New Paradigms for Clinical Drug Development in the Genomic Era

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

“Biomarkers”

- Surrogate endpoints
  - A measurement made on a patient before, during and after treatment to determine whether the treatment is working
- Predictive classifier
  - A measurement made before treatment to predict whether a particular treatment is likely to be beneficial

Surrogate Endpoints

- It is extremely difficult to properly validate a biomarker as a surrogate for clinical outcome. It requires a series of randomized trials with both the candidate biomarker and clinical outcome measured
- Biomarkers for use in phase I/II studies need not be validated as surrogates for clinical outcome

Partial Surrogate Endpoint

- “Unvalidated” partial surrogates can be used for early termination of phase III trials. The trial should continue accrual and follow-up to evaluate true endpoint if treatment effect on partial surrogate is sufficient.

Predictive Biomarkers

- Most 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
  - Improve the efficiency of clinical development
  - Help control medical costs

Oncology Needs Predictive Markers not Prognostic Factors

- Most prognostic factors are not used because they are not therapeutically relevant
- Most prognostic factor studies use a convenience sample of patients for whom tissue is available. Generally the patients are too heterogeneous to support therapeutically relevant conclusions

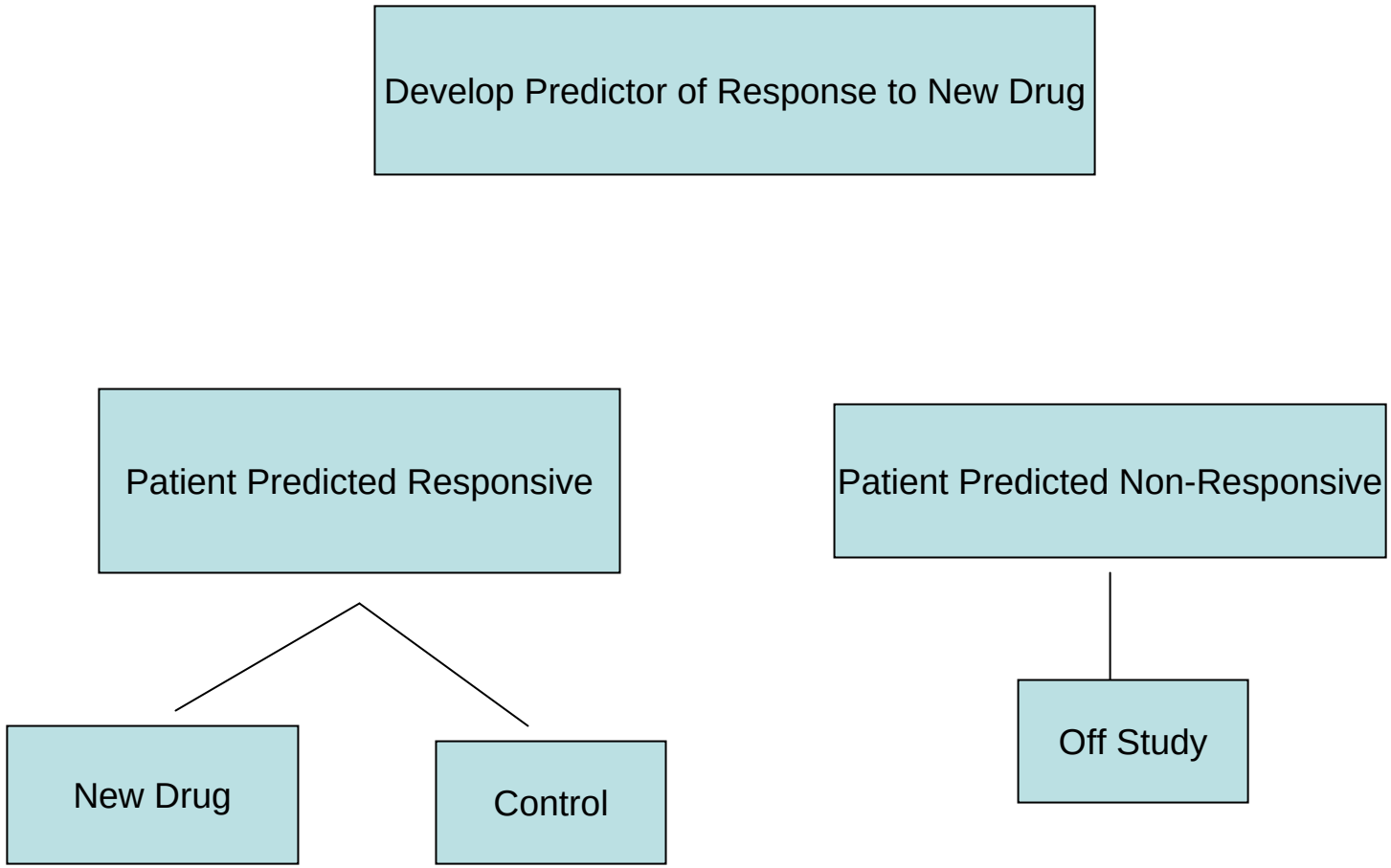
- Criteria for validation of surrogate endpoints should not be applied to biomarkers used in treatment selection

- Targeted clinical trials can be much more efficient than untargeted clinical trials, if we know who to target

- In new drug development, the role of a classifier is to select a target population for treatment
  - The focus should be on evaluating the new drug in a population defined by a predictive classifier, not on “validating” the classifier

Developmental Strategy (I)

- **Develop** a diagnostic classifier that identifies the patients likely to benefit from the new drug
- Develop a reproducible assay for the classifier
- **Use** the diagnostic to restrict eligibility to a prospectively planned evaluation of the new drug
- Demonstrate that the new drug is effective in the prospectively defined set of patients determined by the diagnostic


Develop Predictor of Response to New Drug
Patient Predicted Responsive
Patient Predicted Non-Responsive
New Drug
Control
Off Study

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.
- Maitnourim A and Simon R. On the efficiency of targeted clinical trials. Statistics in Medicine 24:329-339, 2005.
- reprints and interactive sample size calculations at http://linus.nci.nih.gov/brb

You Can Evaluate How This Design Might Work For You

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

One Should Require That

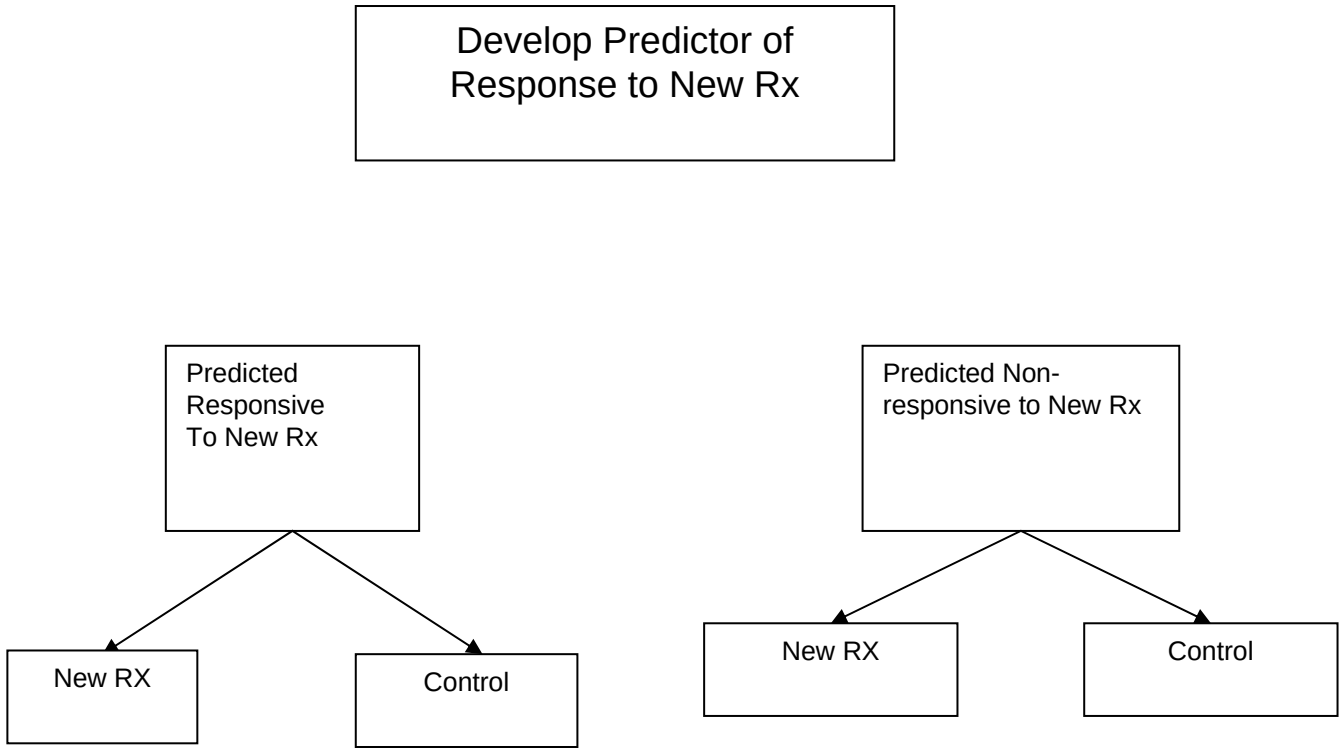
- The classifier, as a whole, be reproducibly measurable

- The classifier identifies a patient population for which the new drug has clinical utility

There Should Be No Requirement For

- Demonstrating that the classifier or any of its components are “validated biomarkers of disease status”
- Demonstrating that repeating the classifier development process on independent data results in the selection of the same components (genes)

Developmental Strategy (II)



Developmental Strategy (II)

- Do not use the diagnostic to restrict eligibility, but to structure a prospective analysis plan.
- Compare the new drug to the control overall for all patients ignoring the classifier.
  - If $p_{overall} \leq 0.04$ claim effectiveness for the eligible population as a whole
- Otherwise perform a single subset analysis evaluating the new drug in the classifier + patients
  - If $p_{subset} \leq 0.01$ claim effectiveness for the classifier + patients.

Key Features of Design (II)

- Pre-specified analysis plan
- Single pre-defined subset
- Overall study type I error of 0.05 is split between overall test and subset test
- Saying that the study should be “stratified” is not sufficient

Key Features of Design (II)

- The purpose of the RCT is to evaluate treatment T vs C overall and for the pre-defined subset; not to re-evaluate the components of the classifier, or to modify or refine the classifier

Guiding Principle

- The data used to develop the classifier must be distinct from the data used to test hypotheses about treatment effect in subsets determined by the classifier
  - Developmental studies are 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

Development of Genomic Classifiers

- Single gene or protein based on knowledge of therapeutic target
- Single gene or protein culled from set of candidate genes identified based on imperfect knowledge of therapeutic target
- Empirically determined based on correlating gene expression to patient outcome after treatment

Development of Genomic Classifiers

- During phase II development or
- After failed phase III trial using archived specimens.
- Adaptively during early portion of phase III trial.

Adaptive Signature Design

An adaptive design for generating and prospectively testing a gene expression signature for sensitive patients

Boris Freidlin and Richard Simon

Clinical Cancer Research 11:7872-8, 2005

Adaptive Signature Design
End of Trial Analysis

- Compare E to C for **all patients** at significance level 0.04
  - If overall H_0 is rejected, then claim effectiveness of E for eligible patients
  - Otherwise

- Otherwise:
  - Using only the first half of patients accrued during the trial, develop a binary classifier that predicts the subset of patients most likely to benefit from the new treatment E compared to control C
  - Compare E to C for patients accrued in second stage who are predicted responsive to E based on classifier
    - Perform test at significance level 0.01
    - If H_0 is rejected, claim effectiveness of E for subset defined by classifier

Simon R. Using DNA microarrays for diagnostic and prognostic prediction. Expert Review of Molecular Diagnostics, 3(5) 587-595, 2003.

Simon R. Diagnostic and prognostic prediction using gene expression profiles in high dimensional microarray data. British Journal of Cancer 89:1599-1604, 2003.

Simon R and Maitnourim A. Evaluating the efficiency of targeted designs for randomized clinical trials. Clinical Cancer Research 10:6759-63, 2004.

Maitnourim A and Simon R. On the efficiency of targeted clinical trials. Statistics in Medicine 24:329-339, 2005.

Simon R. When is a genomic classifier ready for prime time? Nature Clinical Practice – Oncology 1:4-5, 2004.

Simon R. An agenda for Clinical Trials: clinical trials in the genomic era. Clinical Trials 1:468-470, 2004.

Simon R. Development and Validation of Therapeutically Relevant Multi-gene Biomarker Classifiers. Journal of the National Cancer Institute 97:866-867, 2005.

Simon R. A roadmap for developing and validating therapeutically relevant genomic classifiers. Journal of Clinical Oncology 23(29), 2005.

Freidlin B and Simon R. Adaptive signature design. Clinical Cancer Research 11:7872-8, 2005.

Simon R. Guidelines for the design of clinical studies for development and validation of therapeutically relevant biomarkers and biomarker classification systems. In Biomarkers in Breast Cancer, Hayes DF and Gasparini G, Humana Press, pp 3-15, 2005.

Simon R and Wang SJ. Use of genomic signatures in therapeutics development in oncology and other diseases. The Pharmacogenomics Journal, 2006.