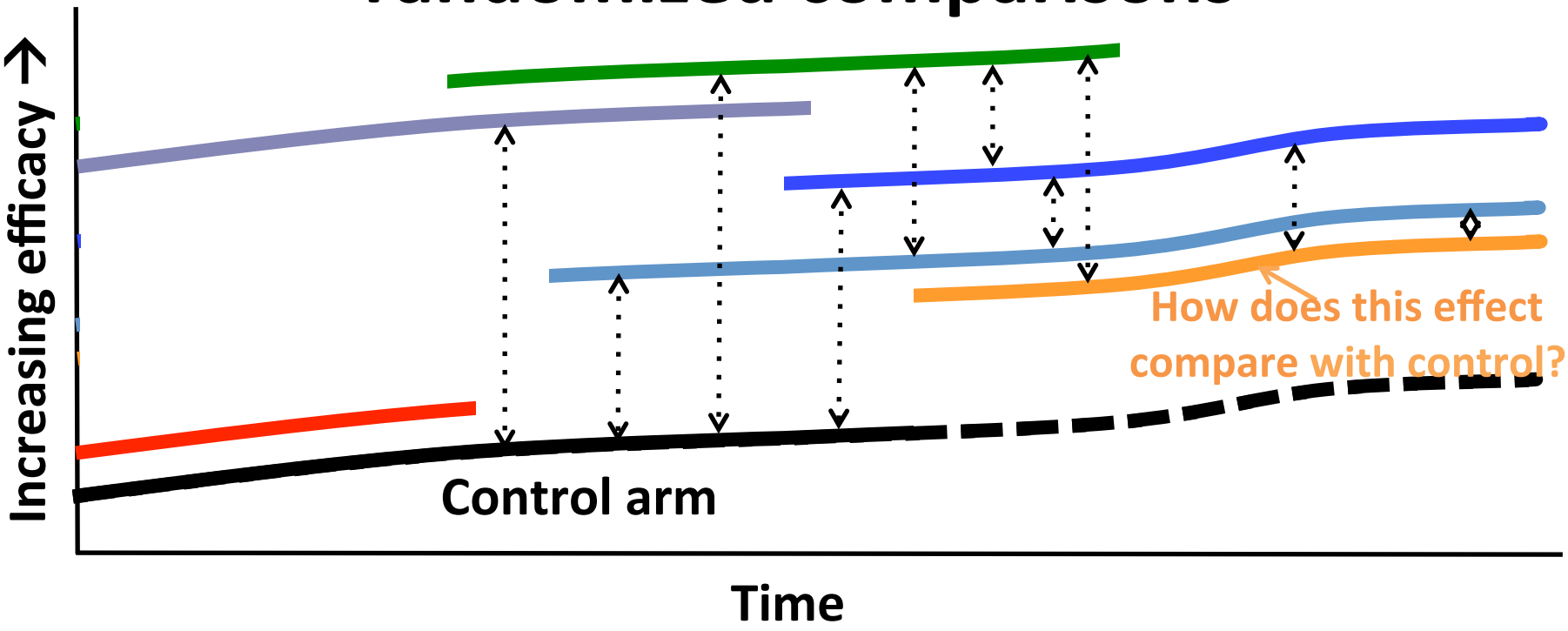
The Time Machine in Platform Trials



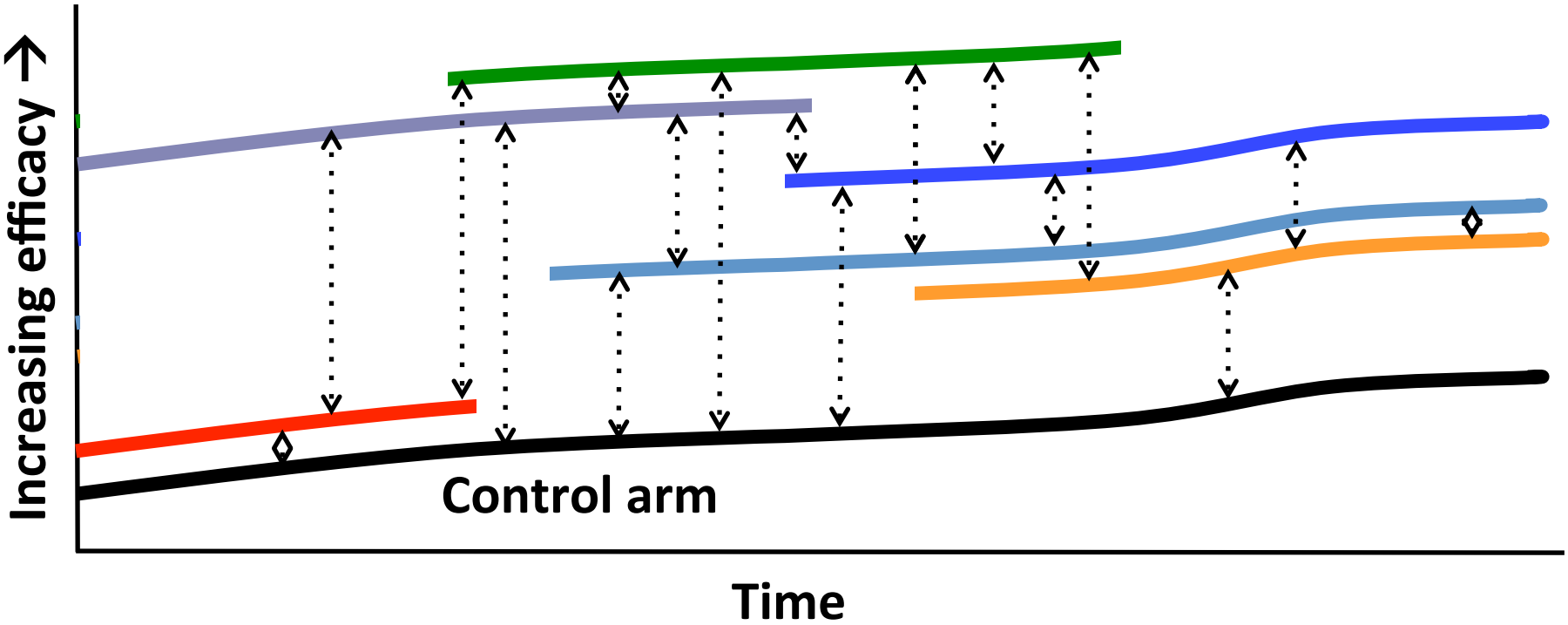
Suppose have to drop control arm



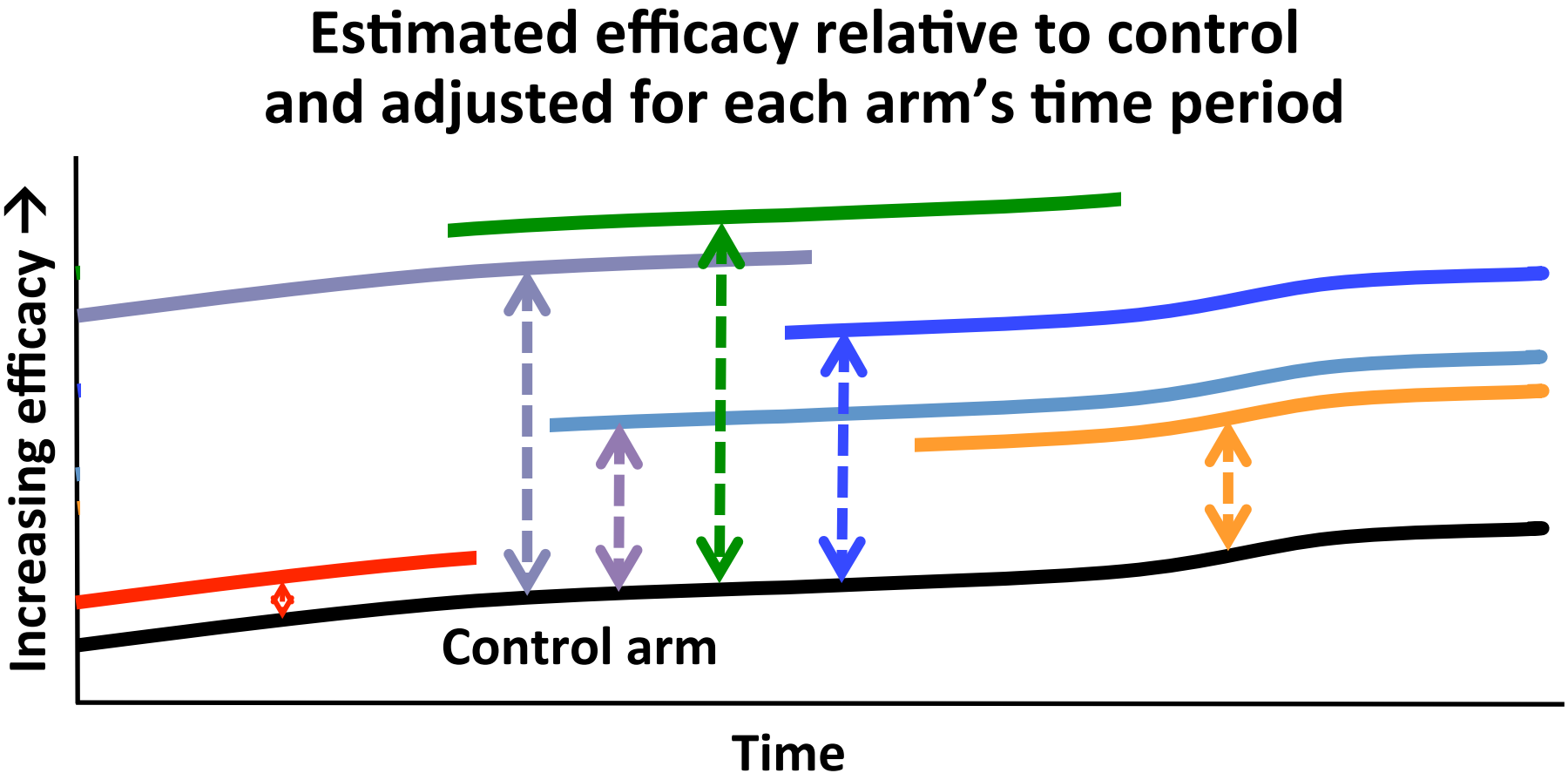
Enhancing controls via concurrently randomized comparisons



Every arm supports every other arm



17 concurrent comparisons of pairs of arms,
of which 6 are versus control


Estimated efficacy relative to control
and adjusted for each arm's time period
Increasing efficacy →
Control arm
Time

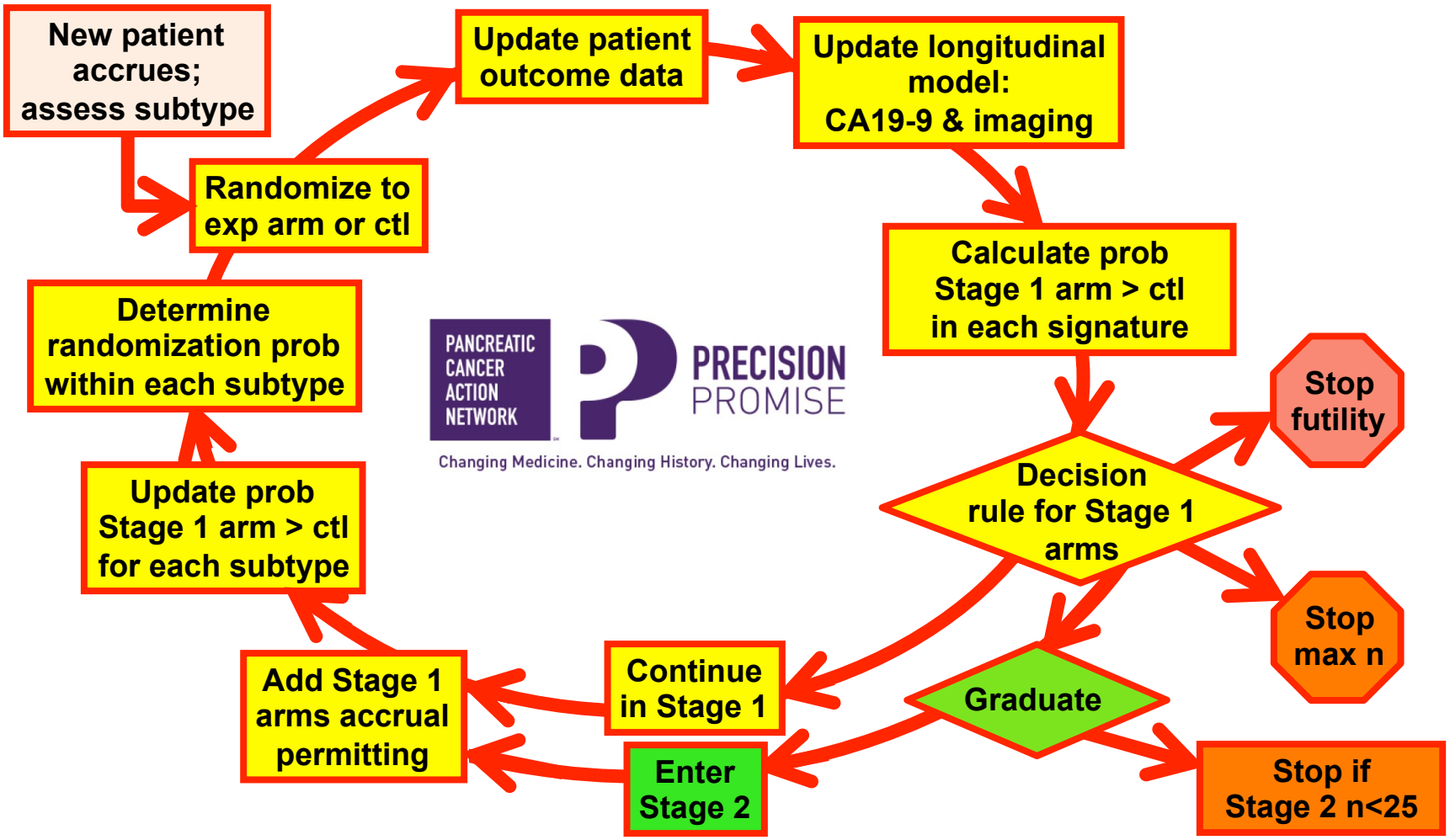
Phase 3 platform trials: GBM AGILE & Precision Promise

Goals

- Phase 2/3 for experimental drugs
- Many arms (& companies), incl combinations
- Stratify by biomarkers
- Identify biomarker signature of each arm
- Adaptively randomize within subtype vs control
- Move better therapies through faster
- Endpoint: Overall survival
- Inform OS using longitudinal model based on MRI and performance status plus …
- Sample size random; max in phase 2: ≤150 pts.

Innovations in Precision Promise

1. Seamless shift, learn (Stage 1) to confirm (Stage 2)
2. All patients, regardless of stage, count in final analysis
3. Many arms, that enter and leave the trial
4. Two controls, with hierarchical borrowing
5. Compare arms with all controls via “time machine”
6. Continuous learning and updating information
7. Adaptive randomization (in learn stage)
8. Identify arms’ indications, if any, including biomarkers
9. Interpretation of Type I error
10. Trial driven by predictive probability
11. Endpoint is OS, but re-randomize for 2nd-line therapy


New patient accrues; assess subtype
Update patient outcome data
Update longitudinal model: CA19-9 & imaging
Randomize to exp arm or ctl
Calculate prob Stage 1 arm > ctl in each signature
Determine randomization prob within each subtype
PANCREATIC CANCER ACTION NETWORK
PRECISION PROMISE
Changing Medicine. Changing History. Changing Lives.
Stop futility
Decision rule for Stage 1 arms
Update prob Stage 1 arm > ctl for each subtype
Stop max n
Graduate
Continue in Stage 1
Add Stage 1 arms accrual permitting
Enter Stage 2
Stop if Stage 2 n<25

Simulations required

- To find operating characteristics:
  - Control Type I error rate
  - Find power -- complicated
  - Sample size distribution
  - Arm's duration in trial
  - Amount of drug required
- Prospective design essential
- Longitudinal modeling (not yet included)
- Many scenarios/examples
- Accrual rate matters
- Other arms and their efficacies matter