Phase III Design Considerations for Molecularly Targeted Agents

Boris Freidlin
National Cancer Institute

November 4, 2014
• Molecularly targeted agents may only benefit a subgroup of a histologically defined population

• Successful evaluation requires co-development of biomarkers to identify sensitive subpopulations

• Various design strategies to integrate treatment and biomarker evaluation are available

• Choice of Phase III design depends on the biomarker’s credentials
Focus of this talk

• Binary biomarker separates the population of interest into biomarker-positive (B+) and biomarker-negative (B-) subgroups

• Analytical validity of the biomarker assay has been established

• Biomarker credentials are sufficient to assume that B- patients benefit only if B+ patients benefit
Biomarker with strong credentials: use Enrichment design

Measure biomarker

Biomarker-positive (B+)
  └ Randomize
    ├ New Drug
    └ Control

Biomarker-negative (B-)
  └ Off Study
Example: BRIM3 study
Vemurafenib in metastatic melanoma
• 2107 patients screened to identify 675 patients with BRAF mutation

• BRAF mutated patients randomized to vemurafenib vs. standard chemotherapy

• Overall survival  HR 0.37, 95% CI 0.26–0.55

Reference: Chapman et al NEJM 2011
Limitations of Enrichment design:
Unless external evidence clearly limits benefit to B+ patients, a positive enrichment study leaves open:

• Whether the treatment benefit extends to biomarker-negative patients

• Whether the costs and inconvenience of routine use of the biomarker to select patients for treatment are justified
Biomarker credentials are not compelling: use Biomarker-stratified (randomize-all) designs

- **Measure biomarker**
  - **Biomarker-positive (B+)**
    - Randomize
      - New Drug
      - Control
  - **Biomarker-negative (B-)**
    - Randomize
      - New Drug
      - Control
Goals of biomarker-stratified Phase III:

• Asses benefit in each biomarker subgroup

• Recommend drug to patients who benefit

• Do not recommend drug to patients who do not benefit
Assess benefit in each biomarker subgroup

For biomarker positive subgroup (B+)

$H_{0+} : \delta_+ = 0 \quad \text{vs.} \quad H_{A+} : \delta_+ > 0$

For biomarker negative subgroup (B-)

$H_{0-} : \delta_- = 0 \quad \text{vs.} \quad H_{A-} : \delta_- > 0$

Where $\delta_+$ and $\delta_-$ are treatment effects in B+ and B-, respectively.
Three possible Null hypotheses

1) Global Null \[ H_0 = H_{0+} \cap H_{0-} \]

2) No benefit in B- \[ H_{A+} \cap H_{0-} \]

3) No benefit in B+ \[ H_{0+} \cap H_{A-} \]
Type I errors to control

1) \( \Pr[\text{Reject } H_{0+} \text{ or } H_{0-} \mid H_0] < \alpha \)

2) \( \Pr[\text{Reject } H_{0-} \mid H_{A+} \cap H_{0-}] < \alpha^* \)
Subgroup-specific parallel strategy

Test B+
If significant at level $\alpha_1$

- **YES**: Recommend treatment for B+
- **NO**: Do not recommend treatment for B+

Test B-
If significant at level $\alpha - \alpha_1$

- **YES**: Recommend treatment for B-
- **NO**: Do not recommend treatment for B-
Subgroup-specific sequential strategy

Test B+
If significant at level $\alpha$

YES vs.
NO

Test B-
If significant at level $\alpha$

YES
Recommends treatment for all patients

NO
Recommends treatment for B+ only

STOP

Do not recommend the treatment
For subgroup-specific designs

\[
\text{Pr}[\text{Reject } H_{0+} \text{ or } H_{0-} \mid H_0] < \alpha
\]

\[
\text{Pr}[\text{Reject } H_{0-} \mid H_{A+} \cap H_{0-}] < \alpha
\]
Example: PRIME study
Panitumumab in metastatic colorectal cancer

- Biomarker: KRAS status

- KRAS WT 656 patients,
  Hazard Ratio 0.80  95%CI (0.66, 0.97)

- KRAS MT 440 patients
  Hazard Ratio 1.29  95%CI (1.04, 1.62)

Reference: Douillard et al  JCO  2010
Biomarker-positive/overall parallel strategy

Test overall population if significant at level $\alpha_1$

YES
- Recommend treatment for all patients
NO
- Do not recommend treatment for all patients

Test B+ if significant at level $\alpha - \alpha_1$

YES
- Recommend treatment for B+
NO
- Do not recommend treatment for B+
Biomarker-positive/overall sequential strategy

Test **B+**
if significant at level $\alpha$

- **YES**
  - Test overall population
    if significant at level $\alpha$
      - **YES**
        - Recommend treatment for all patients
      - **NO**
        - Recommend treatment for B+ only
  - **NO**
    - STOP

- **NO**
  - Do not recommend treatment
For biomarker-positive/overall designs

$$\Pr[\text{Reject } H_{0+} \text{ or } H_{0-} \mid H_0] < \alpha$$

$$\Pr[\text{Reject } H_{0-} \mid H_{A+} \cap H_{0-}]$$ is not controlled
(could be as high as 100%)
Biomarker-positive/overall strategy may formally recommend treatment for biomarker-negative patients even though the treatment is ineffective in these patients.

Reason: even with no benefit in B- patients a statistically significant effect can be still observed in the overall population if the effect in B+ patients is large.
Example: Lapatinib+letrozole vs. Placebo +letrozole in metastatic breast cancer
Biomarker: HER2 status

- HER2-positive (n=219)
  HR=.71   p-value=.019

- Overall population (n=1286)
  HR=0.86  p-value=.026

- HER2-negative (n=952)
  HR=0.9   p-value=.188

Reference: Johnston et al  JCO  2009
**Marker Sequential Test MaST(α, α₁)**

- **Test B⁺**
  - If significant at level $α₁$
    - YES
    - Test B⁻
      - if significant at level $α$
        - YES
          - Recommend treatment for all patients
        - NO
          - Recommend treatment for B⁺ only
    - NO
      - Test overall population
        - if significant at $α₂ = α - α₁$
          - YES
            - Recommend treatment for all patients
          - NO
            - Do not recommend the treatment

Reference: Freidlin et al. Clinical Trials 2013
For MaST($\alpha$, $\alpha_1$) procedure

\[ \Pr[\text{Reject } H_{0+} \text{ or } H_{0-} \mid H_0] < \alpha \]

\[ \Pr[\text{Reject } H_{0-} \mid H_{A+} \cap H_{0-}] \text{ depends on } \alpha_1 \]
Probability of rejecting $H_0$ as a function of $\alpha_1$
(for $\alpha=0.025$)
Probability of rejecting $H_0$ as a function of $\alpha_1$ (for $\alpha=0.05$)
Rejection regions

Sequential subgroup-specific test

MaST(.025, .022)

Reject $H_{0+}$ and $H_{0-}$

Reject $H_{0+}$ and $H_{0-}$

Reject $H_{0+}$
## Power comparison (B+ prevalence 30%)

<table>
<thead>
<tr>
<th>True hazard ratio</th>
<th></th>
<th>Overall test</th>
<th>Sequential subgroup-specific test</th>
<th>MaST(.025, .022)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM+</td>
<td>BM-</td>
<td>BM+</td>
<td>BM-</td>
<td>BM+</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>.0249</td>
<td>.0256</td>
<td>.0007</td>
</tr>
<tr>
<td>.60</td>
<td>1</td>
<td>.469</td>
<td>.902</td>
<td>.023</td>
</tr>
<tr>
<td>.71</td>
<td>1</td>
<td>.243</td>
<td>.600</td>
<td>.015</td>
</tr>
<tr>
<td>.60</td>
<td>.60</td>
<td>1</td>
<td>.902</td>
<td>.901</td>
</tr>
<tr>
<td>.60</td>
<td>.71</td>
<td>.997</td>
<td>.902</td>
<td>.839</td>
</tr>
<tr>
<td>.71</td>
<td>.71</td>
<td>.981</td>
<td>.600</td>
<td>.552</td>
</tr>
</tbody>
</table>
MaST design allows one to

• minimize the probability of recommending ineffective treatment for B- patients
• maximize power for treatments with homogeneous treatment effect

Example: E1910 Blinatumomab in ALL
n=285, MAST(\(\alpha=0.025\) \(\alpha_1=0.02\))

Biomarker: MRD
Interim monitoring
(Group-specific and MaST designs)

**Efficacy**
First B+ subgroup, if positive then B- subgroup
(no overall population testing)

**Futility**
B+ subgroup: if negative the entire study stops
B - subgroup: if negative B- accrual is stopped
Interim monitoring (efficacy and futility)

• Start with B+ patients (enrichment design) → if early signal in B+ expand enrollment to B- (Liu et al, Clin Trials 2010)

• Can use an early endpoint similar to Phase II/III

• Analyze using sequential subgroup-specific strategy
Sample size consideration (MaST)

Biomarker with relatively strong credentials - need enough B+ patients to detect a meaningful benefit in B+ subgroup:

• Subgroup-specific calculation using $\alpha$ to size B+ (minor loss of power)

• Use $\alpha_1$ to size B+ subgroup (minor increase in sample size, e.g., $\leq 4\%$ for a design with overall $\alpha=.025$)
Design considerations: prevalence of B+

- MaST is recommended when B+ prevalence is <70%

- If B+ prevalence is low limit size of B- cohort

- If B+ prevalence is >80% use sequential subgroup-specific strategy (possibly with relaxed $\alpha$ for B- subgroup)
Unavailable biomarker subgroup

- Biomarker status may be unavailable in a fraction study patients

- Subgroup-specific analysis does not include unavailable status patients

- MaST: two options for these patients
  1) do not include
  2) include in the overall analysis
MaST: unavailable status pts in overall test (proportion unavailable – $\rho_{ub}$)

- False-positive for $H_{A+}$ is controlled at .025

- $\Pr[\text{Reject } H_{0-} \mid H_{A+} \cap H_{0-}]$ could exceed .025, e.g., for $\rho_{ub} = 20\%$ it could be as high as .03

- Adjustment to $\alpha_2$ (assuming MCAR)

$$\alpha_2^* = 1 - \Phi\left(\frac{1}{\sqrt{1 - \rho_{UB}} \left(Z_{\alpha_2} - Z_{\beta^*}\right) - Z_{\beta^*}}\right)$$
Recommendations

• Optimize predictive value of biomarker before designing phase III

• Select phase III design based on biomarker credentials

• Ensure adequate control of relevant false-positive error rates
References