Design of Phase II Clinical Trials with a Potential Predictive Biomarker

November 4, 2014

Ed Korn
Biometric Research Branch
National Cancer Institute
Phase II trials are designed to decide whether to take an experimental therapy to a definitive phase III trial.
Possible phase III trial designs with a biomarker

(1) Enrichment design

(2) Biomarker stratified design

(3) Standard phase III design ignoring the biomarker

(4) Biomarker strategy design
Possible phase III trial designs with a biomarker

(1) Enrichment design

Assess biomarker

Biomarker positive → Randomize → Treatment A → Treatment B

Biomarker negative → Off study
Possible phase III trial designs with a biomarker

(2) Biomarker stratified design
Possible phase III trial designs with a biomarker

(3) Standard phase III design ignoring the biomarker
Possible phase III trial designs with a biomarker

(4) Biomarker-strategy design

Possible phases of trial designs with a biomarker

- Biomarker directed arm
  - Biomarker positive → Treatment A
  - Biomarker negative → Treatment B

- Control arm → Treatment B

Assess biomarker → Randomize
Possible phase III trial designs with a biomarker

(4) Biomarker-strategy design – Not Generally Recommended

Assess biomarker → Randomize → Biomarker directed arm

Biomarker positive → Treatment A
Biomarker negative → Treatment B

Control arm → Treatment B
Single-Arm Phase II versus Randomized Screening Phase II Design

When can a single-arm design be used instead of a randomized screening design?
When can a single-arm design be used?

(1) Single agent (or experimental agent combined with inactive agents) and responses expected if agent is active

[(2) New treatment is expected to be much more effective than historical treatments with response rate or time-to-event endpoint]

[(3) New treatment with time-to-event endpoint and data from a large collection of historical trials are available]
When can a single-arm design *with a biomarker* be used?

Unless one knows that the biomarker is not prognostic, restricted to:

(1) Single agent (or experimental agent combined with inactive agents) and responses expected if agent is active
Single-arm designs with a biomarker -- Options

(a) Only perform a phase II trial in the B+ group.

(b) Perform two phase II trials concurrently – one in the B+ group and one in the B- group

(c) Perform phase II trial in the B+ group. If the trial is positive, then perform a phase II trial in the B- group.

(d) Perform two-stage trial in the B+ group. If the trial passes its first stage, continue the B+ trial and also begin a trial in the B- group.
Single-arm designs with a biomarker -- Options

(e) Perform an unrestricted phase II trial. Examine the biomarker status of the patients after the trial.

(f) Perform an unrestricted phase II trial. But, if the trial is negative, continue enrollment of only the B+ patients to obtain a B+ phase II trial.

(g) Perform an unrestricted two-stage phase II trial. But, if the trial is negative at either stage, continue enrollment of only the B+ patients to obtain a B+ phase II trial.
Single-arm designs with a biomarker -- Recommendation?

(g) Perform an unrestricted two-stage phase II trial. But, if the trial is negative at either stage, continue enrollment of only the B+ patients to obtain a B+ phase II trial.

Simon two-stage phase II design:

12 at the first stage, with $\geq 1$ response, continue 25 more at second stage. With $\geq 4$ responses out of 37, trial is positive.
Randomized phase II (screening designs) with a biomarker

--Experimental arm and control arm.

--All comers are randomized at first, but with biomarker status determined for analysis

--Design to be discussed uses a progression-free survival (PFS) endpoint
Possible recommendations for the phase III trial after completing the phase II trial

(1) Enrichment design

(2) Biomarker stratified design

(3) Standard phase III design ignoring the biomarker

(4) No further testing of new therapy
STEP 1
Test $H_{0(+)}$: $HR_{(+)} \leq 1$ vs. $HR_{(+)} > 1$

in biomarker-positive subgroup at one-sided $\alpha = 0.10$ level

Do not reject $H_{0(+)}$

(Targeted therapy no better than standard in biomarker-positive subgroup.)

Reject $H_{0(+)}$

(Targeted therapy better than standard in biomarker-positive subgroup.)

STEP 2A
Test $H_0$: $HR_{(o)} \leq 1$ vs. $HR_{(o)} > 1$ in overall group at one-sided $\alpha = 0.05$ level

Do not reject $H_0$

(Targeted therapy harmful or only marginally helpful in biomarker-negative subgroup)

Reject $H_0$

(STEP 2B)
Form 80% confidence interval for $HR_{(-)}$ in biomarker-negative subgroup, $CI_{HR}$

$CI_{HR}$ below 1.3

(Targeted therapy better than standard in biomarker-negative subgroup)

$CI_{HR}$ includes 1.3 or 1.5

(Inconclusive in biomarker-negative subgroup)

$CI_{HR}$ above 1.5

(Targeted therapy better than standard in biomarker-negative subgroup)

Recommend NO FURTHER TESTING of new therapy

Phase III trial: DROP BIOMARKER, (standard phase III design)

Phase III trial: BIOMARKER-ENRICHMENT DESIGN

Phase III trial: BIOMARKER-STRATIFIED DESIGN

Phase III trial: DROP BIOMARKER (standard phase III design)
Simulations

Trial designed to detect a doubling of the median PFS in the biomarker subgroup (hazard ratio=2) with 90% power at the one-sided 10% significance level.

Trial requires 56 PFS events in the biomarker-positive subgroup, corresponding to 70 biomarker-positive patients.

Cut-off accrual to any biomarker subgroup at 140.

Results presented for 20% prevalence of biomarker-positive patients.

Approximate sample sizes=70 biomarker-positive patients, 140 biomarker-negative patients.
Biomarker Subgroup

<table>
<thead>
<tr>
<th></th>
<th>Positive</th>
<th></th>
<th>Negative</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Expt Tx. Control</td>
<td>Median  Median HR</td>
<td>Expt Tx. Control</td>
<td>Median  Median HR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4</td>
<td>1.0</td>
<td>4</td>
</tr>
</tbody>
</table>

Probability of recommendations for phase III trial design

- Enrichment Design: 6%
- Biomarker-stratified design: 4%
- No biomarker (standard phase III): 3%
- No further testing: 87%
### Biomarker Subgroup

<table>
<thead>
<tr>
<th></th>
<th>Positive</th>
<th></th>
<th>Negative</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Expt Tx. Control</td>
<td>Expt Tx.</td>
<td>Control</td>
<td>Expt Tx.</td>
<td>Control</td>
</tr>
<tr>
<td>Median Median HR</td>
<td>Median</td>
<td>Median HR</td>
<td>Median</td>
<td>Median HR</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Probability of recommendations for phase III trial design

- Enrichment Design: 53%
- Biomarker-stratified design: 36%
- No biomarker (standard phase III): 1%
- No further testing: 10%
### Biomarker Subgroup

<table>
<thead>
<tr>
<th></th>
<th>Positive</th>
<th></th>
<th>Negative</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Expt Tx. Control</td>
<td>Expt Tx. Control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Median</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>HR</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

**Probability of recommendations for phase III trial design**

- Enrichment Design: 1%
- Biomarker-stratified design: 52%
- No biomarker (standard phase III): 38%
- No further testing: 10%
### Biomarker Subgroup

<table>
<thead>
<tr>
<th></th>
<th>Positive</th>
<th></th>
<th>Negative</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Expt Tx.</td>
<td>Control</td>
<td>Median</td>
<td>Expt Tx.</td>
<td>Control</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Median</td>
<td></td>
<td>Median</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HR</td>
<td></td>
<td>Median</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.75</td>
<td>1.75</td>
<td>1.75</td>
</tr>
</tbody>
</table>

Probability of recommendations for phase III trial design

- Enrichment Design: <1%
- Biomarker-stratified design: 51%
- No biomarker (standard phase III): 48%
- No further testing: 1%
### Biomarker Subgroup

<table>
<thead>
<tr>
<th></th>
<th>Positive Mean</th>
<th>Negative Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expt Tx. Control</td>
<td>8 4 2</td>
<td>6 4 1.5</td>
</tr>
<tr>
<td>HR</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Probability of recommendations for phase III trial design

- Enrichment Design: 2%
- Biomarker-stratified design: 79%
- No biomarker (standard phase III): 18%
- No further testing: 2%
## Biomarker Subgroup

<table>
<thead>
<tr>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expt Tx.</td>
<td>Control</td>
</tr>
<tr>
<td>Median</td>
<td>Median</td>
</tr>
<tr>
<td></td>
<td>HR</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Probability of recommendations for phase III trial design

- **Enrichment Design**: 76%
- **Biomarker-stratified design**: 2%
- **No biomarker (standard phase III)**: <1%
- **No further testing**: 22%
## Biomarker Subgroup

### Positive

<table>
<thead>
<tr>
<th>Expt Tx.</th>
<th>Control</th>
<th>Median</th>
<th>Median</th>
<th>HR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>6</td>
<td>1.0</td>
</tr>
</tbody>
</table>

### Negative

<table>
<thead>
<tr>
<th>Expt Tx.</th>
<th>Control</th>
<th>Median</th>
<th>Median</th>
<th>HR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>4</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Probability of recommendations for phase III trial design

- Enrichment Design: 6%
- Biomarker-stratified design: 4%
- No biomarker (standard phase III): 2%
- No further testing: 87%
How would the proposed design work in the real world?

We take some real randomized phase III data, treat the observed hazard ratios for the Biomarker positive and negative groups as if they were truth, and see what our phase II design would have recommended.
Example 1: Phase III trial of gefitinib versus carboplatin-paclitaxel in NSCLC

Biomarker: EGFR mutation

Ref: Mok et al. 2009 (IPASS)
EGFR mutation subgroup

Positive (60%, n=261)  Negative (40%, n=176)

<table>
<thead>
<tr>
<th>Control</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>Median</td>
</tr>
<tr>
<td>6.0 mo</td>
<td>2.0 mo</td>
</tr>
<tr>
<td>2.08</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Probability of recommendations for phase III trial design

- Enrichment Design: 98%
- Biomarker-stratified design: 0%
- No biomarker (standard phase III): 0%
- No further testing: 2%

Average size of phase II trial = 175
Example 2: Phase III trial of radiotherapy with or without temozolomide for glioblastoma

Biomarker: methylation of MGMT promoter

Ref: Hegi et al. 2005
### MGMT Subgroup

<table>
<thead>
<tr>
<th>Positive (45%, n=92)</th>
<th>Negative (54%, n=114)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Control</strong></td>
<td><strong>Control</strong></td>
</tr>
<tr>
<td>Median</td>
<td>Median</td>
</tr>
<tr>
<td>HR</td>
<td>HR</td>
</tr>
<tr>
<td>5.9 mo</td>
<td>4.4 mo</td>
</tr>
<tr>
<td>2.08</td>
<td>1.61</td>
</tr>
</tbody>
</table>

#### Probability of recommendations for phase III trial design

- **Enrichment Design**: 1%
- **Biomarker-stratified design**: 75%
- **No biomarker (standard phase III)**: 21%
- **No further testing**: 2%

**Average size of phase II trial = 157**
Example 3: Phase III trial of cetuximab versus best supportive care for advanced colorectal cancer

Biomarker: K-ras mutation (positive marker=wild type)

Ref: Karapetis et al. 2008 (reanalysis of CO.17)
K-ras Subgroup

Positive (58%, n=215)                 Negative (42%, n=151)

- Control Control
  - Median       HR   Median     HR
    1.9 mo       2.50 1.8 mo     1.01

Probability of recommendations for phase III trial design

- Enrichment Design 38%
- Biomarker-stratified design 61%
- No biomarker (standard phase III) 0%
- No further testing 1%

Average size of phase II trial = 167
Example 4: Phase III trial of FOLIFIRI with or without cetuximab for metastatic EGFR-positive colorectal cancer

Biomarker: K-ras mutation (positive marker=wild type)

Ref: Van Cutsem et al. 2009
### K-ras Subgroup

<table>
<thead>
<tr>
<th></th>
<th>Positive (64%, n=348)</th>
<th>Negative (36%, n=192)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>8.7 mo</td>
<td>8.1 mo</td>
</tr>
<tr>
<td>HR</td>
<td>1.47</td>
<td>0.93</td>
</tr>
</tbody>
</table>

Probability of recommendations for phase III trial design

<table>
<thead>
<tr>
<th>Design</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrichment Design</td>
<td>36%</td>
</tr>
<tr>
<td>Biomarker-stratified design</td>
<td>37%</td>
</tr>
<tr>
<td>No biomarker (standard phase III)</td>
<td>0%</td>
</tr>
<tr>
<td>No further testing</td>
<td>26%</td>
</tr>
</tbody>
</table>

Average size of phase II trial = 191
Summary

When a single-arm response rate trial is appropriate, it is straightforward to include a biomarker.

For randomized phase II designs, it is possible to design a trial with a biomarker and a reasonable sample size, to help determine what type of biomarker phase III trial design to use.
Some Additional References

Freidlin, McShane, Polley, and Korn (2012). Randomized phase II trial designs with biomarkers. Journal of Clinical Oncology

Jones and Holmgen (2007). An adaptive Simon two-stage design for phase 2 studies of targeted therapies. Contemporary Clinical Trials


Karuri and Simon (2012). A Two-stage Bayesian design for co-development of new drugs and companion diagnostics. Statistics in Medicine
