Some experience with biomarker driven cancer clinical trials

Michael LeBlanc

Fred Hutchinson Cancer Research Center
SWOG Statistical Center
Outline

• Statistical Considerations (prior talks)
  • Impact of treatment and biomarker(s) on patient outcome (predictive and prognostic associations)
  • Impact of design choices on inference

• Experience
  • S9704 Prognostic Targeting
  • S1406 Single mutation (or subgroup) targeting
  • S1400 Multiple sub-group targeting
Traditional divisions of treatments by types of cancer

- Sites: Breast, Lung, Gastrointestinal, Genitourinary, Melanoma, Leukemia, Lymphoma, Myeloma, Sarcoma
- Traditional trials in sub-sites, histologies, early stage, advanced stages relapsed disease
- But increasingly disease is characterized molecularly into much finer divisions
Variation in efficacy

- Genetic or protein measurement (designing statistical interactions)
  - HER2 amplification [Herceptin]
  - EGFR mutation [Erlotinib]
  - tyrosine kinase enzyme (c-kit) [Imatinib]
  - BRAF mutation [Vemurafenib]
- Multi-variable genetics predicting treatment efficacy
  - OncotypeDx recurrence score (breast cancer)
  - Other Tumor genomics
Stages of treatment testing (learning)

- **Phase I**
  - The safe dose range, side effects, early activity.

- **Phase II**
  - Sufficient promise for further testing, more side effect assessment, refinement of dose, evidence of disease subtypes with most promise and feasibility.
  - Some design examples: single arm 2-stage, single arm pilot, multi-arm randomized (screening or selection).

- **Phase III**
  - Formal comparison of new treatment to “standard”.

Modeling
Outcome Associations in Trials: Choosing Target Design

- Biomarker - Treatment Interaction Model
  - Two cases:
    - 1) Treatment is essentially equally effective regardless of gene
    - 2) The expression indicates where one treatment is preferred

![Graph showing treatment outcomes and biomarker expression.](image-url)
General Case: Discrete Subgroup Models

For designing treatment trials, summaries based on a subgroup of patients are often useful.

At least 3 components are of interest:

1. Rules to describe a subgroup of patients, \( R \).
2. A model for treatment effect in that group
3. The mass (or the fraction of all patients in that group)

The triple describes future design properties

- Example of subgroup models

\[
M(R) = \theta(Z \mid X \in R) = \alpha(R) + \beta(R)Z
\]

- Eligibility

- Fraction of patients

- Main effect

- Treatment effect
Model Class 1: Targeted Design

Advantages: If treatment is only effective in a subgroup this is powerful. However, if there is broader activity or if the goal is to assess a marker, then this is not a good design.
Model Class 2: Stratified Design
Options: Stratification overall test, subgroup+overall testing, interaction tests
Measure prospectively or retrospectively

- Subgroup \((R_-)\)
  - New Treatment (B)
  - Standard Treatment (A)

- Subgroup \((R_+)\)
  - New Treatment (B)
  - Standard Treatment (A)

This is not a good design if one believes treatment can only be efficacious for \((R_+)\) group.
SWOG: a diverse network and part of US NCTN

- Network of 650+ sites, including:
  - 40 core member institutions
  - ~14 strongly associated Lead Academic Participating Sites
  - 28 NCI-designated cancer centers
  - 27 Community Clinical Oncology Programs
  - 27 SPORES
  - Extensive collaboration within Canada
  - Sites in Europe, Middle East, Latin America, Asia

- Membership includes:
  - More than 5,000 researchers & clinicians
  - Almost 5,000 research nurses & clinical research associates
The Past: A design based on a prognostic model: SWOG 9704
S9432 Phase II pilot study: High Dose Therapy with Transplant for Newly Diagnosed KI67 Positive Diffuse Aggressive Lymphoma

- Based on KI67 proliferation model from prior samples
- Identified a very poor risk group
- KI67>80% cell staining
  - 3 year OS of 18% versus 56%. This population is appropriate for high dose chemotherapy and transplant [optimistic difference]
  - 18% of patients with diffuse aggressive lymphoma have a KI67 > 80% [small subgroup size]
- Frozen tissue/paraffin was sent to University of Arizona
- “Real” time communication back to institution to determine treatment assignment
- Study closed due to poor accrual (3 patients)
Alternative prognostic model and supportive data

- International prognostic index (IPI) for lymphoma developed from a large data base
- Combination of multiple easily measured clinical variables; no need for tissue
- IPI = Stage II vs. III/IV, low vs. high LDH, performance status 0-1 vs. ≥ 2, > 1 extra nodal site
  - High-Int risk ≥ 3 factors, High Risk ≥ 4 factors
- Retrospective analysis of a French Phase III study supporting high dose therapy in poor prognostic group, the high-intermediate risk which was approximately 30% of the patients
S9704: A Randomized Phase III Trial Comparing Early High Dose Therapy and Autologous Stem Cell Transplant to Conventional Dose CHOP/R Chemotherapy for Patients with Diffuse Aggressive Non-Hodgkin's Lymphoma in High-Intermediate and High Risk Groups

- Eligible: 370
- Eligible for randomization: 253
S9704 Timeline

- S9704 Activated 9/15/97
- Results from a large randomized study CHOP vs. CHOP-Rituximab showing improved survival for CHOP-R.
- Rituximab was added for all B-cell CD20+ lymphomas on 4/1/03
- Chose not to redesign the trial to target only B-cell CD20+ patients
- Trial closed 12/17/07 after reaching its randomization accrual goal
# S9704 Results: Grade III–IV Toxicities

<table>
<thead>
<tr>
<th>Toxicities</th>
<th>CHOP (R) x 1 + ASCT (%)</th>
<th>CHOP (R) x 3 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infection</td>
<td>50</td>
<td>13</td>
</tr>
<tr>
<td>GI</td>
<td>26</td>
<td>5</td>
</tr>
<tr>
<td>Metabolic</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>Lung</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>CV</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>Neurologic</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Hypoxia</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Hepatic</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Treatment deaths</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>

N=253 randomized patients
Outcome of randomized patients

- Targeting the poor prognostic subgroup identified a group that benefited for PFS but not OS.
- Some suggestion of greater effect in the highest risk group (interaction p-value .02).
While only exploratory there was suggestion of an effect in the highest risk group.

Was the poor prognostic group targeting not sufficiently aggressive?
Diffuse Large Cell Lymphoma: Gene Expression on archived tissue specimens (same disease as S9704)

- Gene expression arrays (quantitative, large numbers)
  - Fresh or frozen tissue (problematic for multi-institutional studies, also often a problem wrt to use of historical samples)
- Gene expression from paraffin (array plate technology) <100 genes
  - Great for our multi-institutional cooperative group studies
- Data from several clinical trials.
  - Both before and after the introduction of Rituxan therapy to standard chemotherapy
- Analysis focused on overall prognostic effect, no evidence of interactions
Hazard rates for multiple genes for DLBCL

HLA-DRB

CCND2

PRKCB1

SERPINA9

c-MYC

ACTN1

Rimsza et al. 2011
Practical Issues

- The biomarker wasn’t workable yet in S9432.
- The fraction of high risk patients (targeted group was less than expected.
- There were questions of when to hold the design fixed and when to be more flexible. It was a practical choice for S9704 not to redesign mid-trial after the introduction of Rituximab for the B-Cell subgroup.
- Given the limited sample sizes available, we need to consider modeling based on data from multiple sources to guide targeting.
Recent Past and Present

- Recently multiple examples of genomic or other biomarker targeted studies
- Antje Hoering presented SWOG studies
  - Lung Cancer Study S0819
  - Breast Cancer Study S1007
- Many more – but with some general themes
  - Typically a single target group
  - Many issues with respect defining target
S1406 Randomized Phase II study of Irinotecan and Cetuximab with or without Vemurafenib in BRAF Mutant Metastatic Colorectal Cancer

Unk/Not BRAF\textsuperscript{V600E} mutation

Irinotecan and Cetuximab + Vemurafenib

BRAF\textsuperscript{V600E}

Irinotecan and Cetuximab

Example of targeting (on mutation at time): If treatment is only effective in a subgroup this is powerful

Special: Embedded Patient-Derived Xenograft Co-Clinical Trial
A New Present: Lung-Map  S1400

- Special thanks to Mary Redman (slides and more)

Also in Canada in Q1 or Q2 of 2015 (hopefully)!
Unmet needs addressed by a Master Protocol

- How to develop drugs for uncommon-rare genotypes?
- How to apply broad-based screening (NGS)?
- How to achieve acceptable turn-around times for molecular testing for therapy initiation? (<2 weeks)
- How to expedite the new drug-biomarker FDA approval process? (companion diagnostic)

Significantly mutated genes in lung SQCC
PS Hammerman et al. Nature 000, 1-7 (2012)
doi:10.1038/nature11404
Sub-studies assigned based on biomarker results, patients with multiple biomarkers randomly assigned to sub-study.

Exp = Targeted therapy (TT) or TT combinations (TTC), Exp\(^1-4\) are different TT/TTC regimens

MT = non-match study experimental therapy or combinations

oC = docetaxel or erlotinib, SoC\(^1-5\) depends on biomarker and TT/TTC/NMT regimen
Study Design and Objectives

Design:
Independently conducted and analyzed parallel Phase II/III studies

Primary Objectives within each sub-study:

**Phase II Component:**
1. To evaluate if there is sufficient evidence to continue to the Phase III component by comparing progression-free survival (PFS) between patients randomized to experimental therapy versus SoC.

**Phase III Component:**
1. To determine if there is both a statistically and clinically-meaningful difference in PFS between the treatment arms.
2. To compare overall survival (OS) between treatment arms.
Goals

• Improve screening
  ◦ Screening large numbers of patients for multiple targets
  ◦ Reduce screen failure rate
  ◦ Provide a sufficient “hit rate” to engage patients & physicians

• Increase speed of drug evaluation and development:
  ◦ Provide an infrastructure to open new sub-studies faster
  ◦ Rapid drug/biomarker testing for detection of “large effects”
  ◦ Facilitate FDA approval of new drugs and bring safe & effective drugs to patients faster
Lung-MAP current sub-studies

FMI NGS/MET IHC

- **Non-match (Anti-PD-L1)**
  - Arm¹
  - Arm²
  - 1:1
  - ¹ Medi4736
    - ² Docetaxel

- **PI3K PIK3CA mut**
  - Arm¹
  - Arm²
  - 1:1
  - ¹ GDC-0032
    - ² Docetaxel

- **CDK4/6 CCND1, CCND2, CCND3, cdk4 ampl**
  - Arm¹
  - Arm²
  - 1:1
  - ¹ Palbociclib
    - ² Docetaxel

- **FGFR FGFR ampl, mut, fusion**
  - Arm¹
  - Arm²
  - 1:1
  - ¹ AZD4547
    - ² Docetaxel

- **HGF c-Met Expr**
  - Arm¹
  - Arm²
  - 1:1
  - ¹ Rilotumumab + erlotinib
    - ² Erlotinib
**Patient-Sample Schema**

- **Patient Registration Consent**
- **Assign Sub-study by marker**
- **Genomic Screening ≤ 16 days**
- **28 days to register**
- **1:1 Randomization**
- **Investigational Therapy**
- **Standard of Care Therapy**

**Central genomic screening:**
- Foundation Medicine: NGS test platform
- Clariant: c-MET IHC
Study Design Within Each Sub-study

- **Randomization**
  - Phase II Analysis
    - 55 PFS events
  - Phase III Interim Analyses
    - OS for efficacy
    - PFS/OS for futility

- **Complete Accrual**
  - 258 OS events
  - 290 PFS events

- **Final Analysis**

- **12 months follow-up**

- Futility established
## Statistical Design: Phase II Interim Analysis

Each sub-study can choose between Plan A or Plan B to determine “bar” for continuation past Phase 2 interim analysis.

<table>
<thead>
<tr>
<th></th>
<th>Plan A</th>
<th>Plan B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary Outcome</strong></td>
<td>PFS</td>
<td></td>
</tr>
<tr>
<td><strong>Sample Size</strong></td>
<td>55 progression events</td>
<td></td>
</tr>
<tr>
<td><strong>Target HR (% improvement)</strong></td>
<td>HR = 0.5 2-fold increase</td>
<td>HR = 0.4 2.5-fold increase</td>
</tr>
<tr>
<td><strong>Power</strong></td>
<td>90%</td>
<td>95%</td>
</tr>
<tr>
<td><strong>Type I error</strong></td>
<td>10%</td>
<td>4%</td>
</tr>
<tr>
<td><strong>Approx. Threshold to continue:</strong></td>
<td>HR = 0.71 41% increase</td>
<td>HR = 0.61 63% increase</td>
</tr>
</tbody>
</table>
### Statistical Design: Phase III

<table>
<thead>
<tr>
<th></th>
<th>PFS and OS Co-primary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>PFS</strong></td>
</tr>
<tr>
<td>Events</td>
<td>290</td>
</tr>
<tr>
<td>Null Hypothesis (HR)</td>
<td>0.75* (33% improvement)</td>
</tr>
<tr>
<td>Alternative Hypothesis</td>
<td>0.5 (2-fold increase)</td>
</tr>
<tr>
<td>Type I error (1-sided)</td>
<td>0.014 against HR = 1.33 &lt; 0.000001 against HR = 1</td>
</tr>
<tr>
<td>Power</td>
<td>90%</td>
</tr>
</tbody>
</table>

*Non HR = 1 null hypothesis encodes clinical significance

Sample size based on OS for all studies
### Sample Size for the Sub-studies

<table>
<thead>
<tr>
<th>Sub-study ID</th>
<th>Prevalence Estimate$^1$</th>
<th>Approximate Sample Size</th>
<th>Approximate time of analysis</th>
<th>Sample Size</th>
<th>Approximate time of analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>400A(non-match)$^2$</td>
<td>56%</td>
<td>170</td>
<td>8</td>
<td>400</td>
<td>21</td>
</tr>
<tr>
<td>400B(PI3K)$^3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GNE+</td>
<td>6%</td>
<td>78</td>
<td></td>
<td>288</td>
<td></td>
</tr>
<tr>
<td>FMI+</td>
<td>8%</td>
<td>152</td>
<td>19</td>
<td>400</td>
<td>72</td>
</tr>
<tr>
<td>400C(CDK4/6)</td>
<td>12%</td>
<td>124</td>
<td>11</td>
<td>312</td>
<td>45</td>
</tr>
<tr>
<td>400D (FGFR)</td>
<td>9%</td>
<td>112</td>
<td>11</td>
<td>302</td>
<td>53</td>
</tr>
<tr>
<td>400E (HGF)</td>
<td>16%</td>
<td>144</td>
<td>9</td>
<td>326</td>
<td>37</td>
</tr>
</tbody>
</table>

Prevalence estimates: 35% with 1; 8% with 2; 0.8% with 3; 0% with 4 biomarkers

S1400A design and minimum PD-L1+: 50 (phase 2), 114 (phase 3) patients

S1400B design: eligibility based on FMI criteria, but designed around subgroup defined to be GNE+ (assumed ~70% of FMI+)
### Study development time-line

<table>
<thead>
<tr>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>Q2</td>
<td>Q3</td>
<td>Q4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Pre-Study Activities, Planning
- Project Management
- Assay Co. Selection
- Protocol Development
- Master IND application
- Database, systems, forms
- Master IDE application
- Approvals (CTEP, CIRB)
- Contracts
- Study Drug Management
- Clinical Operations Management
- Stat/Data Oversight, Management, and Analysis
- Team Meetings, Teleconferences
- Other Activities
- Initial Meeting March 2013
- FDA Meeting November 21, 2013
- Trial Starts June 16, 2014

---

**Notes:**
- Initial Meeting: March 2013
- FDA Meeting: November 21, 2013
- Trial Starts: June 16, 2014
Design Issues

- Master study - but how much variation by sub-study for design specifications?
- Different target efficacy by sub-study
- Additional assay(s) added to FMI assay
- Frequency of marker subgroups – what sub-study frequency remains feasible?

Sub-study Eligibility $\rightarrow (R, M(R), \nu(R)) \leftarrow$ Fraction of patients in sub-study

Treatment effect
Complex study lessons learned

- Communicate early and often with partners
  - OPEN(registration) saw Lung-MAP as one study, but we were planning to activate it as six.
  - Better specifications for how the marker data would be received. Plan for change (Central IRB, new assays)
  - Improved communication with pharmaceutical partners and institutions regarding SWOG structure, attributes and processes
Learning more from Master protocols

- Impact of dynamic multiple sub-study design and inference (as genotype groups open and close patient population changes)
- Opportunities for modeling of treatment effects are possible based on detailed genomic data and additional use of specimens

\[(R, M(R), \nu(R))\]
Acknowledgments

Collaborators

• Key statistical center Lung-MAP team:
  Lead Statistician: Mary Redman
• Design methods: Antje Hoering, John Crowley
• Target subgroup modeling: Charles Kooperberg
End