Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Use of Biomarkers in Clinical Trial Design

Richard Simon, D.Sc.
Chief, Biometric Research Branch
National Cancer Institute
http://brb.nci.nih.gov

1

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Biometric Research Branch Website

brb.nci.nih.gov

- Powerpoint presentations
- Reprints & Technical Reports
- BRB-ArrayTools software
- Web based Sample Size Planning

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Different Kinds of Biomarkers

- Surrogate endpoints
  - A measurement made on a patient before, during and after treatment to determine whether the treatment is working
- Prognostic biomarkers
  - Measured before treatment to indicate long-term outcome for patients untreated or receiving standard treatment
- Predictive biomarkers
  - Measured before treatment to identify who will benefit from a particular treatment

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Surrogate Endpoints

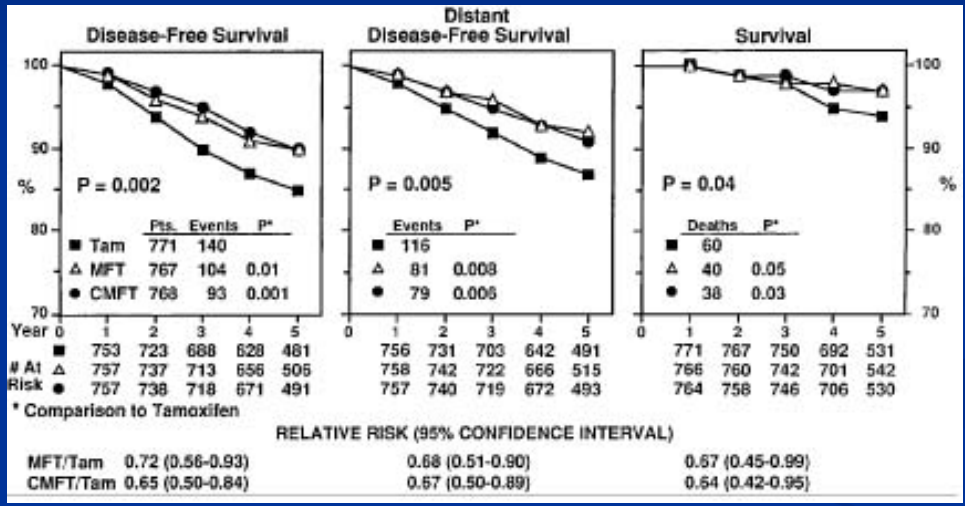
- It is very difficult to properly validate a biomarker as a surrogate of clinical benefit for use as an alternative endpoint in phase III trials
- Biomarkers can be useful in phase I/II studies as measures of treatment effect
  - they need not be validated as surrogates for clinical benefit
  - Unvalidated surrogates can also be used for interime "futility analyses" of phase III trials. The trial should continue accrual and follow-up to evaluate true endpoint if treatment effect on biomarker is sufficient

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Prognostic & Predictive Biomarkers

- Many cancer treatments benefit only a minority of patients to whom they are administered
  - Particularly true for molecularly targeted drugs
- Being able to predict which patients are likely to benefit would
  - save patients from unnecessary toxicity, and enhance their chance of receiving a drug that helps them
  - Help control medical costs
  - Improve the success rate of clinical drug development

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute



Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Prognostic and Predictive Biomarkers in Oncology

- Single gene or protein measurement
  - e.g. HER2 protein staining 2+ or 3+
  - HER2 amplification
  - KRAS mutation
- Scalar index or classifier that summarizes contributions of multiple genes/proteins
  - Empirically determined based on genome-wide correlating gene expression to patient outcome after treatment

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Prognostic Factors in Oncology

- Most prognostic factors are not used because they are not therapeutically relevant
- Most prognostic factor studies do not have a clear medical objective
  - They use a convenience sample of patients for whom tissue is available.
  - Generally the patients are too heterogeneous to support therapeutically relevant conclusions
- Most prognostic factor studies are not reliable because they are not prospectively focused on a single factor

8

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Pusztai et al. The Oncologist 8:252-8, 2003

- 939 articles on "prognostic markers" or "prognostic factors" in breast cancer in past 20 years
- ASCO guidelines only recommend routine testing for ER, PR and HER-2 in breast cancer
- "With the exception of ER or progesterone receptor expression and HER-2 gene amplification, there are no clinically useful molecular predictors of response to any form of anticancer therapy."

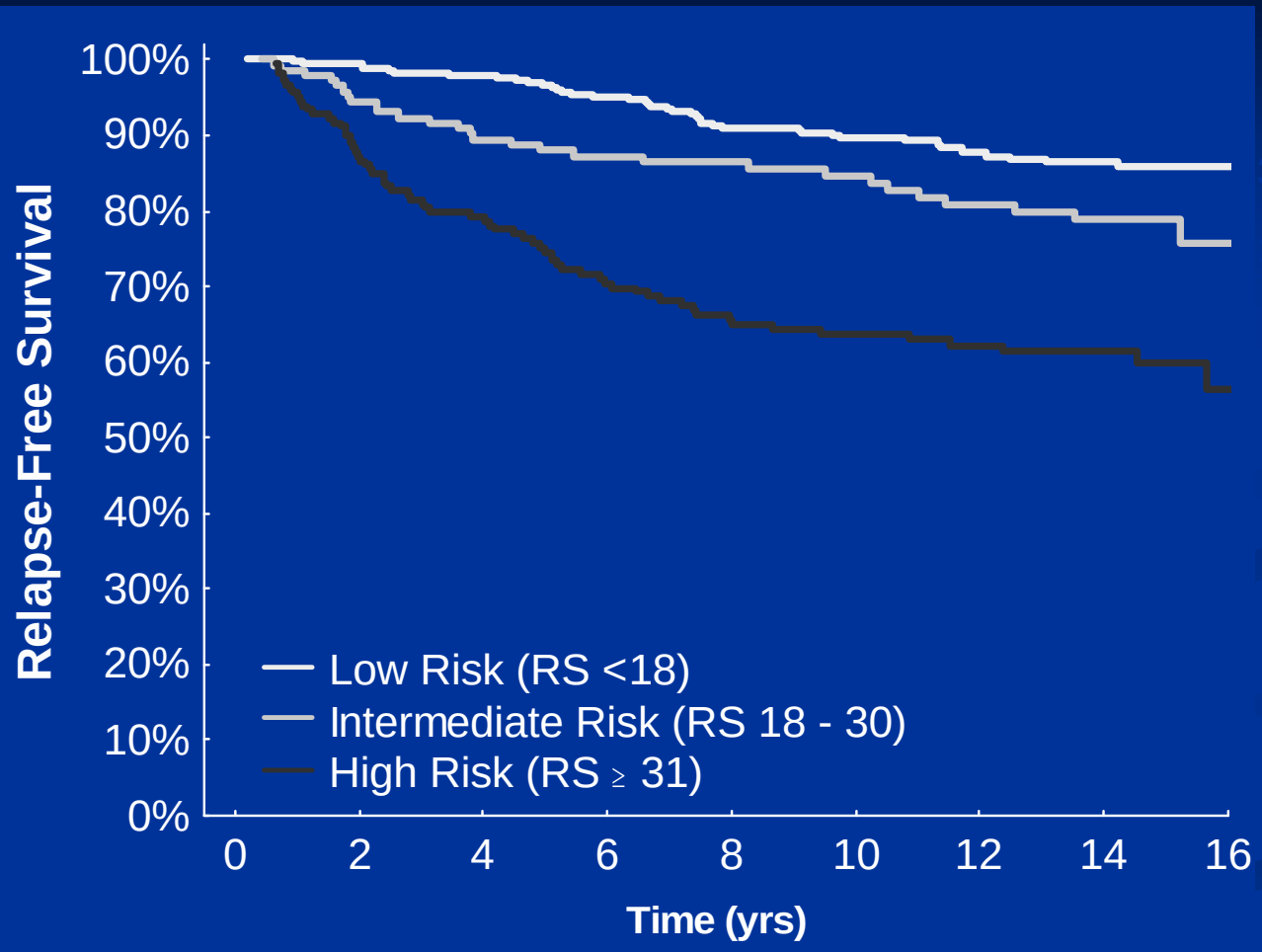
Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Prognostic Biomarkers Can be Therapeutically Relevant

- <10% of node negative ER+ breast cancer patients require or benefit from the cytotoxic chemotherapy that they receive
- OncotypeDx
  - 21 gene RTPCR assay for FFPE tissue

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

B-14 Results—Relapse-Free Survival



11

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Key Features of OncotypeDx Development

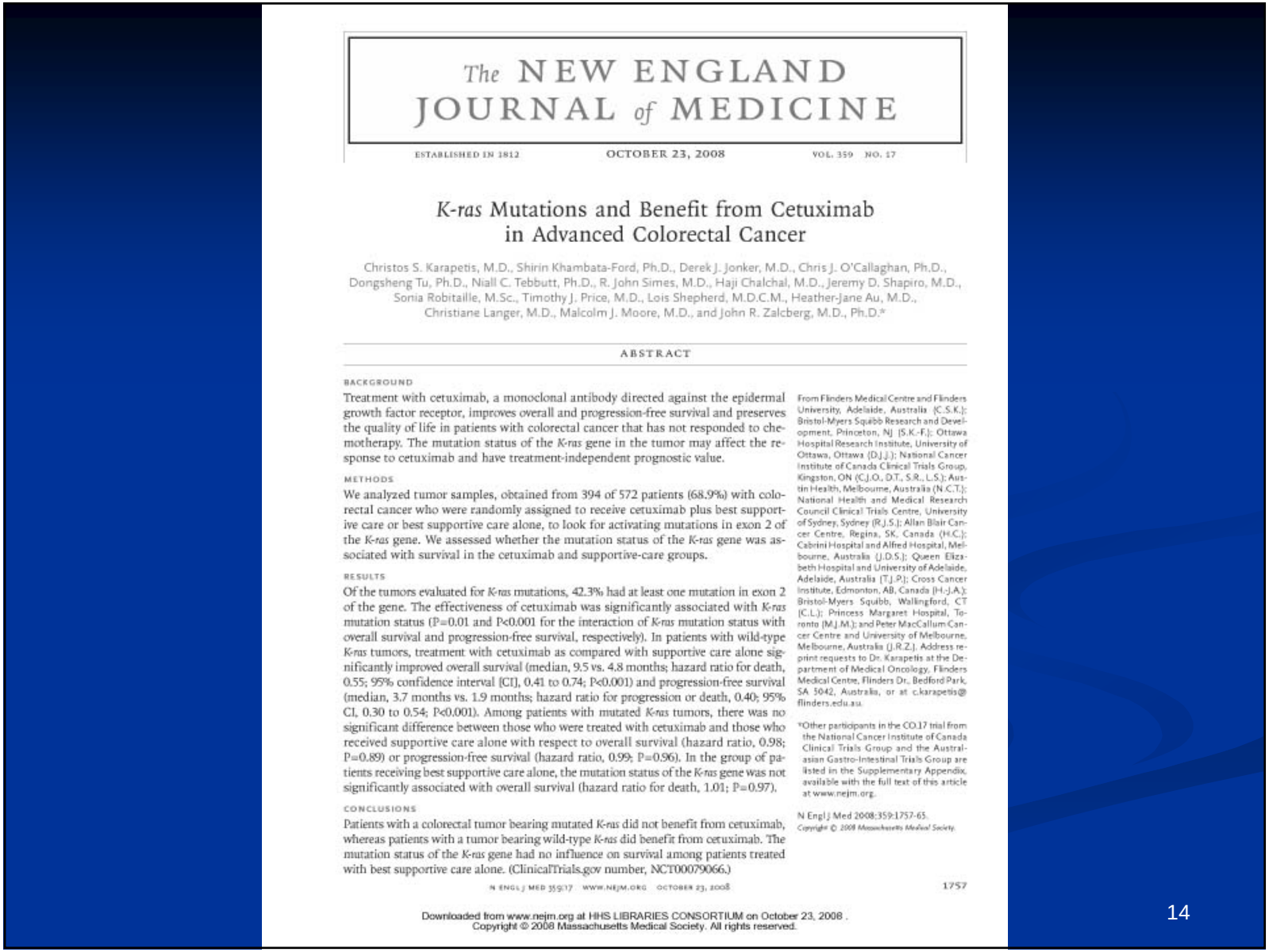
- Identification of important therapeutic decision context
- Prognostic marker development was based on patients with node negative ER positive breast cancer receiving tamoxifen as only systemic treatment
- Staged development and validation
  - Separation of data used for test development from data used for test validation
- Development of robust assay with rigorous analytical validation
  - 21 gene RTPCR assay for FFPE tissue
  - Quality assurance by single reference laboratory operation

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Predictive Biomarkers

- In the past often studied as un-focused post-hoc subset analyses of RCTs.
  - Numerous subsets examined
  - Same data used to define subsets for analysis and for comparing treatments within subsets
  - No control of type I error

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute



The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 OCTOBER 23, 2008 VOL. 359 NO. 17

K-ras Mutations and Benefit from Cetuximab in Advanced Colorectal Cancer

Christos S. Karapetis, M.D., Shirin Khambata-Ford, Ph.D., Derek J. Jonker, M.D., Chris J. O'Callaghan, Ph.D., Dongsheng Tu, Ph.D., Niall C. Tebbutt, Ph.D., R. John Simes, M.D., Haji Chalchal, M.D., Jeremy D. Shapiro, M.D., Sonia Robitaille, M.Sc., Timothy J. Price, M.D., Lois Shepherd, M.D.C.M., Heather-Jane Au, M.D., Christiane Langer, M.D., Malcolm J. Moore, M.D., and John R. Zalcberg, M.D., Ph.D.*

ABSTRACT

BACKGROUND

Treatment with cetuximab, a monoclonal antibody directed against the epidermal growth factor receptor, improves overall and progression-free survival and preserves the quality of life in patients with colorectal cancer that has not responded to chemotherapy. The mutation status of the *K-ras* gene in the tumor may affect the response to cetuximab and have treatment-independent prognostic value.

METHODS

We analyzed tumor samples, obtained from 394 of 572 patients (68.9%) with colorectal cancer who were randomly assigned to receive cetuximab plus best supportive care or best supportive care alone, to look for activating mutations in exon 2 of the *K-ras* gene. We assessed whether the mutation status of the *K-ras* gene was associated with survival in the cetuximab and supportive-care groups.

RESULTS

Of the tumors evaluated for *K-ras* mutations, 42.3% had at least one mutation in exon 2 of the gene. The effectiveness of cetuximab was significantly associated with *K-ras* mutation status ($P=0.01$ and $P<0.001$ for the interaction of *K-ras* mutation status with overall survival and progression-free survival, respectively). In patients with wild-type *K-ras* tumors, treatment with cetuximab as compared with supportive care alone significantly improved overall survival (median, 9.5 vs. 4.8 months; hazard ratio for death, 0.55; 95% confidence interval [CI], 0.41 to 0.74; $P<0.001$) and progression-free survival (median, 3.7 months vs. 1.9 months; hazard ratio for progression or death, 0.40; 95% CI, 0.30 to 0.54; $P<0.001$). Among patients with mutated *K-ras* tumors, there was no significant difference between those who were treated with cetuximab and those who received supportive care alone with respect to overall survival (hazard ratio, 0.98; $P=0.89$) or progression-free survival (hazard ratio, 0.99; $P=0.96$). In the group of patients receiving best supportive care alone, the mutation status of the *K-ras* gene was not significantly associated with overall survival (hazard ratio for death, 1.01; $P=0.97$).

CONCLUSIONS

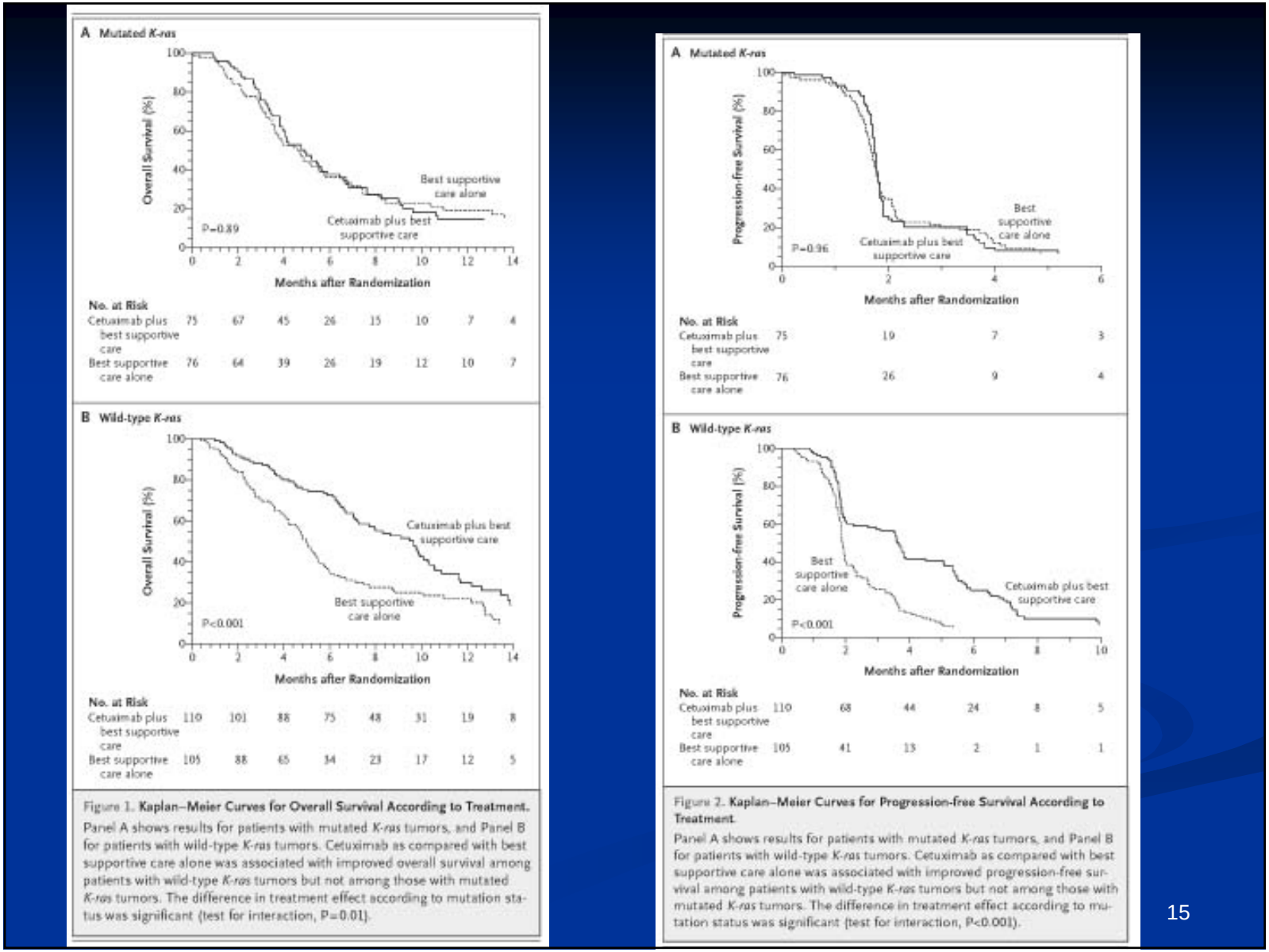
Patients with a colorectal tumor bearing mutated *K-ras* did not benefit from cetuximab, whereas patients with a tumor bearing wild-type *K-ras* did benefit from cetuximab. The mutation status of the *K-ras* gene had no influence on survival among patients treated with best supportive care alone. (ClinicalTrials.gov number, NCT00079066.)

From Flinders Medical Centre and Flinders University, Adelaide, Australia (C.S.K.); Bristol-Myers Squibb Research and Development, Princeton, NJ (S.K.-F.); Ottawa Hospital Research Institute, University of Ottawa, Ottawa (D.J.J.); National Cancer Institute of Canada Clinical Trials Group, Kingston, ON (C.J.O., D.T., S.R., L.S.); Austin Health, Melbourne, Australia (N.C.T.); National Health and Medical Research Council Clinical Trials Centre, University of Sydney, Sydney (R.J.S.); Allan Blair Cancer Centre, Regina, SK, Canada (H.C.); Cabrini Hospital and Alfred Hospital, Melbourne, Australia (J.D.S.); Queen Elizabeth Hospital and University of Adelaide, Adelaide, Australia (T.J.P.); Cross Cancer Institute, Edmonton, AB, Canada (H.-J.A.); Bristol-Myers Squibb, Wallingford, CT (C.L.); Princess Margaret Hospital, Toronto (M.J.M.); and Peter MacCallum Cancer Centre and University of Melbourne, Melbourne, Australia (J.R.Z.). Address reprint requests to Dr. Karapetis at the Department of Medical Oncology, Flinders Medical Centre, Flinders Dr., Bedford Park, SA 5042, Australia, or at c.karapetis@flinders.edu.au.

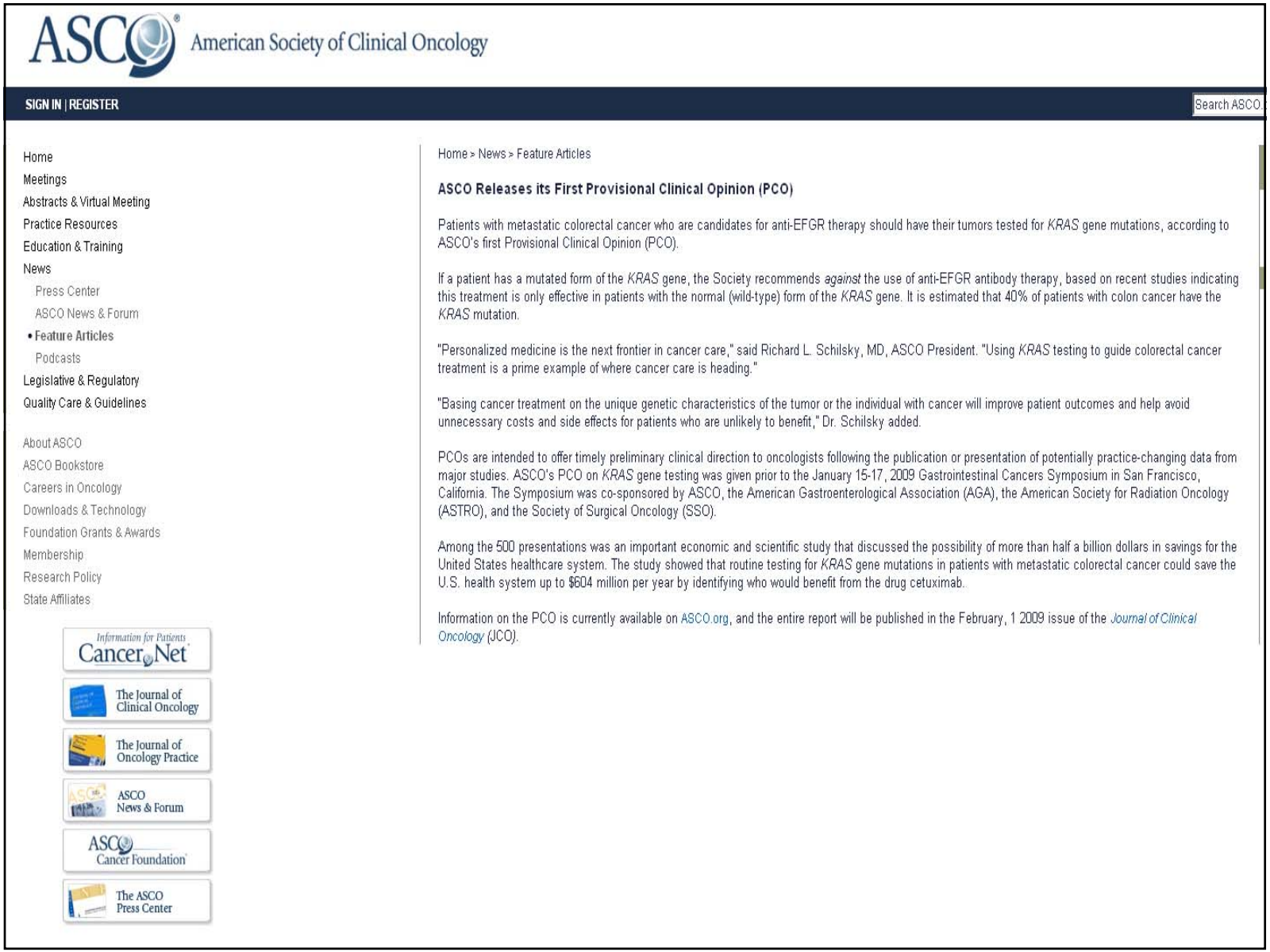
*Other participants in the CO.17 trial from the National Cancer Institute of Canada Clinical Trials Group and the Australasian Gastro-Intestinal Trials Group are listed in the Supplementary Appendix, available with the full text of this article at www.nejm.org.

N Engl J Med 2008;359:1757-65.

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute



Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute



ASCO American Society of Clinical Oncology

SIGN IN | REGISTER

Search ASCO

Home
Meetings
Abstracts & Virtual Meeting
Practice Resources
Education & Training
News
Press Center
ASCO News & Forum
• Feature Articles
Podcasts
Legislative & Regulatory
Quality Care & Guidelines

About ASCO
ASCO Bookstore
Careers in Oncology
Downloads & Technology
Foundation Grants & Awards
Membership
Research Policy
State Affiliates

Information for Patients CancerNet
The Journal of Clinical Oncology
The Journal of Oncology Practice
ASCO News & Forum
ASCO Cancer Foundation
The ASCO Press Center

Home » News » Feature Articles

ASCO Releases its First Provisional Clinical Opinion (PCO)

Patients with metastatic colorectal cancer who are candidates for anti-EFGR therapy should have their tumors tested for *KRAS* gene mutations, according to ASCO's first Provisional Clinical Opinion (PCO).

If a patient has a mutated form of the *KRAS* gene, the Society recommends *against* the use of anti-EFGR antibody therapy, based on recent studies indicating this treatment is only effective in patients with the normal (wild-type) form of the *KRAS* gene. It is estimated that 40% of patients with colon cancer have the *KRAS* mutation.

"Personalized medicine is the next frontier in cancer care," said Richard L. Schilsky, MD, ASCO President. "Using *KRAS* testing to guide colorectal cancer treatment is a prime example of where cancer care is heading."

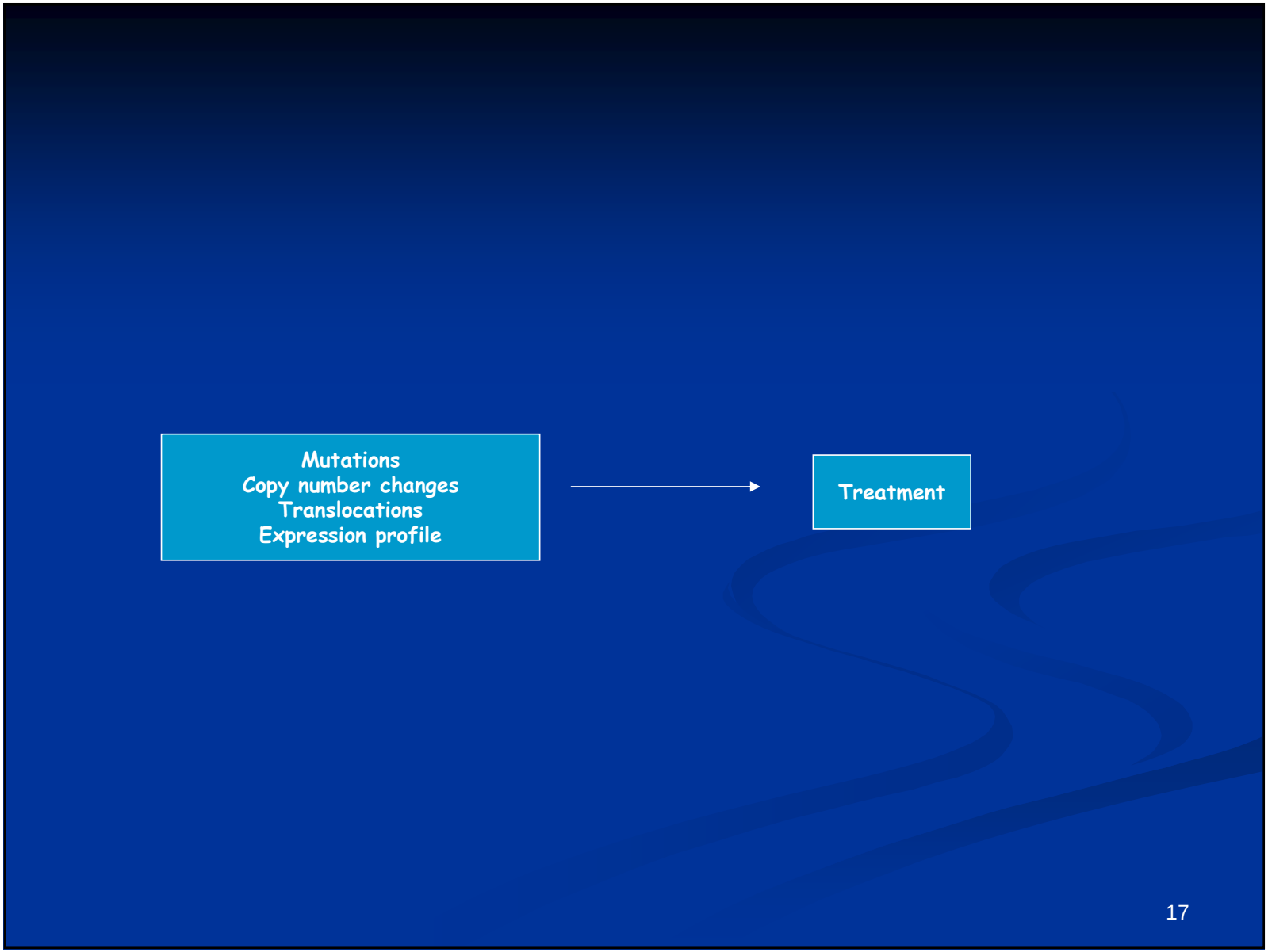
"Basing cancer treatment on the unique genetic characteristics of the tumor or the individual with cancer will improve patient outcomes and help avoid unnecessary costs and side effects for patients who are unlikely to benefit," Dr. Schilsky added.

PCOs are intended to offer timely preliminary clinical direction to oncologists following the publication or presentation of potentially practice-changing data from major studies. ASCO's PCO on *KRAS* gene testing was given prior to the January 15-17, 2009 Gastrointestinal Cancers Symposium in San Francisco, California. The Symposium was co-sponsored by ASCO, the American Gastroenterological Association (AGA), the American Society for Radiation Oncology (ASTRO), and the Society of Surgical Oncology (SSO).

Among the 500 presentations was an important economic and scientific study that discussed the possibility of more than half a billion dollars in savings for the United States healthcare system. The study showed that routine testing for *KRAS* gene mutations in patients with metastatic colorectal cancer could save the U.S. health system up to $604 million per year by identifying who would benefit from the drug cetuximab.

Information on the PCO is currently available on ASCO.org, and the entire report will be published in the February, 1 2009 issue of the *Journal of Clinical Oncology* (JCO).

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute



Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

VOLUME 23 · NUMBER 29 · OCTOBER 10 2005

JOURNAL OF CLINICAL ONCOLOGY

REVIEW ARTICLE

Roadmap for Developing and Validating Therapeutically Relevant Genomic Classifiers

Richard Simon

From the Biometric Research Branch, Division of Cancer Treatment and Diagnosis, National Cancer Institute, National Institutes of Health, Bethesda, MD.

Submitted May 26, 2005; accepted June 18, 2005.

Terms in blue are defined in the glossary, found at the end of this issue and online at www.jco.org.

Author's disclosures of potential conflicts of interest are found at the end of this article.

Address reprint requests to Richard Simon, DSc, National Cancer Institute, 9000 Rockville Pike, MSC 7434, Bethesda, MD 20892; e-mail: rsimon@nih.gov.

0732-183X/05/2329-7332/$20.00

DOI: 10.1200/JCO.2005.02.8712

A B S T R A C T

Oncologists need improved tools for selecting treatments for individual patients. The development of therapeutically relevant prognostic markers has traditionally been slowed by poor study design, inconsistent findings, and lack of proper validation studies. Microarray expression profiling provides an exciting new technology for relating tumor gene expression to patient outcome, but it also provides increased challenges for translating initial research findings into robust diagnostics that benefit patients and physicians in therapeutic decision making. This article attempts to clarify some of the misconceptions about the development and validation of multigene expression signature classifiers and highlights the steps needed to move genomic signatures into clinical application as therapeutically relevant and robust diagnostics.

J Clin Oncol 23:7332-7341.

INTRODUCTION

Oncologists need improved tools for selecting treatments for individual patients. Most cancer treatments benefit only a minority of the patients to whom they are administered. Being able to predict which patients are most likely to benefit would not only save patients from unnecessary toxicity and inconvenience, but might facilitate their receiving drugs that are more likely to help them. In addition, the current overtreatment of patients results in major expense for individuals and society, an expense that may not be indefinitely sustainable.

Microarray expression profiling has provided an exciting new technology for attempting to identify classifiers for tailoring treatments to patients. To date, however, no multigene expression signature has been widely adopted into oncology practice and very few are close to achieving such status. Development of biomarker classifiers useful for improving treatment decisions and sufficiently validated for broad clinical application is difficult, and more difficult for expression signature classifiers. The field of microarray expression profiling is also burdened with both unrealistic hype and excessive skepticism. In this article, I will attempt to clarify some of the misconceptions about the development and validation of multigene expression signature classifiers and highlight the steps needed to move genomic signatures into clinical application as therapeutically relevant and robust diagnostics.

WHY ARE SO FEW PROGNOSTIC FACTORS USED IN ONCOLOGY?

Although there is a large literature on prognostic factors for cancer patients, very few such factors are used in clinical practice. Prognostic factors are unlikely to be used unless they are therapeutically relevant, and most publications do not establish such relevance. Most prognostic factor studies are conducted using a convenience sample of patients for whom tissue is available, but the cohort is often far too heterogeneous with regard to stage and treatment to support therapeutically relevant conclusions. Additional problems in the prognostic marker literature derive from the fact that most studies develop prognostic

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Prospective Co-Development of Drugs and Companion Diagnostics

1. Develop a completely specified genomic classifier of the patients likely to benefit from a new drug
   - Single gene/protein
   - Gene expression signature
     - Screen genes using microarrays
     - Develop classifier for RT-PCR platform
   - Pre-clinical, phase II data, archived specimens from previous phase III studies
2. Establish analytical validity of the classifier
3. Use the completely specified classifier to design and analyze a new clinical trial to evaluate effectiveness of the new treatment with a pre-defined analysis plan that preserves the overall type-I error of the study.

19

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Guiding Principle

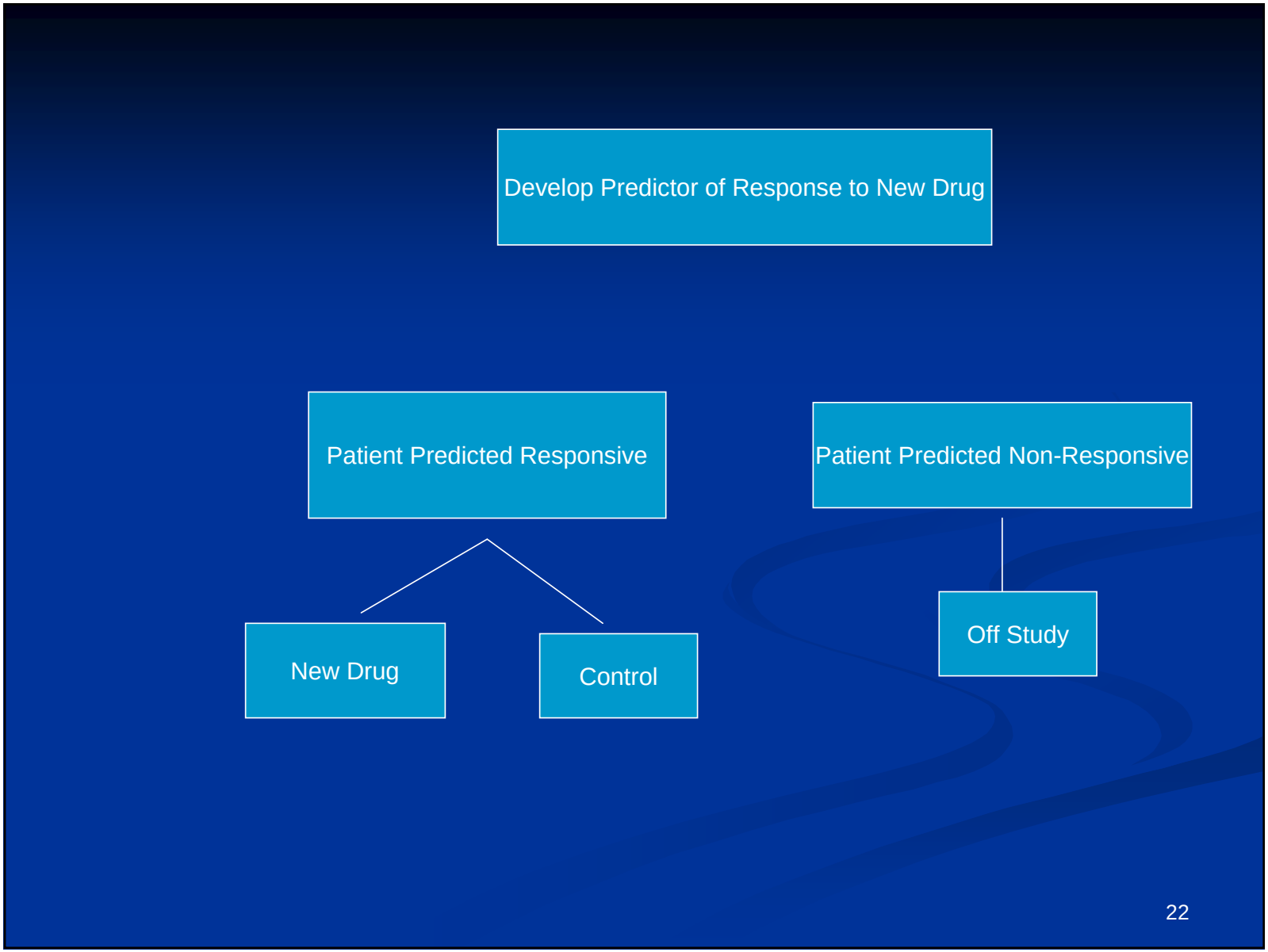
- The data used to develop the classifier should be distinct from the data used to test hypotheses about treatment effect in subsets determined by the classifier
  - Developmental studies can be exploratory
  - Studies on which treatment effectiveness claims are to be based should be definitive studies that test a treatment hypothesis in a patient population completely pre-specified by the classifier

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

New Drug Developmental Strategy I

- Restrict entry to the phase III trial based on the binary predictive classifier, i.e. targeted design

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute



Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Applicability of Design I

- Primarily for settings where the classifier is based on a single gene whose protein product is the target of the drug
  - eg Herceptin

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Evaluating the Efficiency of Strategy (I)

- Simon R and Maitnourim A. Evaluating the efficiency of targeted designs for randomized clinical trials. Clinical Cancer Research 10:6759-63, 2004; Correction and supplement 12:3229, 2006
- Maitnourim A and Simon R. On the efficiency of targeted clinical trials. Statistics in Medicine 24:329-339, 2005

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

- Relative efficiency of targeted design depends on
  - proportion of patients test positive
  - effectiveness of new drug (compared to control) for test negative patients
- When less than half of patients are test positive and the drug has little or no benefit for test negative patients, the targeted design requires dramatically fewer randomized patients
- The targeted design may require fewer or more screened patients than the standard design

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Trastuzumab
Herceptin

- Metastatic breast cancer
- 234 randomized patients per arm
- 90% power for 13.5% improvement in 1-year survival over 67% baseline at 2-sided .05 level
- If benefit were limited to the 25% test + patients, overall improvement in survival would have been 3.375%
  - 4025 patients/arm would have been required

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

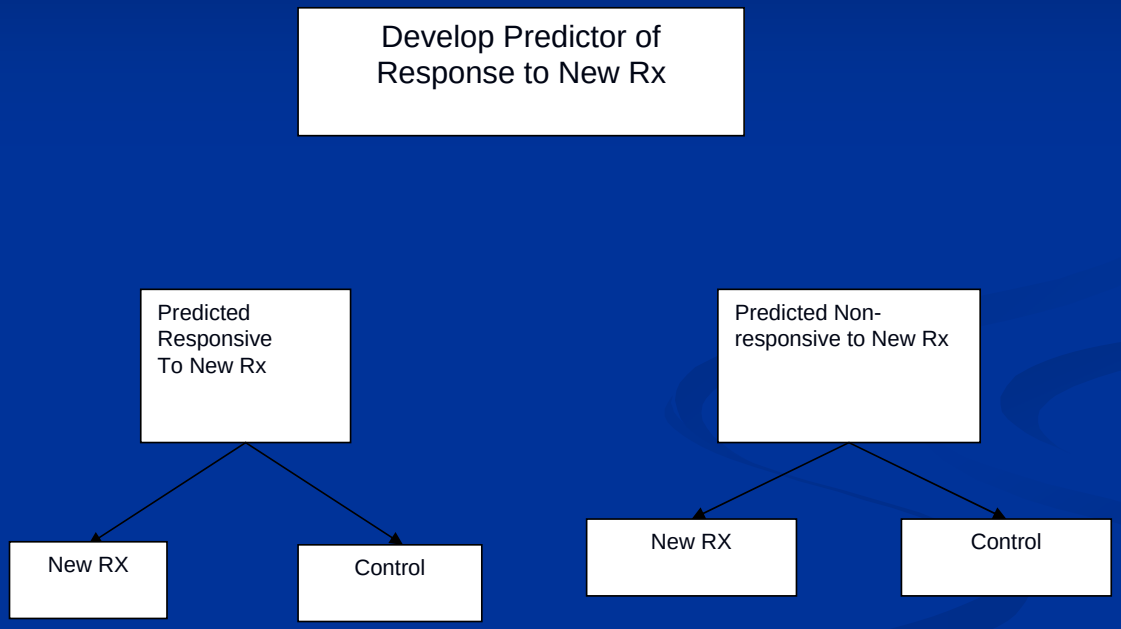
Web Based Software for Comparing Sample Size Requirements

- http://brb.nci.nih.gov

27

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Developmental Strategy (II)



28

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Developmental Strategy (II)

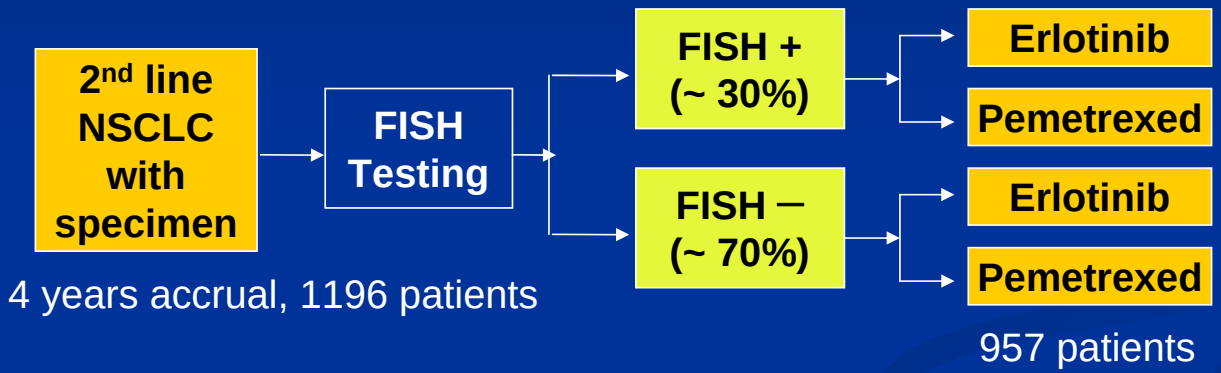
- Do not use the test to restrict eligibility, but to structure a prospective analysis plan
- Having a prospective analysis plan is essential
- "Stratifying" (balancing) the randomization is useful to ensure that all randomized patients have tissue available but is not a substitute for a prospective analysis plan
- The purpose of the study is to evaluate the new treatment overall and for the pre-defined subsets; not to modify or refine the classifier
- The purpose is not to demonstrate that repeating the classifier development process on independent data results in the same classifier

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

- R Simon. Using genomics in clinical trial design, Clinical Cancer Research 14:5984-93, 2008
- R Simon. Designs and adaptive analysis plans for pivotal clinical trials of therapeutics and companion diagnostics, Expert Opinion in Medical Diagnostics 2:721-29, 2008

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Validation of EGFR biomarkers for selection of EGFR-TK inhibitor therapy for previously treated NSCLC patients



Outcome

1° PFS
2° OS, ORR

1-2 years minimum additional follow-up

- PFS endpoint
  - 90% power to detect 50% PFS improvement in FISH+
  - 90% power to detect 30% PFS improvement in FISH–
- Evaluate EGFR IHC and mutations as predictive markers
- Evaluate the role of RAS mutation as a negative predictive marker

31

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Analysis Plan B
(Limited confidence in test)

- Compare the new drug to the control overall for all patients ignoring the classifier.
  - If $p_{overall} \leq 0.03$ claim effectiveness for the eligible population as a whole
- Otherwise perform a single subset analysis evaluating the new drug in the classifier + patients
  - If $p_{subset} \leq 0.02$ claim effectiveness for the classifier + patients.

32

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Analysis Plan C

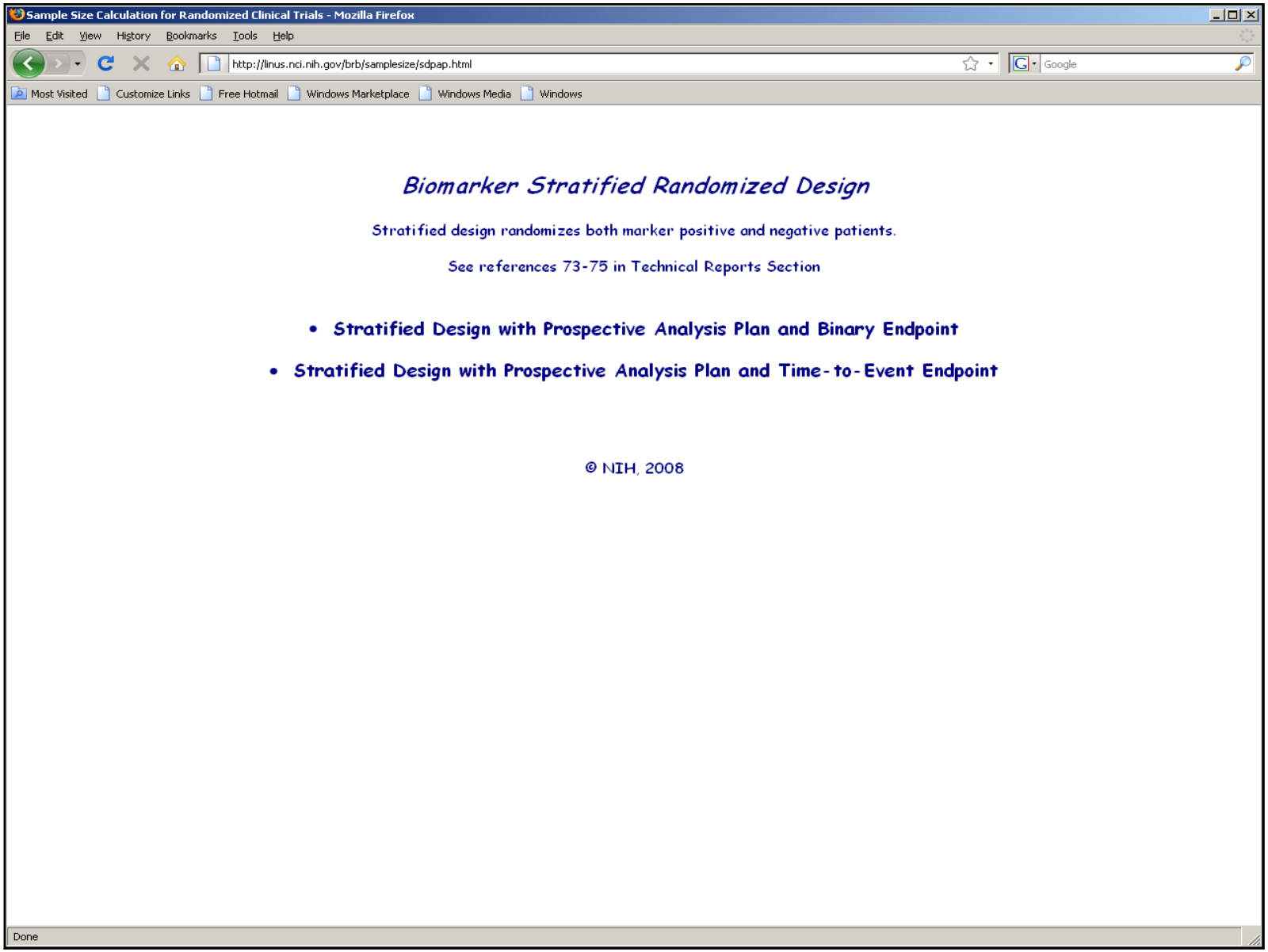
- Test for difference (interaction) between treatment effect in test positive patients and treatment effect in test negative patients
- If interaction is significant at level α_{int} then compare treatments separately for test positive patients and test negative patients
- Otherwise, compare treatments overall

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

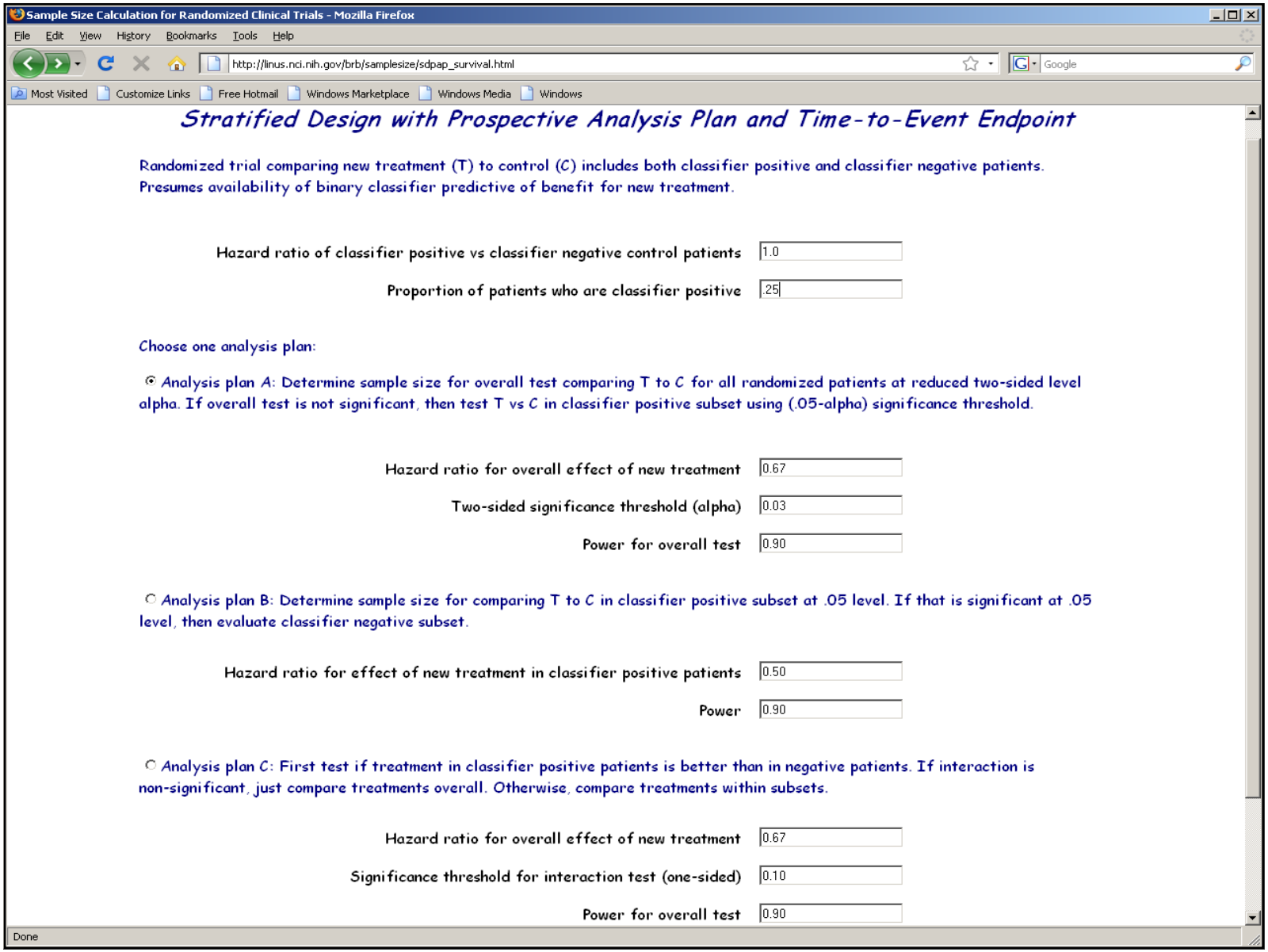
Sample Size Planning for Analysis Plan C

- 88 events in test + patients needed to detect 50% reduction in hazard at 5% two-sided significance level with 90% power
- If 25% of patients are positive, when there are 88 events in positive patients there will be about 264 events in negative patients
  - 264 events provides 90% power for detecting 33% reduction in hazard at 5% two-sided significance level

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute



Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute



Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

JNCI | JOURNAL NAME | djn022 | MS No. | HS | CE Code

ARTICLE

Biomarker-Adaptive Threshold Design: A Procedure for Evaluating Treatment With Possible Biomarker-Defined Subset Effect

Wenyu Jiang, Boris Freidlin, Richard Simon

Background Many molecularly targeted anticancer agents entering the definitive stage of clinical development benefit only a subset of treated patients. This may lead to missing effective agents by the traditional broad-eligibility randomized trials due to the dilution of the overall treatment effect. We propose a statistically rigorous biomarker-adaptive threshold phase III design for settings in which a putative biomarker to identify patients who are sensitive to the new agent is measured on a continuous or graded scale.

Methods The design combines a test for overall treatment effect in all randomly assigned patients with the establishment and validation of a cut point for a prespecified biomarker of the sensitive subpopulation. The performance of the biomarker-adaptive design, relative to a traditional design that ignores the biomarker, was evaluated in a simulation study. The biomarker-adaptive design was also used to analyze data from a prostate cancer trial.

Results In the simulation study, the biomarker-adaptive design preserved the power to detect the overall effect when the new treatment is broadly effective. When the proportion of sensitive patients as identified by the biomarker is low, the proposed design provided a substantial improvement in efficiency compared with the traditional trial design. Recommendations for sample size planning and implementation of the biomarker-adaptive design are provided.

Conclusions A statistically valid test for a biomarker-defined subset effect can be prospectively incorporated into a randomized phase III design without compromising the ability to detect an overall effect if the intervention is beneficial in a broad population.

J Natl Cancer Inst 2007;99:1–8

Human cancers are heterogeneous with regard to their molecular and genomic properties. Recent advances in biotechnology have resulted in a shift toward molecularly targeted anticancer agents. These new therapeutics are likely to benefit only a subset of the patients with a given cancer. Definitive testing of such targeted agents requires the identification of the appropriate "sensitive" population. When biomarkers to identify the patients who are likely to benefit from the new therapy are available, targeted clinical trials that restrict eligibility to sensitive patients should be used (1). However, reliable assays to identify sensitive patients are often unavailable. In the absence of a reliable biomarker, broad-eligibility clinical trials are used routinely. Most of these trials use a conventional design, in which the primary analysis is based on comparison of all randomly assigned patients. This often leads to the failure to recognize effective agents due to dilution of the treatment effect by the presence of the patients who do not benefit from the agent. Retrospective analysis of trials with a conventional design can be used as an initial step in identifying biomarkers for the sensitive subpopulation. However, retrospectively identified biomarkers typically have to be validated in a confirmatory prospective randomized phase III clinical trial (2). This approach is inefficient and may considerably prolong clinical development.

Previously, we have proposed a design [adaptive signature design (3)] that combines a definitive test for treatment effect in a broad population with identification and validation of a genomic signature for the subset of sensitive patients if the broad population test is negative. The adaptive signature design was developed for high-dimensional data such as gene expression microarrays, where only a few unknown genes among thousands assayed may be relevant and where a classifier (signature) to identify sensitive patients is not available. The design incorporates both the identification and the validation of a pharmacogenomic signature for sensitive patients.

Often, preliminary information on a biomarker to identify the sensitive subset of patients is available but an appropriate cutoff

Affiliation of authors: Biometric Research Branch, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, MD.

Correspondence to: Boris Freidlin, PhD, Biometric Research Branch, Division of Cancer Treatment and Diagnosis, EPN-8122, National Cancer Institute, Bethesda, MD 20892 (e-mail: freidlinb@ctep.nci.nih.gov).

See "Notes" following "References."

DOI: 10.1093/jnci/djm022

Adaptive Signature Design: An Adaptive Clinical Trial Design for Generating and Prospectively Testing A Gene Expression Signature for Sensitive Patients

Boris Freidlin and Richard Simon

Abstract **Purpose:** A new generation of molecularly targeted agents is entering the definitive stage of clinical evaluation. Many of these drugs benefit only a subset of treated patients and may be overlooked by the traditional, broad-eligibility approach to randomized clinical trials. Thus, there is a need for development of novel statistical methodology for rapid evaluation of these agents.
Experimental Design: We propose a new adaptive design for randomized clinical trials of targeted agents in settings where an assay or signature that identifies sensitive patients is not available at the outset of the study. The design combines prospective development of a gene expression–based classifier to select sensitive patients with a properly powered test for overall effect.
Results: Performance of the adaptive design, relative to the more traditional design, is evaluated in a simulation study. It is shown that when the proportion of patients sensitive to the new drug is low, the adaptive design substantially reduces the chance of false rejection of effective new treatments. When the new treatment is broadly effective, the adaptive design has power to detect the overall effect similar to the traditional design. Formulas are provided to determine the situations in which the new design is advantageous.
Conclusion: Development of a gene expression–based classifier to identify the subset of sensitive patients can be prospectively incorporated into a randomized phase III design without compromising the ability to detect an overall effect.

Developments in tumor biology have resulted in shift toward molecularly targeted drugs (1–3). Most human tumor types are heterogeneous with regard to molecular pathogenesis, genomic signatures, and phenotypic properties. As a result, only a subset of the patients with a given cancer is likely to benefit from a targeted agent (4). This complicates all stages of clinical development, especially randomized phase III trials (5, 6). In some cases, predictive assays that can accurately identify patients who are likely to benefit from the new therapy have been developed. Then, targeted randomized designs that restrict eligibility to patients with sensitive tumors should be used (7). However, reliable assays to select sensitive patients are often not available (8, 9). Consequently, traditional randomized clinical trails with broad eligibility criteria are routinely used to evaluate such agents. This is generally inefficient and may lead to missing effective agents.

Genomic technologies, such as microarrays and single nucleotide polymorphism genotyping, are powerful tools that hold a great potential for identifying patients who are likely to benefit from a targeted agent (10, 11). However, due to the large number of genes available for analysis, interpretation of these data is complicated. Separation of reliable evidence from the random patterns inherent in high-dimensional data requires specialized statistical methodology that is prospectively incorporated in the trial design. Practical implementation of such designs has been lagging. In particular, analysis of microarray data from phase III randomized studies is usually considered secondary to the primary overall comparison of all eligible patients. Many analyses are not explicitly written into protocols and done retrospectively, mainly as "hypothesis-generating" tools.

We propose a new adaptive design for randomized clinical trials of molecularly targeted agents in settings where an assay or signature that identifies sensitive patients is not available. Our approach includes three components: (*a*) a statistically valid identification, based on the first stage of the trial, of the subset of patients who are most likely to benefit from the new agent; (*b*) a properly powered test of overall treatment effect at the end of the trial using all randomized patients; and (*c*) a test of treatment effect for the subset identified in the first stage, but using only patients randomized in the remainder of the trial. The components are prospectively incorporated into a single phase III randomized clinical trial with the overall false-positive error rate controlled at a prespecified level.

Authors' Affiliation: Biometric Research Branch, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, Maryland
Received 3/18/05; revised 7/18/05; accepted 8/4/05.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
Requests for reprints: Boris Freidlin, Biometric Research Branch, Division of Cancer Treatment and Diagnosis, National Cancer Institute, 6130 Executive Boulevard, EPN 8122, MSC 7434, Bethesda, MD 20892-7434. Phone: 301-402-0640; Fax: 301-402-0560; E-mail: freidlinb@ctep.nci.nih.gov.
doi:10.1158/1078-0432.CCR-05-0606

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Prospective-Retrospective Evaluation of Prognostic or Predictive Classifier

1. Analytically validate a single completely specified classifier
2. Design a prospective clinical trial that definitvely addresses the hypothesis of interest about the medical utility of the completely specified classifier
    1. Write a detailed protocol for the prospective study, including sample size justification and detailed statistical analysis plan addressing a single hypothesis about the prognostic or predictive utility of a single completely specified classifier
3. Find a previously performed clinical trial that matches as closely as possible the prospective protocol developed above
    1. Adequate design
    2. Adequate sample size
    3. Adequate proportion of patients with archived tissue
    4. Not used in any way in developing the classifier or analytically validating it
4. Perform the assay on the archived samples and then analyze the data as defined in the prospective analysis plan

38

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Types of Validation for Prognostic and Predictive Biomarkers

- Analytical validation
  - Pre-analytical and post-analytical robustness
- Clinical validation
  - Does the biomarker predict what it's supposed to predict for independent data
- Clinical utility
  - Does use of the biomarker result in patient benefit

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Clinical Utility

- Benefits patient by improving treatment decisions
- Depends on context of use of the biomarker
  - Treatment options and practice guidelines
  - Other prognostic factors

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Clinical Utility of Prognostic Biomarker

- Prognostic biomarker for identifying patients
  - for whom practice standards imply cytotoxic chemotherapy
  - who have good prognosis without chemotherapy
- Prospective trial to identify such patients and withhold chemotherapy
    - TAILORx
- "Prospective plan" for analysis of archived specimens from previous clinical trial in which patients did not receive chemotherapy
  - OncotypeDx

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Clinical Utility of Predictive Biomarker

- Predictive biomarker for identifying the patients who benefit from a specific regimen and/or the patients who do not
- Prospective RCT of new regimen versus control with tissue prospectively collected and assayed and patients classified as test + or test –
  - Sample size established to have enough test + patients for separate analysis of new regimen versus control and enough test – patients for separate analysis of new regimen versus control
  - Focused analysis on a single completely prospectively defined biomarker classifier

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Use of Archived Specimens in Evaluation of Prognostic and Predictive Biomarkers

Richard M. Simon, Soonmyung Paik and Daniel F. Hayes

- We propose modified guidelines for the conduct of reliable analyses of prognostic and predictive biomarkers using archived specimens. These guidelines stipulate that:
- (i) archived tissue adequate for a successful assay must be available on a sufficiently large number of patients from a phase III trial that the appropriate analyses have adequate statistical power and that the patients included in the evaluation are clearly representative of the patients in the trial.
- (ii) The test should be analytically and pre-analytically validated for use with archived tissue.
- (iii) The analysis plan for the biomarker evaluation should be completely specified in writing prior to the performance of the biomarker assays on archived tissue and should be focused on evaluation of a single completely defined classifier.
- iv) the results from archived specimens should be validated using specimens from a similar, but separate, study.

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Use of Archived Specimens in Evaluation of Prognostic and Predictive Biomarkers

Richard M. Simon, Soonmyung Paik and Daniel F. Hayes

- Conclusions
- Claims of medical utility for prognostic and predictive biomarkers based on analysis of archived tissues can be considered to have either a high or low level of evidence depending on several key factors.
- These factors include the analytical and pre-analytical validation of the assay, the nature of the study from which the specimens were archived, the number and condition of the specimens, and the development prior to assaying tissue of a focused written plan for analysis of a completely specified biomarker classifier.
- Studies using archived tissues, when conducted under ideal conditions and independently confirmed can provide the highest level of evidence.
- Traditional analyses of prognostic or predictive factors, using non analytically validated assays on a convenience sample of tissues and conducted in an exploratory and unfocused manner provide a very low level of evidence for clinical utility.

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Conclusions

- New technology makes it increasingly feasible to identify which patients require systemic treatment and which are most likely to benefit from a specified regimen
- We are rapidly proceeding on the way to predictive oncology based on genomic characterization of a patient's tumor
- Rate limiting steps are
  - Identifying key oncogenic mutations
  - Access to tissue from patients in key clinical trials
  - Preforming the appropriate clinical trials

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Conclusions

- Targeting treatment can provide
  - Patient benefit
  - Economic benefit for society
  - Improved chance of success for new drug development
    - Not necessarily simpler or less expensive development
- Achieving the potential of new technology requires
  - Paradigm changes in focus and methods of "correlative science."
  - New approaches to trans-disciplinary training and collaboration
  - Effective collaboration between academic research and industry
  - Appropriate standards for regulation of in-vitro diagnostics

Richard Simon, D.Sc.
Biometric Research Branch
Division of Cancer Treatment and Diagnosis, National Cancer Institute

Acknowledgements

- NCI Biometric Research Branch
  - Boris Freidlin
  - Yingdong Zhao
  - Wenyu Jiang
  - Aboubakar Maitournam
- Soonmyung Paik, NSABP
- Daniel Hayes, U. Michigan