Randomized Trial Designs for Evaluating Predictive Biomarker Tests: What’s the Estimand?

Gene Pennello, Ph.D.
Division of Biostatistics
FDA/CDRH/OSB

Acknowledgements: Qin Li, Ph.D., Jingjing Ye, Ph.D.
Outline

- Predictive Biomarker
- Companion Diagnostic Test (CDx)
- CDx Clinical Trial Design
  - Biomarker-Stratified
  - Biomarker-Strategy
  - Enrichment Design
  - Discordant Risk Randomization
- Estimands
Predictive Biomarkers

- **Predictive biomarker** informs on likely outcomes with specific treatments (e.g., relative sensitivity or resistance).
  - Other names: treatment selection biomarker, CDx

- **Prognostic biomarker** is biological characteristic indicating likelihood of disease progression in a homogeneous population of patients, either not receiving therapy (natural course) or on a standard therapy.
  - Inform on outcomes independent of specific treatment (i.e., in oncology, ability of tumor to proliferate, invade, and/or spread)
Intended Uses / Claims

• **Companion Diagnostic:**
  – Provides information that is essential for the safe and effective use of a corresponding therapeutic product, allowing its benefits to exceed its risks.
  – EX. Defines the population for whom a therapeutic product is indicated.

• **Complementary Diagnostic:**
  – Provides clinically useful information about a therapeutic product yet is not a prerequisite for the therapeutic product’s use (*not an official FDA definition*).
FDA Guidance, Predictive Markers

Beaver JA; Tzou A; Blumenthal GM; McKee AE; Kim G; Pazdur R; Philip R. An FDA Perspective on the Regulatory Implications of Complex Signatures to Predict Response to Targeted Therapies. Clin Cancer Res. 2017, 23 (6), 1368-1372.

US FDA. Guidance on Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products. US FDA: Silver Spring, MD, 2012.


US FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource [Internet]. Silver Spring (MD): Food and Drug Administration (US); 2016.
Qualitative Interaction (NSCLC)

A  EGFR-Mutation-Positive

B  EGFR-Mutation-Negative

<table>
<thead>
<tr>
<th>No. of patients at risk</th>
<th>Time (months)</th>
<th>0</th>
<th>4</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gefitinib</td>
<td>132</td>
<td>108</td>
<td>71</td>
<td>31</td>
<td>11</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Carboplatin plus paclitaxel</td>
<td>129</td>
<td>103</td>
<td>37</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. of patients at risk</th>
<th>Time (months)</th>
<th>0</th>
<th>4</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gefitinib</td>
<td>91</td>
<td>21</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Carboplatin plus paclitaxel</td>
<td>85</td>
<td>58</td>
<td>14</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Examples of qualitative interactions. Gefitinib vs carboplatin + paclitaxel for first-line treatment of non–small cell lung cancer patients with EGFR mutation–positive tumors (A) and EGFR mutation–negative tumors (B) [adapted from Figure 2 of Mok et al (11)]. Reprinted with permission. Copyright 2009 Massachusetts Medical Society. Cetuximab + chemo vs chemotherapy for first-line treatment of non–small cell lung cancer patients with high-expressing EGFR immunohistochemistry (IHC)–positive tumors (C) and low-expressing EGFR IHC–positive tumors (D) [adapted from Figure 4 of Pirker et al. (13)]. Reprinted with permission. Copyright 2012 Elsevier. PFS = progression-free survival.
Quantitative Interaction (NSCLC)

C  EGFR mutation

HR = 0.10 (95% CI = 0.04 to 0.25
Log-rank $P < .0001$

D  EGFR wild-type

HR = 0.78 (95% CI = 0.63 to 0.96
Log-rank $P < .02$

**Figure 2.** Examples of quantitative interaction: pazopanib vs placebo for locally advanced or metastatic renal cell carcinoma patients with high interleukin 6 (IL-6) values (A) and low IL-6 values (B) [adapted from Figure 2 of Tran et al. (14). Reprinted with permission. Copyright 2012 Elsevier]. Erlotinib maintenance therapy vs placebo for non–small cell lung cancer patients with *EGFR* mutation–positive tumors (C) and *EGFR* wild-type tumors (D) [adapted from Figure 3 of Brugger et al. (15). Reprinted with permission. Copyright 2011 American Society of Clinical Oncology]. Note that data were not available from Brugger et al. (15) to provide the number of patients at risk for (C) and (D). CI = confidence interval; HR = hazard ratio; PFS = progression-free survival.
References


Bossuyt PM; Lijmer JG; Mol BW. Randomised comparisons of medical tests: sometimes invalid, not always efficient. Lancet 2000, 356 (9244), 1844-1847.


References


Biomarker-Drug Trials

marker-positive, +
marker-negative, –

Treatment A is experimental arm.
Treatment B is control arm, typically standard-of-care or placebo.
A. MARVEL (NSCLC, 2nd Line)

Assess $EGFR$ gene copy number by FISH

- **FISH(+)**
  - Randomize
  - Erlotinib
  - Pemetrexed

- **FISH(-)**
  - Randomize
  - Erlotinib
  - Pemetrexed
Freidlin; B.; McShane; L.M., Korn, E.L. Randomized Clinical Trials with Biomarkers: Design Issues. J Natl Cancer I. 2010, 102 (3), 152-160

**B. CALGB-30506 (NSCLC, Stage 1)**

- **High risk**
  - Randomize
  - Vinrolebin+cisplatin
  - Physician’s choice: Docetaxel+cisplatin
  - Observation

- **Low risk**
  - Randomize
  - Vinrolebin+cisplatin
  - Physician’s choice: Docetaxel+cisplatin
  - Gemcitabine+cisplatin
  - Observation

Assess lung metastase score

[Diagram showing the flow of patients through different treatment options based on risk level.]
Biomarker-Stratified
Retrospective analysis of RCT

- Apply CDx biomarker test to stored specimens
- Evaluate treatment effect in CDx subgroups (−, +)
Prospective-Retrospective Biomarker-Stratified Design


Biomarker Strategy Design

C. Marker-Strategy design

- Study Population
- Randomize
- Biomarker directed arm

No biomarker: Treatment B

Marker +

- Treatment A

Marker -

- Treatment B
B. TCA ovarian cancer trial

Assess ATP-TCA

Randomize

Biomarker strategy arm Arm
ATP-based assay directed choice of chemotherapy

Control Arm
Physician’s choice of chemotherapy
Freidlin; B.; McShane; L.M., Korn, E.L. Randomized Clinical Trials with Biomarkers: Design Issues. J Natl Cancer I. 2010, 102 (3), 152-160

A. ERCC1 trial

**NSCLC**

- **Randomize**
  - Assess *ERCC1* gene expression
  - 2:1
  - **Biomarker Strategy Arm**
    - Low *ERCC1*: → cisplatin+docetaxel
    - High *ERCC1*: → gemcitabine+docetaxel
  - **Control Arm**
    - Cisplatin+docetaxel
Freidlin; B.; McShane; L.M., Korn, E.L. Randomized Clinical Trials with Biomarkers: Design Issues. J Natl Cancer I. 2010, 102 (3), 152-160

A. SLCG0601 NSCLC, Stage IV

- Assess
  1. *EGFR* mutation exon 19 or 21
  2. methylation 14-3-3σ gene

- Randomize
  - *EGFR* 19 or 21 mutated
    - Biomarker Strategy
      - 14-3-3σ gene status: methylated → gemcitabine+cisplatin
      - unmethylated → docetaxel+cisplatin
    - Control arm
    - Erlotinib
  - *EGFR* 19 or 21 wild type
    - Off study
Enrichment Design

1. Study Population
2. Biomarker Measurement
3. Randomize
4. Treatment A
5. Treatment B

B. Enrichment design

Off Study

Marker -

Marker +
Enrichment Design

Randomize a subset of subjects defined by diagnostic test value (TRAP-LRTI on PCT ≤ .1)

Discordant Risk Randomization


Figure 2: Trial designs to compare two tests
IHD=ischaemic heart disease; PTCA=percutaneous transluminal coronary angioplasty; R=randomisation process. Abnormal scintigraphy=reversible perfusion defect; abnormal intracoronary flow velocity=insufficient reserve.
Notation

- $\theta_{ab} = E_{ab}(Y) = \text{expectation of } Y \text{ for treatment } A = a, \text{ biomarker status } B = b \ (A, B = 0,1)$. 
  
  - objective response (0,1), event-free survival time

- $\theta_{at}^* = E_{at}(Y) = \text{expectation of } Y \text{ for treatment } A = a, \text{ biomarker test result } T = t \ (A, T = 0,1)$

\[
\theta_{at}^{\text{NDME}} = \sum_{b=0}^{1} \theta_{ab} \Pr(B = b | T = t) = \theta_{a0}(1 - p_t) + \theta_{a1}p_t,
\]

\[
p_t = \Pr(B = 1 | T = t)
\]
Notation

• $\delta_b = \theta_{1b} - \theta_{0b} = \text{treatment effect (difference) in parameter value between treatment arms } a = 0,1 \text{ given biomarker status } B = b (= 0,1)$

• $\delta_t^* = \theta_{1t}^* - \theta_{0t}^* = \text{treatment effect (difference) in parameter value between treatment arms } a = 0,1 \text{ given test result } T = t (= 0,1)$

• $\Delta_{A.B} = \delta_1 - \delta_0 = \text{predictive biomarker capacity.}$
Biomarker Stratified Design

\[ \delta_t^* = \theta_{1t}^* - \theta_{0t}^* \]
\[ \delta_b = \theta_{1b} - \theta_{0b} \]

Estimand

\[ \Delta_{A.T}^* = \delta_1^* - \delta_0^* \]
\[ = (p_1 - p_0)(\delta_1 - \delta_0) \]
\[ = (PPV + NPV - 1)\Delta_{A.B}, \]

= treatment arm by biomarker interaction

\[ \Delta_{A.B} \] attenuated by the factor \( PPV + NPV - 1 \).
Biomarker Strategy Design

- $\Delta^*_{T-S} =$ difference between test-strategy arm and SoC arm in outcome $Y$. Let $\tau = \Pr(T = 1)$.

Estimand

$$\Delta^*_{T-S} = \tau \delta_1^*$$

$$= \tau [\delta_0 + p_1 (\delta_1 - \delta_0)]$$

$$\delta_1 = \delta_0 = \delta$$

$$= \tau \delta$$

$$\delta_0 = 0$$

$$= \tau p_1 \delta_1 = \Pr(TP) \delta_1$$

$$p_1 = p_0 = p$$

Remarks. **Inefficient**: $\delta_1^*$ diluted by $\tau$.

**Invalid**: $\Delta^*_{T-S}$ can be $> 0$ wo HTE or if test random.
Enrichment Design

• $\Delta_{T+}^* = \delta_1^* = \text{treatment effect given test result}$

$T = 1$:  

Estimand  

$$\Delta_{T+}^* = \delta_1^*$$

$$= \delta_0 + p_1 (\delta_1 - \delta_0)$$

$$\delta_1 = \delta_0 = \delta = \delta$$

$\delta_0 = 0$  

$= p_1 \delta_1 = \text{PPV} \delta_1$  

$p_1 = p_0 = p$  

$= p \delta_1$

Remarks.  **Efficient:** $\delta_1^*$ not diluted by $\tau$.

**Invalid?**: $\Delta_{T+}^*$ can be $> 0$ wo HTE or if test random
Discordant Risk Design

- \( \Delta_{A.R}^* \) = difference in treatment effect between two discordant risk arms

Estimand

\[ \Delta_{A.R}^* = \delta_{10}^* - \delta_{01}^* \]

\[ = (p_{10} - p_{01})(\delta_1 - \delta_0) \]

\( \delta_{ts}^* \) = treatment effect given new and standard test results \( T = t \) (= 0,1) and \( S = s \) (= 0,1),

\( p_{ts} = \Pr(B = 1|S = s, T = t) \)

Remarks. If \( T \) is not associated with \( B \) given \( S \), then \( p_{10} - p_{01} = p_{0} - p_{1} < 0 \)
Study Design of RCTs in Literature
Marker Strategy Design

Study Group

Randomize

No PCT

PCT

PCT ≤ .25 ng/ml

AB initiation according to Standard care

AB initiation

No AB initiation

PCT > .25 ng/ml
### Ex. Length of Hospital Stay (Days)

#### Infection Status

<table>
<thead>
<tr>
<th>PCT</th>
<th>Not Bact</th>
<th>Bact</th>
<th>Pr</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \leq 0.25 (-) )</td>
<td>( TN )</td>
<td>( FN )</td>
<td>0.75</td>
</tr>
<tr>
<td>( &gt; 0.25 (+) )</td>
<td>( FP )</td>
<td>( TP )</td>
<td>0.25</td>
</tr>
</tbody>
</table>

#### Random Test

- \( 1 - Sp = Se = 0.25 \)
- \( p_0 = p_1 = 0.10 \)

#### PCT Value

<table>
<thead>
<tr>
<th>PCT Value</th>
<th>PCT Directed</th>
<th>AB Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \leq 0.25 (-) )</td>
<td>( TN(6) + FN(18) = 5.4 )</td>
<td>0.9(6)</td>
</tr>
<tr>
<td>( &gt; 0.25 (+) )</td>
<td>( FP(6) + TP(6) = 1.5 )</td>
<td>0.1(6)</td>
</tr>
</tbody>
</table>

Difference in mean length of stay between PCT-directed and AB always groups is \( \Delta_{Dir-ABI}^{*} = 6.9 - 6.0 = 0.9 \) days.
### Ex. Length of Hospital Stay (Days)

#### Infection Status

<table>
<thead>
<tr>
<th>PCT</th>
<th>Not Bact</th>
<th>Bact</th>
<th>Pr</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 0.25(−)</td>
<td>0.675</td>
<td>0.075</td>
<td>0.75</td>
</tr>
<tr>
<td>&gt; 0.25(+)</td>
<td>0.225</td>
<td>0.025</td>
<td>0.25</td>
</tr>
<tr>
<td>Prev</td>
<td>0.9</td>
<td>0.1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No AB</th>
<th>6</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

#### PCT Value

<table>
<thead>
<tr>
<th>PCT Value</th>
<th>PCT Directed</th>
<th>AB Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 0.25(−)</td>
<td>0.675(6) + 0.075(18) = 5.4</td>
<td>0.9(6)</td>
</tr>
<tr>
<td>&gt; 0.25(+)</td>
<td>0.225(6) + 0.025(6) = 1.5</td>
<td>0.1(6)</td>
</tr>
</tbody>
</table>

Difference in mean length of stay between PCT-directed and AB always groups is $\Delta_{Dir-AB}^* = 6.9 - 6.0 = 0.9$ days.

**Random Test**

$1 - Sp = Se = 0.25$

$p_0 = p_1 = 0.10$
### Ex. Length of Hospital Stay (Days)

<table>
<thead>
<tr>
<th>Infection Status</th>
<th>PCT</th>
<th>Not Bact</th>
<th>Bact</th>
<th>Pr</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 0.25 (−)</td>
<td></td>
<td>0.675</td>
<td>0.010</td>
<td>0.685</td>
</tr>
<tr>
<td>&gt; 0.25 (+)</td>
<td></td>
<td>0.225</td>
<td>0.090</td>
<td>0.315</td>
</tr>
<tr>
<td>Prev</td>
<td></td>
<td>0.9</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>No AB</td>
<td></td>
<td>6</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>AB</td>
<td></td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

**Informative Test**  
\[ Sp = .25 \quad Se = 0.90 \quad p_0 = .0146 \quad p_1 = .2857 \]

<table>
<thead>
<tr>
<th>PCT Value</th>
<th>PCT Directed</th>
<th>AB Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 0.25 (−)</td>
<td>0.675(6) + 0.010(18) = 4.23</td>
<td>0.9(6)</td>
</tr>
<tr>
<td>&gt; 0.25 (+)</td>
<td>0.225(6) + 0.090(6) = 1.89</td>
<td>0.1(6)</td>
</tr>
<tr>
<td></td>
<td>6.12</td>
<td>6</td>
</tr>
</tbody>
</table>

Difference in mean length of stay between PCT-directed and AB always groups is \( \Delta^*_{Dir-ABI} = 6.12 - 6.0 = 0.12 \) days.
### Estimand Values, PCT EX.

<table>
<thead>
<tr>
<th>Trial Design</th>
<th>Estimand</th>
<th>Perfect Test</th>
<th>Random Test</th>
<th>Decent Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\tau_0 = 0$</td>
<td>$\tau_1 = 0.25$</td>
<td>$\tau_0 = 0.25$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\tau_1 = 1$</td>
<td>$\tau_1 = 0.25$</td>
<td>$\tau_1 = 0.90$</td>
</tr>
<tr>
<td>Biomarker</td>
<td>$\delta'_1 - \delta'_0$</td>
<td>$p_0 = 0$</td>
<td>$p_0 = 0.1$</td>
<td>$p_0 = .0146$</td>
</tr>
<tr>
<td>Capacity</td>
<td></td>
<td>$p_1 = 1$</td>
<td>$p_1 = 0.1$</td>
<td>$p_1 = .2857$</td>
</tr>
<tr>
<td>Biomarker</td>
<td>$\Delta^*_{A,T} = \delta'_1 - \delta'_0 = (p_1 - p_0)(\delta'_1 - \delta'_0)$</td>
<td>12</td>
<td>0</td>
<td>3.25</td>
</tr>
<tr>
<td>Stratified</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biomarker</td>
<td>$\Delta^*_{Dir-ABI} = (1 - \tau)\delta'_0 = (1 - \tau)[\delta'_0 + p_0(\delta'_1 - \delta'_0)]$</td>
<td>0</td>
<td>0.9</td>
<td>0.12</td>
</tr>
<tr>
<td>Strategy</td>
<td>$\delta'_0 = (1 - \tau)p_0\delta'_1 = \Pr(FN)\delta'_1$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enrichment</td>
<td>$\Delta^*_{T-} = \delta'_0 = \delta'_0 + p_0(\delta'_1 - \delta'_0)$</td>
<td>0</td>
<td>1.2</td>
<td>0.1752</td>
</tr>
<tr>
<td></td>
<td>$\delta'_0 = p_0\delta'_1 = (1 - NPV)\delta'_1$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$\delta'_t = -\delta_t, \delta^*_t = -\delta^*_t, \tau_b = \Pr(T = 1|B = b)$
Randomize subjects for whom test result and clinician disagree on treatment decision.

Meeting Materials of the FDA’s Microbiology Devices Panel, 10 November 2016.
https://www.fda.gov/advisorycommittees/committeesmeetingmaterials/medicaldevices/medicaldevicesadvisorycommittee/microbiologydevicespanel/ucm515517.htm
Key Subgroups for Adjunctive Tests

- Marker-strategy design compares PCT + SoC and SoC groups on whole population.

- Alternatively, the comparison can be restricted to those subgroups for whom PCT mattered (changed the treatment decision):

<table>
<thead>
<tr>
<th>SoC + PCT</th>
<th>SoC</th>
<th>no ABI</th>
<th>ABI</th>
</tr>
</thead>
<tbody>
<tr>
<td>no ABI</td>
<td>No Change</td>
<td>Change</td>
<td></td>
</tr>
<tr>
<td>ABI</td>
<td>Change</td>
<td>No Change</td>
<td></td>
</tr>
</tbody>
</table>

ABI = antibiotic initiation
Estimand Values, PCT EX.

<table>
<thead>
<tr>
<th>Test</th>
<th>Test Positive Fractions</th>
<th>Predictive Value</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>$\tau_0 = 0.40$ $\tau_1 = 0.90$</td>
<td>$p_0 = .0182$ $p_1 = .2000$</td>
<td>$\rho_0 = .0000$ $\rho_1 = .5556$</td>
</tr>
<tr>
<td>T</td>
<td>$\tau_0 = 0.25$ $\tau_1 = 0.90$</td>
<td>$p_0 = .0146$ $p_1 = .2857$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Estimand</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{10}$</td>
<td>0.0288</td>
</tr>
<tr>
<td>$p_{01}$</td>
<td>0.0146</td>
</tr>
<tr>
<td>$p_{10} - p_{01}$</td>
<td>0.0142</td>
</tr>
<tr>
<td>$\Delta_{A,R}^{*'} = (p_{10} - p_{01})(\delta_1' - \delta_0')$</td>
<td>0.1701</td>
</tr>
</tbody>
</table>

$$\delta_t' = -\delta_t, \; \delta_t'^* = -\delta_t^*, \; \tau_b = \Pr(T = 1|B = b), \; \delta_1' - \delta_0' = 12$$
Discussion

• Clinical trials are conducted in efforts to translate trial results to clinical practice.
• A predictive biomarker test has direct clinical consequences because it is essential for the safe and effective use of a corresponding therapeutic product.
• A predictive biomarker test for a therapeutic product should be evaluated using an estimand that provides a clear link between test results and the outcomes of treatment decisions.
• Unfortunately, some trial designs do not separate predictive accuracy of the test from treatment effect, leading to an inefficient, misleading, or otherwise uninterpretable estimand for evaluating the test / treatment combination.
• Some new performance measures (estimands) have similar interpretability problems.
Global measure of marker performance

Change in response rate under marker-based treatment:

$$\Theta = P(R = 1|T = 1, \Delta(Y) > 0) \cdot P(\Delta(Y) > 0)$$
$$+ P(R = 1|T = 0, \Delta(Y) < 0) \cdot P(\Delta(Y) < 0) - P(R = 1|T = 1)$$
$$= 0.82 - 0.79 = 0.03$$

- 3% increase in 5-year DFS rate under marker-based treatment

http://biostats.bepress.com/uwbiostat/paper389/
Global measure of marker performance

Change in response rate under marker-based treatment:

$$\Theta = P(R = 1|T = 1, \Delta(Y) > 0) P(\Delta(Y) > 0)$$
$$+ P(R = 1|T = 0, \Delta(Y) < 0) P(\Delta(Y) < 0) - P(R = 1|T = 1)$$
$$= 0.82 - 0.79 = 0.03$$

- 3% increase in 5-year DFS rate under marker-based treatment


http://biostats.bepress.com/uwbiostat/paper389/