Re-Examination of Study Designs for Early Clinical Trials of Molecularly Targeted Drugs

Richard Simon, D.Sc.
National Cancer Institute
linus.nci.nih.gov/brb
I have no financial relationships to disclose.

I will not discuss off label use and/or investigational use in my presentation.
Objectives of Initial Trials

• Find safe dose at which target pathway is inhibited
Conventional Phase I Approach

• Many current targets are not specific to cancer cells and hence targeted drugs are toxic
• Few examples of drugs whose effectiveness at inhibiting target decreases with dose after maximum
• Titrating dose for maximum inhibition of target is difficult due to assay variability and need for tumor biopsies
• Conventional phase I trial to establish dose just below MTD which can be delivered repeatedly

• Accrue an additional cohort of patients at that selected dose to determine whether the target is inhibited
Traditional Phase II Trials

• Estimate the proportion of tumors that shrink by 50% or more when the drug is administered singly or in combination to patients with advanced stage tumors of a specific primary site
Problem With Traditional Approach

• Phase II single agent responses do not predict well for phase III success of combination regimens

• Some drugs that have effectiveness in phase III did not produce many responses in single agent phase II trials
Possible Reasons for Inadequacy of Traditional Phase II Trials

- Drugs active as single agents may not contribute to activity of combinations
- Partial response may not represent sufficient anti-tumor effect to prolong survival
- Substantial anti-tumor effect for minority of patients is so diluted in broad phase III trials that it would take huge sample sizes to have adequate statistical power
Design Approaches for Phase II Evaluation of New Drug in Combination with PFS Endpoint

- Single arm trial of standard + new drug using historical controls
- Randomized phase 2.5 design
- Randomized discontinuation design
- Seamless phase II/III design
Single arm trial of standard + new drug using historical controls

- Often un-interpretable
  - Inherent limitations
  - Poor execution
- Requires selecting specific controls matched for prognostic factors
- Meta-analysis of previous phase II trials demonstrating sufficiency of matching criteria
Randomized Phase 2.5 Design

• Standard regimen ± new drug

• Differs from Phase III Design
  – Significance level 10% 1-sided
  – PFS endpoint not necessarily accepted as reflecting clinical benefit
Phase 2.5 Trial Design


Total Sample Size  
Randomized Phase 2.5  
2 years accrual, 1.5 years follow-up

<table>
<thead>
<tr>
<th>Improvement in median PFS</th>
<th>Hazard Ratio</th>
<th>$\alpha=.05$</th>
<th>$\alpha=.10$</th>
<th>$\alpha=.20$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 $\rightarrow$ 6 months</td>
<td>1.5</td>
<td>216</td>
<td>168</td>
<td>116</td>
</tr>
<tr>
<td>4 $\rightarrow$ 8 months</td>
<td>2.0</td>
<td>76</td>
<td>60</td>
<td>40</td>
</tr>
</tbody>
</table>
Randomized Discontinuation Design (RDD)


Randomized Discontinuation Design (RDD)

• The RDD requires a large sample size
• The RDD is not a phase III design because it does not establish the clinical utility of administering the drug to the patient compared to not administering it
Seamless Phase II/III Trial

• Randomized comparison of standard treatment ± new drug
• Size trial using phase III (e.g. survival) endpoint
• Perform interim futility analysis using phase II endpoint (e.g. biomarker or PFS)
  – If treatment vs control results are not significant for phase II endpoint, terminate accrual and do not claim any benefit of new treatment
  – If results are significant for intermediate endpoint, continue accrual and follow-up and do analysis of phase III endpoint at end of trial
• Interim analysis does not “consume” any of the significance level for the trial
New Objective for Developmental Studies

• It is important to better characterize in phase II studies which tumors are most likely to be sensitive to the drug

• Conducting a phase III trial of a molecularly targeted agent in the traditional way is likely to result in a false negative trial
  – Unless a sufficiently large proportion of the patients have tumors driven by the targeted pathway
New Drug 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
Using phase II data, develop a predictor of response to a new drug.
Evaluating the Efficiency of Strategy (I)


- reprints and interactive sample size calculations at http://linus.nci.nih.gov/brb
Efficiency of Targeted Design Depends On

- **Treatment specificity**
  - $\delta_1 =$ treatment effect for Target + patients
  - $\delta_0 =$ treatment effect for Target - patients

- **Assay performance**
  - Sensitivity = $\text{Prob}\{\text{Assay}+ \mid \text{Target} +\}$
  - Specificity = $\text{Prob}\{\text{Assay}- \mid \text{Target} -\}$

- **Prevalence of target + patients**
Randomized Ratio

# randomized: standard design / targeted design
sensitivity=specificity=0.9

<table>
<thead>
<tr>
<th>Proportion Expressing Target</th>
<th>$\delta_0=0$</th>
<th>$\delta_0=\delta_1/2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.75</td>
<td>1.29</td>
<td>1.26</td>
</tr>
<tr>
<td>0.5</td>
<td>1.8</td>
<td>1.6</td>
</tr>
<tr>
<td>0.25</td>
<td>3.0</td>
<td>1.96</td>
</tr>
<tr>
<td>0.1</td>
<td>25.0</td>
<td>1.86</td>
</tr>
</tbody>
</table>
## Screened Ratio

*# screened standard design / targeted design*

sensitivity=specificity=0.9

<table>
<thead>
<tr>
<th>Proportion Expressing Target</th>
<th>$\delta_0=0$</th>
<th>$\delta_0=\delta_1/2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.75</td>
<td>0.9</td>
<td>0.88</td>
</tr>
<tr>
<td>0.5</td>
<td>0.9</td>
<td>0.80</td>
</tr>
<tr>
<td>0.25</td>
<td>0.9</td>
<td>0.59</td>
</tr>
<tr>
<td>0.1</td>
<td>4.5</td>
<td>0.33</td>
</tr>
</tbody>
</table>
Trastuzumab

- 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% assay + patients, overall improvement in survival would have been 3.375%
  - 4025 patients/arm would have been required
- If assay – patients benefited half as much, 627 patients per arm would have been required
Gefitinib

- Two negative untargeted randomized trials first line advanced NSCLC
  - 2130 patients
- 10% have EGFR mutations
- If only mutation + patients benefit by 20% increase of 1-year survival, then 12,806 patients/arm are needed
- For trial targeted to patients with mutations, 138 are needed
Web Based Software for Comparing Sample Size Requirements

Research Areas

- Clinical trials
- Drug Discovery
- Molecular Cancer Diagnosis
- Biomedical Imaging
- Computational and Systems Biology
- Biostatistical Research

Technical Reports and Talks

- Download the PDF version of our most recent research papers and PowerPoint version of the talk slides

BRB Staff

- Investigators and contact information

BRB Alumni

BRB Annual Report 2005

Position Available

- Post-doctoral fellow positions available

Sample Size Calculation

Mathematics and Oncology

- The Norton-Simon Hypothesis
- The Norton-Simon Hypothesis and Breast Cancer Mortality in a National Randomized Trial

Software Download

- Accelerated Titration Design Software
- Optimal Two-Stage Phase II Design Software
Sample Size Calculation for Randomized Clinical Trials

- Optimal Two-Stage Phase II Design

- Biomarker Targeted Randomized Design*
  1. Binary Outcome Endpoint
  2. Survival and Time-to-Event Endpoint

* Targeted design randomizes only marker positive patients to treatment or control arm. Untargeted design does not measure marker and randomizes all who otherwise are eligible.

© NIH, 2006
Sample Size Calculation: Binary Outcome Endpoint


pc = probability of "response" for control arm

\[ \text{gamma} \]

\[ \text{delta1} \]

\[ \text{delta0} \]

\[ \alpha = 0.05 \]

\[ \text{power} = 0.90 \]

Submit

pc = probability of "response" for control arm

\[ \text{gamma} = \text{proportion of patients who are classifier negative (i.e. less responsive to new treatment)} \]

\[ \text{delta1} = \text{improvement in response probability for new treatment in classifier positive patients} \]

\[ \text{delta0} = \text{improvement in response probability for new treatment in classifier negative patients} \]

\[ \alpha = \text{two-sided significance level} \]

© NIH, 2006
Predictive Medicine not Correlative Science

• The purpose of the RCT is to evaluate the new treatment for the pre-defined subset
• The purpose is not to modify or refine the classifier
The Roadmap

1. Develop a completely specified genomic classifier of the patients likely to benefit from a new drug
2. Establish reproducibility of measurement 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.
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(s)
  - HER2 amplification
  - EGFR mutation or amplification
- Empirically determined based on correlating gene expression to patient response
  - Genome-wide
  - Candidate genes
Development of Empirical Gene Expression Based Classifier

• 20-30 phase II responders are needed to compare to non-responders in order to develop signature for predicting response
Adaptive Signature Design: An Adaptive Clinical Trial Design for Generating and Prospectively Testing A Gene Expression Signature for Sensitive Patients

Borns 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 identify sensitive patients with an already 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 trials 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.
Collaborators

• Kevin Dobbin
• Boris Freidlin
• Sally Hunsberger
• Wenyu Jiang
• Aboubakar Maitournam
• Yingdong Zhao
Using Genomic Classifiers In Clinical Trials


Using Genomic Classifiers In Clinical Trials


