Statistical Issues in Evaluating the Efficacy of Personalized Medicine

Motomi (Tomi) Mori, PhD
Walter & Clora Brownfield Professor of Cancer Biostatistics
Knight Cancer Institute
Professor and Head, Division of Biostatistics
Department of Public Health & Preventive Medicine
Oregon Health & Science University
Outline

• Motivating Project
• Review of Clinical Trial Designs
• Proposed Designs
• Discussion
Motivating Project
Motivating Project

- Leukemia Lymphoma Society Specialized Center of Research (LLS SCOR) Grant “Novel Tyrosine Kinase Targets in Leukemia and Myeloproliferative Neoplasms” (PI: Brian Druker, MD)

- The major goal is to identify novel kinase targets for the treatment of leukemia and myeloproliferative neoplasms.

- This grant consists of three disease-specific projects and six cores, including RNAi, Kinase Inhibitor Screen, Bioinformatics, Biostatistics, and Clinical Trials.

- The Biostatistics Core is responsible for design and analysis of clinical trials and molecular correlative studies.
Kinase Inhibitor Screen

• It is used to identify kinase inhibitors to which patient primary leukemia sample is sensitive.
  - The best kinase inhibitor is the one that kills most leukemia cells at the lowest concentration.

• It is a high-throughput assay.
  - 384-well plate format
  - 90 small-molecule inhibitor drugs that are FDA approved or in clinical trials
  - 8 different concentrations of each drug including the expected IC50 concentration
  - 3 technical replicates
  - Positive and negative control wells
Kinase Inhibitor Screen (cont.)

• Patient leukemia sample are tested against 90 kinase inhibitor drugs at 8 different drug concentrations.
• Dose-response curve is generated with the response being % of live leukemia cells at each concentration.
• IC50 ("half maximal inhibitory concentration") is a dose that kills 50% of leukemia cells and is computed from the dose-response curve.
• % of the median IC50 is computed by dividing the observed IC50 by the median IC50 of the drug based on all the past samples.
• A drug with the lowest % of the median IC50 is considered as the most sensitive drug, a drug target for the patient.
• In practice, the algorithm is more complex and works in conjunction with RNAi screen (genetic mutation screen).
Kinase Inhibitor Screen
Percent of Median IC50s for Sample 08-024
Research Question

• Does the treatment assignment based on the kinase inhibitor screen improve patient clinical outcome?
• How do we design a clinical trial to obtain the preliminary efficacy of this approach?
Review of Clinical Trial Designs
Drug Development Pipeline

Preclinical Development ➔ Phase I Trial MTD and Safety ➔ Phase II Trial Anti-Tumor Activity ➔ Phase III Trial Comparison with Standard Treatment

Phase IV Trial Post Marketing Surveillance ➔ Licensing
Typical Clinical Trial Designs

• **Phase I trials**
  - Focus on toxicity and safety
  - Determination of the Maximally Tolerated Dose (MTD)
  - Typical designs: classical 3+3 design, accelerated titration design, continual reassessment design

• **Phase II trial**
  - Demonstration of anti-tumor activity
  - Often single-arm trial with short-term clinical endpoint (e.g., tumor response)
  - Typical designs: fixed sample size single arm trial, Simon’s 2-stage design

• **Phase III trial**
  - Randomized controlled trial with long-term clinical endpoint (e.g., disease free survival)
Newer Clinical Trial Designs

• Phase 0 trials
  o First-in-human to assess drug effect on a molecular target, by means of a pharmacodynamic assay in a very small number of patients (i.e., 10-15).

• Phase I trials
  o Designs to determine dose based on both safety and biological endpoints
    • Trinomial Continual Assessment Method (triCRM)
  o Dose escalation designs based on molecular target effects (biological endpoint rather than toxicity)

• Phase II trials
  o Randomized phase II trials
  o Selection designs
  o Bayesian adaptive phase II trials

• Combining phases
  o Phase 0/I trial, phase I/II trial, phase II/III trial
Genomic and Biomarker Guided Clinical Trials

- Enrichment design
- Marker-by-treatment interaction design
- Marker-based strategy design
- Adaptive designs
- Hybrid designs
Enrichment Design

Example: Trastuzumab in HER2 positive breast cancer

Marker-by-Treatment Interaction Design

Example: MARVEL (Marker Validation in Erlotinib in Lung Cancer) for NSCLC.
Example of Marker-by-Treatment Interaction Design
MARVEL (Marker Validation of Erlotinib in Lung Cancer), second-line therapy for NSCLC patients (Mandrekar and Sargent, JCO, 2009)
Marker-Based Strategy Design

Randomization

Use Test

Do Not Use Test

Marker +

Marker -

Marker +

Marker -

Treatment

Control

Randomization

Treatment

Control
Example of Marker-Based Strategy Design. MINDACT (Microarray in Node-Negative Disease May Avoid Chemotherapy) trial for patients with node-negative breast cancer designed to evaluate MammaPrint, the 70-gene expression profile discovered at the Netherlands Cancer Institute (Mandrekar and Sargent, JCO, 2009).
Example of Adaptive Design. I-SPY 2 (Investigation of Serial Studies to Predict Your therapeutic response with imaging and molecular analysis 2) to compare the efficacy of novel drugs in combination with standard chemotherapy vs. the standard therapy alone. (Barker AD et al., Nature Clinical Pharmacology & Therapeutics 2009)
Example of Adaptive Design

BATTLE Trial

- BATTLE (Biomarker-integrated Approaches of Targeted Therapy of Lung cancer Elimination) trial is a response adaptive randomization trial for patients with advanced stage non-small cell lung cancer.
- The BATTLE trial had an initial equal randomization stage followed by an adaptive randomization stage where patients were adaptively randomized to one of four treatments based on molecular biomarkers found in biopsy.
- They showed an impressive benefit from sorafenib among mutant-KRAS patients.
Comments

• Currently available genomic/biomarker guided trial designs are intended for testing a limited number of treatments with a limited number of markers.
• But we have several markers (targets identified by the assay) and several treatments (drugs).
• What we want to test is a strategy, not individual marker-drug combinations.
Proposed Designs
Proposed Phase II Designs

• Fully randomized trial designs
  o Assay Guided Strategy Design
  o Assay Adaptive Randomization Design
  o Two-Stage Bayesian Response-Adaptive Randomization Design

• Partially or non-randomized trial designs
Assay-Guided Strategy Design

Pre-Registration

Kinase-inhibitor and RNAi screens

Randomization

Screen Failure (~25%)

Non-assay guided

Assay guided

Randomization

A  B  C  D  E

A  B  C  D  E
Assay-Guided Strategy Design

- First randomize patients to an assay guided group (experiment) vs. non-assay guided group (control) in a 1:1 ratio.
- Assay guided group: drug assignment according to the assay result.
- Non-assay guided group: randomize patients to one of the five drug groups.
  - Randomization ratio
    - Equal probability (i.e., .20 for each drug group)
    - Expected frequency of the assay results based on the historical data
    - Adaptive randomization where the randomization ratio can change according to the overall frequency of the assay results observed during the trial.
Example Data

Assay-guided group (n=39)

<table>
<thead>
<tr>
<th>Drug/Assay Results</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Non-assay guided group (n=41)

<table>
<thead>
<tr>
<th>Drug/Assay Result</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>8</td>
<td>7</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>9</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Assumptions:
1) Expected frequency of the assay results: p(A)=.50; p(B)=.30; p(C)=.10; p(D)=.05; p(E)=.05.
2) Randomization ratio in the non-assay guided group based on the expected frequency of the assay results.
### Assay-Matched vs. Mismatched Comparison

<table>
<thead>
<tr>
<th>Drug/Assay Results</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>X</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>B</td>
<td>O</td>
<td>X</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>C</td>
<td>O</td>
<td>O</td>
<td>X</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>D</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>X</td>
<td>O</td>
</tr>
<tr>
<td>E</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>X</td>
</tr>
</tbody>
</table>

Comparison of response rate between those in diagonal vs. off-diagonal.
Most of the patients in diagonal entries are in the assay-guided group, but some are in the non-assay guided group.
Assay-Guided Strategy Design Comparisons

• Assay-guided vs. non-assay guided comparison
  o Testing whether the assay-based drug assignment strategy is better than a random drug assignment

• Assay-matched vs. mismatched comparison
  o Testing whether a drug assignment consistent with the kinase-inhibitor assay is better than a drug assignment inconsistent with the kinase-inhibitor assay
Covariate adaptive randomization according to the assay result. For example, a patient is randomized to the drug group consistent with the assay result with the probability of .50 and other drugs with the probability of .125 each.
**Randomized Phase II Trial**

**Assay Adaptive Design**

Inhibitor Assay & RNAi Screening

- Matched Drug
- Mismatched Drug

Randomization probability of drug given the assay results

<table>
<thead>
<tr>
<th>Drug/Assay Results</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.5</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
</tr>
<tr>
<td>B</td>
<td>0.125</td>
<td>0.5</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
</tr>
<tr>
<td>C</td>
<td>0.125</td>
<td>0.125</td>
<td>0.5</td>
<td>0.125</td>
<td>0.125</td>
</tr>
<tr>
<td>D</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
<td>0.5</td>
<td>0.125</td>
</tr>
<tr>
<td>E</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
<td>0.5</td>
</tr>
</tbody>
</table>
NOTE: One-sided, alpha = .10; 20 % response rate for those receiving a drug inconsistent with the assay vs. 40% response rate for those receiving a drug consistent with the assay.
Comments

• For the assay-guided strategy design,
  - A comparison of assay guided arm vs. non-assay guided arm is inefficient.
  - A comparison of matched vs. mismatched groups is efficient, but two groups may not be comparable.

• For the assay adaptive design,
  - If the randomization ratio is 1:1 for matched vs. mismatched drug groups, it is most efficient.
  - It may be non-ethical to assign patients to a treatment contra-indicated by the assay.

• In reality, clinical response may be different for those with different kinase inhibitor results as well as for each drug.

• Excluding those without a drug target may create selection bias. We may want to include them and treat them as a separate control group.
Two-Stage Bayesian Response Adaptive Randomization Design

Stage 1: all patients are randomized equally to sensitive and insensitive drugs.
Stage 2: patients are randomized according to the randomization ratios based on the observed response rates in Stage 1.
Two-Stage Bayesian Response Adaptive Randomization Design

- We evaluated six different Stage 2 randomization allocation ratios through simulation and identified the optimum allocation ratio based on the response rates observed in Stage 1.
- The final proposed design has a reasonable power when the total sample size is 150 with the Stage 1 sample size of 30.
- The design is robust with respect to multiple sensitive drugs a patient may have and randomizes patients to more promising drugs (i.e., promising based on the response rate observed in Stage 1).
Challenges

• Unlikely to be able to randomize patients to non-target drugs
• Substantial % of patients without a target or target drug available
Alternative Design 1: no randomization; inclusion of patients without a target or available target drug

<table>
<thead>
<tr>
<th>Assay Result</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>SOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>No Target</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OO</td>
</tr>
</tbody>
</table>
Alternative Design 2: Randomization of a target drug vs. SOC; inclusion of patients without a target or available target drug

Drug Assignment

<table>
<thead>
<tr>
<th>Assay Result</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>SOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>O</td>
</tr>
<tr>
<td>No Target</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OO</td>
</tr>
</tbody>
</table>
Alternative Design 3: Randomization of target drug vs. SOC in both patient with and without target identified

Drug Assignment

<table>
<thead>
<tr>
<th>Assay Result</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>SOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>O</td>
</tr>
<tr>
<td>No Target</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>OO</td>
</tr>
</tbody>
</table>
Sample Size and Power

Assumption: Response rate of marker +/SOC vs. marker +/target drug: 20% vs. 40%, 80% power, 1-sided 10% significance level

Comparison of target drug vs. SOC within marker + patients.
1:1 randomization: Total N=114 (57 per group)
1:2 randomization: Total N=125 (42 in the marker +/SOC group and 83 in the marker +/target drug group)
Discussion
Summary

• We proposed several partially and fully randomized phase II trial designs to evaluate clinical efficacy of the assay-guided treatment strategy.

• A key is to focus on evaluating a strategy, not individual marker-drug combinations.

• The most efficient design is to randomly assign a patient to a matched vs. one of the unmatched drugs in a 1:1 ratio.

• The basic design can be extended to a two-stage response adaptive design where the randomization ratios change in the second stage, depending on the response rates observed in the first stage.
Future Direction

• Clinical validation of assays
  o FDA/IDE application required for an investigational diagnostic assay
    • IVDMIA (In Vitro Diagnostic Multivariate Index Assays)

• Improve the algorithm and QA/QC assessment of the kinase inhibitor and RNAi assays
  o Assessment of goodness-of-fit and technical variability

• Incorporate actual scores in treatment assignment probabilities and clinical outcome assessment
Current Status
Proof-of-Concept Phase II Trial

• Design: single arm trial
• Eligibility: adult patients with relapsed or refractory acute leukemias (ALL and AML)
• Primary endpoint: ≥ 25% decrease in bone marrow blast count at day 28
• Targeted drugs: dasatinib, sunitinib, soratinib, nilotinib, ponatinib (AP24534)
• Treatment: 28 days
• Sample size: 24 patients based on the clinical response rate of 22% vs. 5% (the standard of care)
Issues – Clinical Trials

- Patient eligibility
  - Previously treated, relapsed patients
  - Previously untreated, newly diagnosed patients

- Primary and secondary endpoints
  - Complete hematological response at 6 month?
  - Molecular endpoints?

- Drug availability and dose

- Single arm or randomized phase II?
  - Single arm
    - What should be the historical comparison group?
  - Randomized phase II
    - What should be the control arm?
    - What is the usual or standard of care?

- Accrual rate, sample size and power
Issues - Assays

- Kinase inhibitor screen results
  - May not identify a drug target
  - May identify a drug target, but may not be available (e.g., not FDA approved for leukemias, out of the approved dose range, etc)
  - May have multiple drug targets
  - The score distribution for drug targets may be low or flat, showing ambiguity of true targets
Challenges and Opportunities

- More than 800 anti-cancer agents are in active clinical development each year.
- The failure rate of these drugs is reported to be 95%.
- Typical R&D cost is over 1 billion dollars per approved drug [DiMasi and Grabowski, J CO 25(2) 2007].
- The current drug development pipeline is intended for cytotoxic drugs.
  - Higher the dose, better the anti-tumor activities. However, the dose needs to be limited due to safety concerns.
  - One drug and one dose work for all.
- Emergence of molecularly targeted drugs and availability of high throughput technologies force us to change the traditional drug development pipeline.
New Technologies and Perspectives

• Rapid development of high-throughput molecular profiling technologies
  - DNA copy number aberrations, DNA mutation detection, epigenetic profiling, gene expression profiling, detection of splicing RNA forms, protein arrays

• Systems biology approach to identify molecular drivers and biomarkers

• Personalized medicine
  - 5Rs - “Right drug at the right dose for the right indication to the right patient at the right time” [Gonzalez-Angulo et al, JCO 28(18) 2010]
Personalized Medicine: Beyond Therapeutic Trials

- Personalized medicine approach can be applied to cancer control and prevention research
- Similar challenges exist in testing individualized approach to cancer control and prevention strategies
- Example
  - Survey instrument to measure psychological constructs and dimensions (e.g., based on the theory of planned behaviors)
  - Individualized intervention may be developed according to the psychological profile revealed by the survey
  - How do we test individualized approach vs. standard approach? How do we know that a receipt of any individualized approach is responsible for the better outcome?
Acknowledgment

• Collaborators
  o Biostatistics Shared Resource, Knight Cancer Institute
    • Yiyi Chen; Byung S. Park; Solange Mongoue-Tchokote
  o Center for Hematological Malignancies
    • Brian Druker; Jeff Tyner; Marc Loriaux
    • Stephen Spurgen; Bill Chang
  o Collaborators at other institutions
    • Mike Deininger, Tibor Kovacssovic (U of Utah)
    • Markus Muschen (U of Southern California)

• Grant Support
  o Leukemia Lymphoma Society, Specialized Center of Research
  o National Cancer Institute, Cancer Center Support Grant