Robust Experimental Design Is Critical for Effective Biomarker Discovery and Development

Donald A. Berry
dberry@mdanderson.org

Financial Disclosure

- Donald Berry is co-owner of Berry Consultants, LLC.
- Berry Consultants designs adaptive clinical trials for
  - Pharmaceutical companies
  - Medical device companies
  - NIH cooperative groups

2

COMMENTARY

Multiplicities in Cancer Research: Ubiquitous and Necessary Evils

Donald Berry

Manuscript received October 4, 2011; revised January 27, 2012; accepted May 31, 2012.

Correspondence to: Donald Berry, PhD, Department of Biostatistics, The University of Texas M.D. Anderson Cancer Center, 1400 Pressler St, 4-5062 Pickens Academic Tower, Houston, TX 77030-1402 (e-mail: dberry@mdanderson.org).

Scientific inquiry involves observations and measurements, some of which are planned and some of which are not. The most interesting or unusual observations might be regarded as discoveries and therefore particularly worthy of publication. However, the observation ous. Multiple observations increase t sleading publications. Statisticians rec itives, a strategy that increases the ra ould not be more difficult to determi right? How does one decide that any ll approach. The goal of this commen g results of empirical studies, with en ing what I call "silent multiplicities." B e or addressed in epidemiology and However, the lessons apply quite generally. I consider both frequentist and Bayesian statistical approaches.

A major source of bias In biomarker research: Multiplicities

J Natl Cancer Inst 2012;104:1124–1132

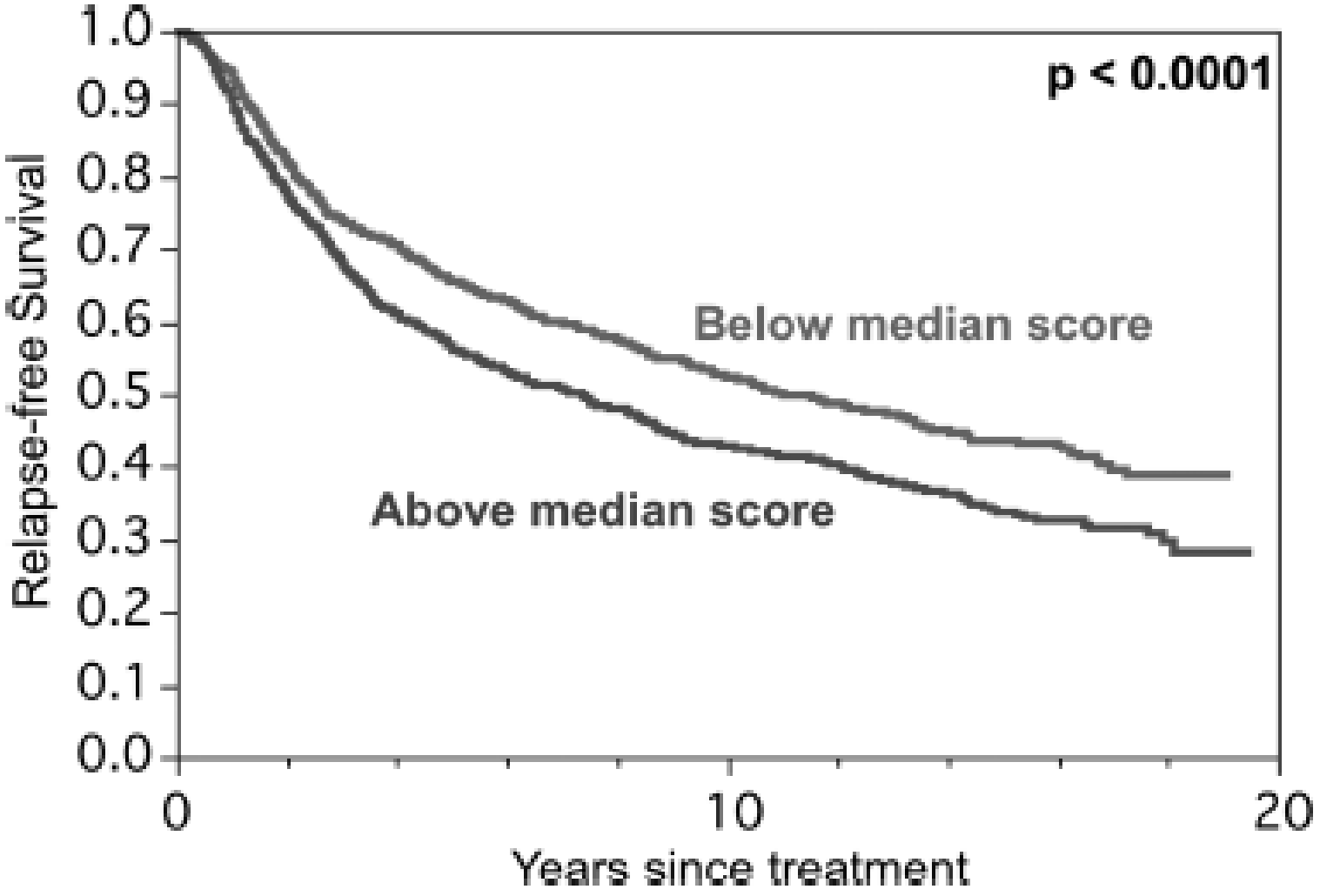
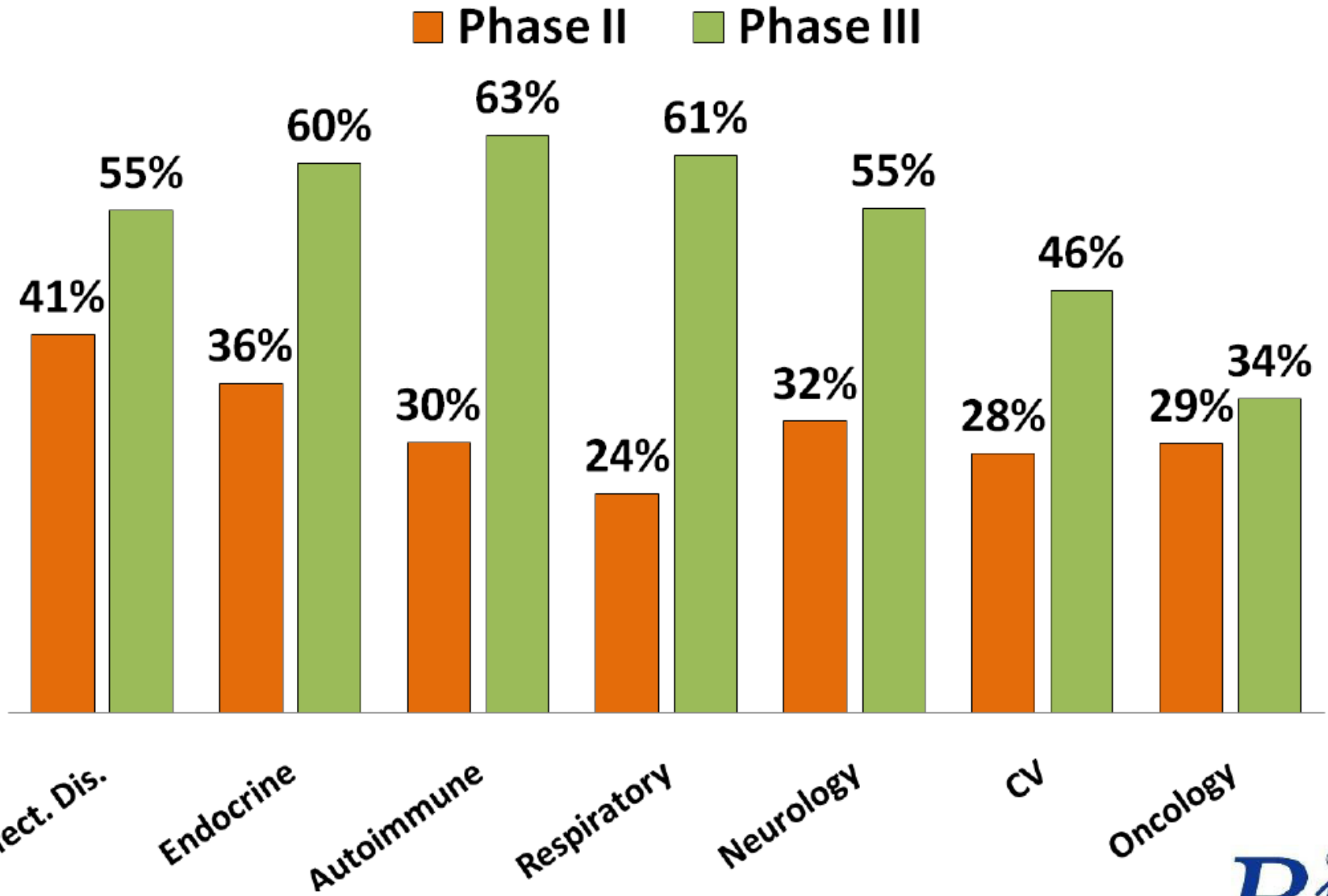


Figure 1. Results of building a prognostic score using five markers selected from 20 biomarkers in a 1550-patient clinical trial of node-positive breast cancer. The hazard ratio when comparing patients below the median with those above the median is 0.77.

OVERALL SUCCESS AT PHASE II AND III



biomed tracker

16 February 2011

Bio BIOTECHNOLOGY INDUSTRY ORGANIZATION

Why Phase III Failures?

- Ineffective d
  - Wrong endpoint in phase II: 70%
  - No randomization in phase II: 30%
  - Silly subsetting: 50%
  - Lottery: 30%
- Effective drug, lousy strategy (50%): of those:
  - Underpowered: 30%
  - Wrong dose/schedule/concomitant Rx: 60%
  - Wrong population: 80%

Need “robust design”

Verify “target” in phase II

FDA's Critical Path Opportunities Report (2006)

"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/SpecialTopics/CriticalPathInitiative/default.htm

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

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 circumstance?

Experimenting with Experiments

- It is ironic that as clinical trials investigate therapies with increasing biological complexity, the clinical trial itself has remained largely unchanged for 65 years.
- Except that they keep getting bigger!

We need many clinical trial experimenters

- Pharma
- Academia
- NCI
- Patient advocates!

Perhaps I-SPY 2 went too far

- Early “curable” disease
- Learning which experimental regimens benefit which biomarker subsets
- Regimens enter and leave the trial
- Screening all patients for markers ✓
- Master protocol ✓
- Registration endpoint
- Longitudinal modeling
- Combination therapies
- Adaptive sample size
- Adaptive randomization
- Learns about off-target effects
- Common control arm

Recent NCI/FNIH/FOCR Lung Cancer Master Protocol

S1400

Perhaps I-SPY 2 went too far

- ☐ Early “curable” disease
- ☐ Learning which experimental regimens benefit which biomarker subsets ✓
- ☐ Regimens enter and leave the trial ✓
- ☐ Screening all patients for markers ✓
- ☐ Master protocol ✓
- ☐ Registration endpoint
- ☐ Longitudinal modeling ✓
- ☐ Combination therapies ✓
- ☐ Adaptive sample size ✓
- ☐ Adaptive randomization ✓
- ☐ Learns about off-target effects ✓
- ☐ Common control arm ✓

Melanoma International Collaboration for Adaptive Trials

Multicompany trials adapt to disciplines beyond cancer

Asher Mullard

Nature Medicine **20**, 3 (2014) doi:10.1038/nm0114-3
Published online 07 January 2014

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 ($73 million) call for proposals for a similarly designed trial in Alzheimer's disease. Already, at least 12 drug companies are keen to participate.

“I think this is a step in a right direction to really try to beat Alzheimer's in a serious way,” says Randall Bateman, a neurologist at the Washington University School of Medicine in St. Louis. Bateman is coordinating an ongoing collaborative and adaptive trial called DIAN-TU, which is testing two experimental antibody drugs from Eli Lilly and Roche in patients with dominantly inherited forms of early-onset Alzheimer's disease. The bigger IMI-initiated trial, he says, is likely to offer hope to a much larger patient population.

Fanatic Studio / Alamy

Across the divide: Drug companies are teaming up to run trials in many disease areas.

Basket Trials (Tumor Agnostic)

- Ongoing in pharma
- Targeted drug, develop simultaneously across organ-specific cancers
- Borrow across histologies (not pool)
- Restrict to tumors expressing target
- Population sizes small means trial sample sizes must be small (n=10-15)

Formal, reproducible approach: Hierarchical modeling

- Formalizes the “Gleevec phenomenon”
- Prospective borrowing across histologies based on similarity of results
- Identify clusters of histologies, responsive and not responsive

Final words

- Biomarkers are notoriously difficult inferential beasts, but we most address them and conquer them if we are to defeat cancer
- We must build clinical trials that learn & confirm the roles of biomarkers
- Cancer trials must become tuned to biomarkers and be smaller, more accurate, and adaptive