Integrating Biomarkers in Clinical Trials: an overview

Stefan Michiels, PhD
stefan.michiels@bordet.be
Jules Bordet Institute, Brussels
Université Libre de Bruxelles

Ref: Piquette-Miller & Grant, Clin Pharmacol Ther 2007
TYPES OF BIOMARKERS (1)

“A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention.”

- **Prognostic** biomarkers, which affect the outcome of patients in terms of a clinical endpoint
- **Predictive** biomarkers, which affect the effect of a specific treatment on a clinical endpoint
- **Surrogate** biomarkers, which may replace a clinical endpoint in clinical trials carried out to evaluate the effect of a specific treatment

Ref: Buyse & Michiels, in Kelly & Halabi, Eds: Oncology Clinical Trials 2010

---

TYPES OF BIOMARKERS (2)

- **Once before treatment**
- **When measured**
- **Several times before, during & after treatment**

**Prognostic**

**Predictive**

**Pharmacodynamic**

**Surrogate**
<table>
<thead>
<tr>
<th>Trial phase</th>
<th>Effect of treatment</th>
<th>Type of biomarker</th>
<th>Effect of biomarker</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Known</td>
<td>Prognostic</td>
<td>Known</td>
</tr>
<tr>
<td>III</td>
<td>Unknown</td>
<td>Predictive</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

**Trial design**

**Examples**


---

<table>
<thead>
<tr>
<th>Trial phase</th>
<th>Effect of treatment</th>
<th>Type of biomarker</th>
<th>Effect of biomarker</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Known</td>
<td>Prognostic</td>
<td>Known</td>
</tr>
<tr>
<td>III</td>
<td>Unknown</td>
<td>Predictive</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

**Trial designs**

**Examples**

Retrospective series

Gene signatures (Mammprint™, OncotypeDX®)
**MAMMAPRINT® DEVELOPMENT**

- Training set: 78 pts, 34 with distant metastasis at 5 years
- Gene expression levels ranked by correlation coefficient with metastatic status at 5 years
- 70 genes with highest correlations = « molecular signature »

*Ref: van’t Veer, Nature 2002*

**EXTERNAL VALIDATION OF MAMMAPRINT®**

**Validation study 1**
- N=295 breast cancers from one single center
- Potential bias: inclusion of 61 pts of the training set
- Accuracy of MammaPrint for distant relapse at 5-years:
  - Se (correctly classifying pts who relapsed) = 93%  (CI95% 81% to 99%)
  - Sp (correctly classifying pts who did not) = 53%  (CI95% 44% to 61%)

**Validation study 2**
- N=307 breast cancers from 5 European centers
- Accuracy of MammaPrint for distant relapse at 5-years:
  - Se = 90%  (CI95% 78% to 95%)
  - Sp = 42%  (CI95% 36% to 48%)

*Ref: Van de Vijver, NEJM 2002; Buyse, JNCI 2006; Michiels, BJC 2007*
### BIOMARKER-BASED TRIAL DESIGNS

<table>
<thead>
<tr>
<th>Trial phase</th>
<th>Effect of treatment</th>
<th>Type of biomarker</th>
<th>Effect of biomarker</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Known</td>
<td>Prognostic</td>
<td>Known</td>
</tr>
<tr>
<td>III</td>
<td>Unknown</td>
<td>Predictive</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

#### Trial designs

- Retrospective analyses of randomized phase III trials

#### Examples

- KRAS mutations in CRC
- EGFR mutations in NSCLC

#### LACK OF RESPONSE OF MUTANT KRAS TUMORS: 1ST EVIDENCE IN SMALL RETROSPECTIVE SERIES

<table>
<thead>
<tr>
<th>Reference</th>
<th>Responses among KRAS mutant</th>
<th>Responses among KRAS wild-type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lièvre 2006</td>
<td>0/13</td>
<td>11/17 (65%)</td>
</tr>
<tr>
<td>Di Fiore 2007</td>
<td>0/16</td>
<td>12/43 (28%)</td>
</tr>
<tr>
<td>Khambata-Ford 2007</td>
<td>3/30</td>
<td>24/50 (48%)</td>
</tr>
<tr>
<td>Benvenuti 2007</td>
<td>1/16</td>
<td>10/32 (31%)</td>
</tr>
<tr>
<td>Frattini 2007</td>
<td>1/10</td>
<td>9/17 (53%)</td>
</tr>
<tr>
<td>De Roock 2008</td>
<td>0/42</td>
<td>27/66 (41%)</td>
</tr>
<tr>
<td>Lièvre 2008</td>
<td>0/36</td>
<td>34/78 (44%)</td>
</tr>
<tr>
<td>Amado 2008</td>
<td>0/84</td>
<td>21/124 (17%)</td>
</tr>
</tbody>
</table>
RETROSPECTIVE ANALYSIS OF PHASE III TRIALS

Ref: Amado, JCO 2008

Interaction Test p<0.0001 (progression-free survival)

BIOMARKER-BASED TRIAL DESIGNS

<table>
<thead>
<tr>
<th>Trial phase</th>
<th>Effect of treatment</th>
<th>Type of biomarker</th>
<th>Effect of biomarker</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Known</td>
<td>Prognostic</td>
<td>Known</td>
</tr>
<tr>
<td>III</td>
<td>Unknown</td>
<td>Predictive</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

→ Trial designs
« Clinical utility » trials

→ Examples
MINDACT and TAILORx in early breast cancer
Evaluate Clinical-Pathological risk and 70-gene signature risk in 6000 patients

- **Clinicopathological and 70-gene both HIGH risk**
  - **Clin-Path High**
  - **70-gene Low**: CTx
  - Use clinicopathological risk to decide Chemo or not
  - **Clin-Path High**
  - **70-gene Low**: no CTx

- **Clinicopathological and 70-gene both LOW risk**
  - **Clin-Path Low**
  - **70-gene High**: no Ctx
  - Use 70-gene signature risk to decide Chemo or not
  - **Clin-Path Low**
  - **70-gene High**: Ctx

- **Chemotherapy** 4350 patients
  - Anthracycline-based
  - Taxane Capecitabine-based

- **Endocrine therapy (≤ 6000 patients)**
  - 2yrs Tam → 5yrs Letrozole
  - 7yrs Letrozole
INTERMEDIATE RISK RANDOMIZED DESIGN

Biomarker risk

- Low risk: Std
- Intermediate risk: R
- High risk: Exp

THE TAILOR-X TRIAL

Assess genomic risk using Oncotype DX in 10,000 patients

- HIGH genomic risk
- INTERMEDIATE genomic risk
- LOW genomic risk

Chemotherapy

R1

No chemotherapy
ISSUES WITH CLINICAL UTILITY TRIALS

- Very large sample sizes are required (pragmatic trials)
- Unclear what hypothesis is being tested (treatment effect known)
- Prognostic biomarkers less useful to choose treatment than predictive biomarkers (hence emphasis from now on will be on predictive biomarkers)


BIOMARKER-BASED TRIAL DESIGNS

<table>
<thead>
<tr>
<th>Trial phase</th>
<th>Effect of treatment</th>
<th>Type of biomarker</th>
<th>Effect of biomarker</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Known</td>
<td>Prognostic</td>
<td>Known</td>
</tr>
<tr>
<td>III</td>
<td>Unknown</td>
<td>Predictive</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

⇒ Trial designs
- Cross-over designs
- Bayesian

⇒ Examples
- Molecular profiling
- BATTLE in NSCLC
TRIAL OF MOLECULAR PROFILING

At least two prior lines of therapy for advanced disease, no further therapy available

Molecular profiling of tumor biopsy by IHC, FISH or micro-array to identify target

Targeted agents were all approved agents (cytotoxics or biologicals)

Ref: Von Hoff, AACR 100th Annual Meeting, Denver, CO, April 18-22, 2009

TRIAL OF MOLECULAR PROFILING

Define the “Time To Progression Ratio” as

\[ \text{TTPR} = \frac{\text{TTP}_2}{\text{TTP}_1} \]

The natural history of most advanced tumors suggests that TTPR < 1 (patients tend to progress faster on successive lines of treatment)

Trial designed to test the hypothesis that at least 15% of the patients have TTPR > 1.3

Ref: Mick, Contr ClinTrials 2000
TRIAL OF MOLECULAR PROFILING

Proportion of patients with TTPR > 1.3:
18 / 66 (27%, 95% C.I. 17% - 38%, P = 0.007)

Breast 8 / 18 (44%)
Colorectal 4 / 11 (36%)
Ovarian 1 / 5 (20%)
Others 5 / 32 (16%)

Among the 18 patients with TTPR > 1.3, none would have received same drug through physician’s choice


ISSUES WITH TRIAL OF MOLECULAR PROFILING

• Only 66 patients of 106 could have molecular profiling
• Non-randomized trial, hence no evidence that physician’s choice would have yielded inferior results
• Is TTPR > 1.3 a relevant endpoint?
• Cross-over design inefficient if low correlation between TTP₁ and TTP₂

BAYESIAN ADAPTIVE PHASE II DESIGN
(PUTATIVE BIOMARKER)

- Signature 1 - - - - - • Treatment 1 ($P_1$)
- Signature 2 - - - - • Treatment 2 ($P_2$)
- Signature 3 - - - - • Treatment 3 ($P_3$)
- Signature 4 - - - - • Treatment 4 ($P_4$)
  etc…

$P_i$: probability of allocating treatment $i$

BATTLE (Biomarker-based Approaches of Targeted Therapy for Lung Cancer Elimination)

- Evaluate targeted therapy agent(s) with different biomarker profiles
- Stage IV recurrent non-small cell lung cancer with endpoint: 8-week disease-free survival (DFS) rate

Ref: Lee, Clin Trials 2008; Kim, AACR 2010
BATTLE CHARACTERISTICS (1)

- Treat more patients in promising groups according to each pt’s biomarker profile
- Start with equal randomization but switch to « adaptive » randomization after 20 pts in each treatment arm
- Early stopping for lack of efficacy in a biomarker × treatment group
- Provide an accurate estimate of true DFS in each of the biomarker × treatment groups
- Borrow strength from patients treated with same agent but different biomarker profile

BATTLE CHARACTERISTICS (2)

*Four Molecular Pathways Targeted in NSCLC: BATTLE Program*

**Enrollment into BATTLE Umbrella Protocol**

**Biomarker Profile and Adaptive Randomization**

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>MTH</th>
<th>EGFR</th>
<th>K-Ras and/or B-raf</th>
<th>VEGF and/or VEGFR</th>
<th>KRAS and/or cyclin D1</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.15</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.3</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>0.25</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>0.1</td>
</tr>
</tbody>
</table>

**Agents**
- Erlotinib
- Sorafenib
- Vandetanib
- Erlotinib + Bexarotene
ISSUES WITH BATTLE DESIGN

• Needs prior information to form and rank biomarker groups
• Needs prior distribution for treatment effects
• Choice of adaptive randomization (when to start, what allocation ratio)
• Ethics of adaptive randomization
• Need fast outcome assessment
• Missing or incomplete biomarker profile (with inadequate amount of tissue)
• No power gain as compared to equal randomization

Ref: Lee, Clin Trials 2008

BIOMARKER-BASED TRIAL DESIGNS

<table>
<thead>
<tr>
<th>Trial phase</th>
<th>Effect of treatment</th>
<th>Type of biomarker</th>
<th>Effect of biomarker</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Known</td>
<td>Prognostic</td>
<td>Known</td>
</tr>
<tr>
<td>III</td>
<td>Unknown</td>
<td>Predictive</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

- Trial designs
  - Randomize-all
  - Interaction
  - Biomarker-strategy
- Examples
  - P53 in ABC
  - MARVEL in NCSLC
  - ERCC1 in NSCLC
**RANDOMIZE-ALL DESIGN**

- Biomarker -
  - Std
  - Exp
- Biomarker +

**EORTC 10994 ("P53") TRIAL IN ADVANCED BREAST CANCER**

- + Taxanes
  - p53+
  - p53-
- - Taxanes
  - p53+
  - p53-

- Assessed p53 (assume 15% NA)

- Assumed prevalence = 0.33

- Powered for detecting a larger effect in p53+ subgroup
- Overall: 80% power of detecting a HR=1.25 at $\alpha=0.02$
- Also planned to test for interaction
- Anticipated 5.5 yr accrual, analyses 2.5 yrs after last entry
ISSUES WITH RANDOMIZE-ALL DESIGN

- Randomize-all design not efficient if
  - there is a strong biological rationale for no or little treatment effect in the biomarker - subgroup
  - the prevalence of biomarker + is low
  - the positive predictive value of the assay is high (> 0.90)

ISSUES WITH INTERACTION DESIGNS

- Biomarker often unknown or poorly defined (e.g. EGFR mutations in NSCLC, KRAS mutations in colorectal cancer) for prospective stratification
- The power of the “interaction test” is very low, hence huge sample sizes are required and/or a very sensitive endpoint (e.g. PSA, tumor measurements or functional imaging)

Ref: Buyse, Nature Rev Clin Oncol 2010
BIOMARKER-BASED STRATEGY DESIGN WITH STANDARD CONTROL

Ref: Sargent, JCO 2005

SPANISH ERCC1 TRIAL

Ref: Cobo, JCO 2007
ISSUES WITH BIOMARKER-BASED STRATEGY WITH STANDARD CONTROL

• Low statistical power, especially if prevalence of biomarker is low, since in that case few patients benefit from the marker-based treatment optimization
• Confounding between predictive effect of biomarker and effect of experimental therapy

ISSUES WITH BIOMARKER-BASED STRATEGY WITH RANDOMIZED CONTROL

• Very low statistical power, since random strategy will lead to correct treatment for many patients
• Randomization ratio must reflect biomarker prevalence, often unknown in advance

Ref: Mandrekar, JCO 2009; Freidlin, JNCI 2010; Hoering, CCR 2008; Young, Clin Trials 2010; Lee, Clin Trials 2010

BIOMARKER-BASED TRIAL DESIGNS

<table>
<thead>
<tr>
<th>Trial phase</th>
<th>Effect of treatment</th>
<th>Type of biomarker</th>
<th>Effect of biomarker</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Known</td>
<td>Prognostic</td>
<td>Known</td>
</tr>
<tr>
<td>III</td>
<td>Unknown</td>
<td>Predictive</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

- Trial designs
  - Targeted (selection)

- Examples
  - Herceptin trials in ABC
TARGETED (SELECTION) SINGLE-ARM DESIGN

Marker -

Biomarker

Marker + — Exp

TARGETED (SELECTION) RANDOMIZED DESIGN

Marker -

Biomarker

Marker + — R

Std

Exp

Exp

Exp
### BIOMARKER-BASED TRIAL DESIGNS

<table>
<thead>
<tr>
<th>Trial phase</th>
<th>Effect of treatment</th>
<th>Type of biomarker</th>
<th>Effect of biomarker</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Known</td>
<td>Prognostic</td>
<td>Known</td>
</tr>
<tr>
<td>III</td>
<td>Unknown</td>
<td>Predictive</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

#### Trial designs
- **Targeted**

#### Examples
- **TOGA in AGC**

---

#### TOGA TRIAL DESIGN

Phase III, open-label international study

- **3807 patients with advanced gastric cancer screened**
- **HER2-positive advanced GC** *(n=584)*
- **HER2-negative and non-tested patients**

- **TOGA**: 5-FU or capecitabine + cisplatin *(n=290)*
- **TOGA + Trastuzumab**: 5-FU or capecitabine + cisplatin + trastuzumab *(n=294)*

Ref: Bang, ASCO 2009; Abstract 4556
ISSUES WITH TARGETED DESIGN

- Targeted design may be less efficient than randomize-all design if drug has at least some activity in biomarker - patients
- Effect in biomarker - patients may never be known (e.g. effect of adjuvant trastuzumab in patients with early breast cancer)
- Loss of opportunity for biomarker - patients
- In a randomize-all design, multiple candidate biomarkers can be tested

### BIOMARKER-BASED TRIAL DESIGNS

<table>
<thead>
<tr>
<th>Trial phase</th>
<th>Effect of treatment</th>
<th>Type of biomarker</th>
<th>Effect of biomarker</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Known</td>
<td>Prognostic</td>
<td>Known</td>
</tr>
<tr>
<td>III</td>
<td>Unknown</td>
<td>Predictive</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

- **Trial designs**
  - Adaptive parallel
  - Tandem two-stage

- **Examples**
  - FGFR1 inhibitor in BC
  - saracatinib in PC

---

**ADAPTIVE PARALLEL**

**PUTATIVE BIOMARKER**

- Marker - Simon 2-stage phase II
- Marker + Simon 2-stage phase II

Example: phase II trial of dovitinib in FGFR1-amplified and nonamplified HER2-negative metastatic breast cancer

Ref: Jones 2007; McShane 2009; André et al ASCO 2010
TANDEM TWO-STAGE
(BIOMARKER NOT KNOWN AT ALL)

• Simon 2-stage design in overall population
• End of stage 1:
  – if enough responses ⇒ continue in overall population
  – else ⇒ develop molecular predictor & only recruit patients predicted to be « responders »
• Example: phase II clinical trial of saracatinib as monotherapy in previously treated patients with metastatic pancreatic cancer

Ref: Pusztai, CCR 2007; Nallaparedy, ASCO GI 2010

BIOMARKER-BASED TRIAL DESIGNS

<table>
<thead>
<tr>
<th>Trial phase</th>
<th>Effect of treatment</th>
<th>Type of biomarker</th>
<th>Effect of biomarker</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Known</td>
<td>Prognostic</td>
<td>Known</td>
</tr>
<tr>
<td>III</td>
<td>Unknown</td>
<td>Predictive</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

 Trial designs
• Enrichment
• Prospective subset

 Examples
• IPASS in NSCLC
• SATURN in NSCLC
IPASS TRIAL ENRICHED FOR MUTATION + PATIENTS (ASIAN POPULATION)

EGFR mutation +

- Gefitinib (n=132)
- Carboptatin / paclitaxel (n=129)

HR (95% CI) = 0.48 (0.36, 0.64)
p<0.0001

Nr. events gefitinib: 97
Nr. events Chemo: 111

EGFR mutation -

- Gefitinib (n=91)
- Carboptatin / paclitaxel (n=85)

HR (95% CI) = 2.85 (2.05, 3.98)
p<0.0001

Nr. events gefitinib: 88
Nr. events Chemo: 70

Ref: Mok, NEJM 2009
PROSPECTIVE SUBSET

• Simplest approach: split significance level:
  \[ \alpha = \alpha_{\text{all}} + \alpha_{\text{biomarker+}} \]
  - the new treatment is compared with the control in the overall population, ignoring the biomarker
  - if \( p_{\text{all}} \leq \alpha_{\text{all}} \), claim effectiveness for all patients
  - if not, the new treatment is compared with the control in biomarker + patients only, and if \( p_{\text{biomarker+}} \leq \alpha_{\text{biomarker+}} \), claim effectiveness for biomarker + patients only

• There are less conservative, yet properly controlled, ways of adjusting \( \alpha \) for both (correlated) tests


SATURN DESIGN

1:1 Chemonaïve advanced NSCLC n=1,949

4 cycles of first-line platinum doublet chemotherapy*

Non-PD n=889

Mandatory tumour sampling

Erlotinib 150mg/day

Placebo

PD

PD

Co-primary endpoints:
• PFS in all patients
• PFS in patients with EGFR IHC+ tumours

Stratification factors:
• EGFR IHC (+ vs - vs indeterminate)
• ...

PFS in all patients
PFS in patients with EGFR IHC+ tumours
SATURN OVERALL RESULTS

<table>
<thead>
<tr>
<th></th>
<th>Erlotinib (n = 438)</th>
<th>Placebo (n = 451)</th>
<th>HR</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>PFS</td>
<td>(n = 437)</td>
<td>(n = 447)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PFS at 12 weeks</td>
<td>53%</td>
<td>40%</td>
<td>0.71</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>PFS at 24 weeks</td>
<td>31%</td>
<td>17%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median PFS</td>
<td>12.3 weeks</td>
<td>11.1 weeks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median OS</td>
<td>12 months</td>
<td>11 months</td>
<td>0.81</td>
<td>0.0088</td>
</tr>
<tr>
<td></td>
<td>(n = 436)</td>
<td>(n = 445)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

KEY QUESTIONS FOR DESIGN CHOICE

- What do we know reliably before the trial starts (in terms of treatment and biomarker effects)?
- What do we want to know reliably after the trial is done (i.e. is the trial for discovery or confirmation?)
- What are the consequences of type I and type II errors in terms of biomarker usefulness?
- Will there be repeat trials?