Biomarkers and Surrogate Endpoints in Clinical Trials

July 28, 2016

Thomas R. Fleming, Ph.D.
Professor, Dept. of Biostatistics
University of Washington


Challenge Questions

~ If an outcome measure, such as a biomarker, is expected to be sensitive to differences between the effects of the control regimen and the experimental regimen, is this condition sufficient to justify its use as the primary endpoint in a registration trial? …Hemodynamics, 6MWD in Pediatric PAH…

~ If responders live longer than non-responders, does this justify the conclusion that treatment effect on response rate reliably predicts treatment effect on patient survival?

~ Similarly, if biomarker status is strongly correlated with a direct measure of how a patient feels, functions or survives, does this justify a conclusion that treatment effect on biomarker status reliably predicts treatment effect on the direct measure of how a patient feels, functions or survives?
Challenge Questions

~ For proper rigorous validation of a biomarker as a replacement (or surrogate) endpoint, are the following of integral importance?

- An in depth clinical understanding of
  - the causal pathways of the disease process; and
  - intervention’s intended & unintended mechanisms of action;
- Meta-analyses of clinical trials showing the relationship between:
  - the net effect of treatment on the biomarker, and
  - the net effect of treatment on direct measures of how a patient feels, functions and survives?

~ Is it feasible that a biomarker could be validated in the early phases of a clinical development plan, and then be used as a replacement (or surrogate) endpoint in the phase 3 trial that is designed to provide a reliable and interpretable assessment of efficacy?
Issues in Replacement (Surrogate) Endpoints

- Criteria for Choosing Endpoints
- “A Correlate does not a Surrogate Make”
- Validation of Replacement (Surrogate) Endpoints
Some Characteristics for Study Endpoints in Phase 3 Clinical Trials

- Consistently & readily measurable
- Sensitive
- Well defined & reliable
- Clinically meaningful

A “Clinically Meaningful Endpoint”:
...a direct measure of how a patient “feels, functions or survives”...

... Robert Temple, FDA

Invasive Procedures: E.g., Biopsy, RHC
Biomarkers & ‘Feels, Functions, Survives’ Endpoints

**Biological Activity:** Hemodynamic Measures in PAH: PVRI, mPAP, CO SBP, DBP, NT-proBNP

**Clinical Meaningful Benefit**

~ **Functions:** Ability to conduct normal activities
  - Ability to walk, Ability to engage in recreational activities,
    Ability for self care, Risk of syncope
  - Time in hospital or missing school (overall, or cause specific)

~ **Feels:**
  - Chest pain, breathlessness, fatigue, dizziness

~ **Survives**
  ...Physician or Observer administered & PROs...
Potential ‘Feels, Functions, Survives’ Endpoints

Patient Reported Outcomes (PROs):

“Any report of the status of a patient’s health condition that comes directly from the patient, without interpretation of the patient’s response by a clinician or anyone else”.

Patient Reported Outcomes (PROs)

...Direct Measures of ‘Feels’, but with need to confirm:

Reliability, Sensitivity, Validity (Content, Construct, etc) 
Clinical Relevance, Interpretability

Integrity, including need for:
  blinded assessment & control of missing data…

…Mobilize disease specific interest groups, 
before sponsors plan clinical trials…

Biomarkers & ‘Feels, Functions, Survives’ Endpoints

- **Biological Activity:** Hemodynamic Measures in PAH: \( PVRI, \ mPAP, \ CO \) \( SBP, \ DBP, \ NT-proBNP \)

- **Clinical Meaningful Benefit**

  - **Functions:** Ability to conduct normal activities
    - Ability to walk, Ability to engage in recreational activities, Ability for self care, Risk of syncope
    - Time in hospital or missing school (overall, or cause specific)

  - **Feels:**
    - Chest pain, breathlessness, fatigue, dizziness

  - **Survives**

    ...*Physician or Observer administered & PROs...*
Biomarkers as Replacement Endpoints

“Biomarkers are measurements of biological processes. Biomarkers include physiological measurements, blood tests and other chemical analyses of tissue or bodily fluids, genetic or metabolic data, and measurements from images.

Cholesterol and blood sugar levels are biomarkers, as are blood pressure, enzyme levels, measurements of tumor size from MRI or CT, and the biochemical and genetic variations observed in age-related macular degeneration...”

Categorization of Nomenclature
Outcome Assessments

Direct Measures of Patient “Functions, Feels, Survives”

- Patient (symptoms: chest pain, dyspnea, fatigue, dizziness)
- Clinician (PANNS for schizophrenia syndrome, Clinician Global Measures)
- Observer (seizures, infant behavior, stroke, death)

Indirect Measures

- Measures depending on patient motivation or clinician judgment to perform the test

Biomarkers

- e.g. $H_bA_1c$, CD-4, PSA, PVRI, NT-proBNP, CO HR, Blood Pressure Pulm Arterial Pressure
- TIMI-III flow
- HDL, LDL, body temperature, urine GAG, urine KS cardiac rhythm, blood cultures, PCR, quantitative measures from radiology imaging.

# John Powers, Dave DeMets, Marc Walton, Laurie Burke, Bob Temple...
Biomarkers (as Replacement Endpoints)

…“Post hoc, ergo, Propter hoc”...

Treatment effects on Biomarkers:

• Establish *Biological Activity*

• But not necessarily the net effects on
  ~ How a patient feels
  ~ The ability to conduct normal activities
  ~ Overall Survival
Issues in Replacement (Surrogate) Endpoints

~ Criteria for Choosing Endpoints

~ “A Correlate does not a Surrogate Make”

~ Validation of Replacement (Surrogate) Endpoints
The Biomarker Endpoint is not in the Causal Pathway of Disease Process
The Biomarker Endpoint is not in the Causal Pathway of the Disease Process.

Disease → Biomarker (e.g., CD4) → Mother-to-Child Trans of HIV

HIV Viral Load

Disease → Biomarker (e.g., CEA, PSA) → Ca. Symptoms & Death

Tumor Burden (NT-proBNP in PAH)

• “Correlates”: Useful for Disease Diagnosis, or Assessing Prognosis
• “Valid Surrogates”: Replacement Endpoints
Multiple Pathways of the Disease Process

Disease → Biomarker Endpoint → Feels, Functions or Survives Endpoint

Intervention

Disease → Biomarker Endpoint → Feels, Functions or Survives Endpoint
Biomarker (as a Surrogate) in Chronic Granulomatous Disease

- CGD $\rightarrow$ Recurrent Serious Infections
- Interferon $\gamma$ …Increase Bacterial Killing and Superoxide Production?

- International CGD Study Group Trial
  Interferon $\gamma$:
  - 70% Reduction in Recurrent Serious Infections
  - Essentially No Effect on Biological Markers
Multiple Pathways of the Disease Process

Intervention

Disease

Surrogate Endpoint

Interferon γ

CGD

Bacterial Killing

Recurrent Serious Infections

Feels, Functions or Survives Endpoint
Biomarkers in Acellular Pertussis Vaccines

(Sweden I Trial with DT control: 10,000 subjects)

- **VE**

<table>
<thead>
<tr>
<th>Vaccine</th>
<th>VE</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>SKB</td>
<td>58%</td>
<td>(51%, 66%)</td>
</tr>
<tr>
<td>Aventis Pasteur</td>
<td>85%</td>
<td>(81%, 89%)</td>
</tr>
</tbody>
</table>

- **Biomarkers**

  Filamentous Haemagglutinin (FHA) and Pertussis Toxoid (PT) antibody responses were superior with the SKB vaccine.
Multiple Pathways of the Disease Process

- AP Vaccine
- FHA & PT Antibodies
- Confirmed Pertussis

- Other Immune Responses, including those resulting from additional antigens in the vaccines:
  - ~ Pertactin
  - ~ Fimbriae (types 2 and 3)

- Durability of effect
Multiple Pathways of the Disease Process

Thrombolytic

What magnitude and what duration is needed?

M.I. (Rapid II / Gusto III)

TIMI III

30-Day Mortality

Intervention

Recurrent Serious Infections

CGD

Biomarker Endpoint
Interventions having Mechanisms of Action Independent of the Disease Process

- **Intervention**
- **Disease**
- **Biomarker Endpoint**
- **Feels, Functions or Survives Endpoint**
Illustration:
Ventricular Arrhythmia after M.I.

- **Arrhythmia:**
  - Risk factor for Sudden Death

- **Antiarrhythmic Drugs:**
  - Class IC antiarrhythmic agents
    …Strong Sodium-Channel Blockade
Illustration:
Ventricular Arrhythmia after M.I.

- Arrhythmia:
  - Risk factor for Sudden Death
- Antiarrhythmic Drugs:
  - Class IC antiarrhythmic agents
    …Strong Sodium-Channel Blockade

Cardiac Arrhythmia Suppression Trial:
The drugs, relative to placebo, TRIPLE the death rate.
Interventions having Mechanisms of Action Independent of the Disease Process

"Deadly Medicine" by Thomas Moore
Interventions having Mechanisms of Action Independent of the Disease Process

ESAs: ↑ Thrombosis ⇒ ↑ Mortality
Cox-2s, Muraglitazar, Rosiglitazone: ↑ CV Risk Factors ⇒ ↑ CV Death/ MI /Stroke
Troglitazone: ↑ Serious Hepatic Risks ⇒ ↑ Morbidity
Natalizumab: ↑ Prog. Multifocal Leukoencephalopathy ⇒ ↑ Morbidity / Mortality
Ezetimibe/Simvastatin: Block pathways linked to CA prot. ⇒ ↑ Cancer Mortality?
Long Acting β-Agonists: ↑ Asthma-related deaths
Torcetrapib: Activates renin angiotensin system ⇒ ↑ BP ⇒ ↑ Mortality
Revatio in Pediatric PAH: ↑ doses ⇒ Improved hemodynamics yet ⇒ ↑ Mortality
“FDA Drug Safety Communication: FDA recommends against use of Revatio in children with pulmonary hypertension”

“The hazard ratio for high dose compared to low dose was 3.5 ($p=0.015$)”
Issues in Replacement (Surrogate) Endpoints

~ Criteria for Choosing Endpoints

~ “A Correlate does not a Surrogate Make”

~ Validation of

   Replacement (Surrogate) Endpoints
End Stage Renal Disease

Goal: Normalize Hematocrit Values
⇒ reduce Death and MI

Standard Dose ESA* ⇒ Hematocrit 30%

High Dose ESA* ⇒ Hematocrit 42%

* Erythropoietin stimulating agent
Patient Distribution & Percent Deaths by Hematocrit %

STANDARD DOSE ESA

- 30% ↓ death RR for 10 pt ↑ in hem.
- ↑ in hematocrit

HIGH DOSE ESA
Patient Distribution & Percent Deaths by Hematocrit %

**STANDARD DOSE ESA**

- 30% ↓ death RR for 10 pt ↑ in hem.
- ↑ in hematocrit
- 30% ↑ in death RR

**HIGH DOSE ESA**
### End Stage Renal Disease

<table>
<thead>
<tr>
<th></th>
<th>Standard Dose ESA</th>
<th>High Dose ESA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Death/MI</td>
</tr>
<tr>
<td>Standard Dose</td>
<td>615</td>
<td>164</td>
</tr>
<tr>
<td>High Dose</td>
<td>618</td>
<td>202</td>
</tr>
</tbody>
</table>

Death / MI relative risk: **1.30** (0.94, 1.79)

Besarab et al, NEJM 339:584-590, 1998:

“↑ in incidence of thrombosis of vascular access sites”
Validation of Replacement (Surrogate) Endpoints

Property of a Valid Replacement (Surrogate) Endpoint:

- Net effect of the Intervention on the Replacement (Surrogate) Endpoint reliably predicts the Net effect of the Intervention on the ‘Feels, Functions, or Survives’ Endpoint
Using Replacement (Surrogate) Endpoints for Direct Measures of ‘Feels, Functions, Survives’

Clinical
- Comprehensive understanding of the
  - Causal pathways of the disease process
  - Intervention’s intended and unintended mechanisms of action

Statistical
- Meta-analyses of clinical trials data
Mechanisms of Action of the Intervention & Causal Pathways of the Disease Process

Torcetrapib

CHD

HDL Cholesterol

LDL Cholesterol

SBP / DBP

CV Morbidity & Mortality
Using Replacement (Surrogate) Endpoints for Direct Measures of ‘Feels, Functions, Survives’

Clinical
- Comprehensive understanding of the
  - Causal pathways of the disease process
  - Intervention’s intended and unintended mechanisms of action

Statistical
- Meta-analyses of clinical trials data
Illustration of Validating a Surrogate

- Anti-Hypertensives
  (>500,000 patients from rand trials)

  ...β-blockers, low dose diuretics, ACE-I, CCBs, ARBs...

  FDA Cardio-Renal Advisory Committee: 6/15/2005

  • Effects on *Blood Pressure* predicting effects on each of the following, considered individually:

    ✓ *Stroke, MI, CVD, Mortality, Heart Failure*
Odds Ratio for CV Events and Systolic BP Difference: Recent and Older Trials


Slide: Henry Black’s lecture
Illustration of Validating a Surrogate

- **Anti-Hypertensives**
  (>500,000 patients from rand trials)
  
  …β-blockers, low dose diuretics, ACE-I, CCBs, ARBs…
  
  FDA Cardio-Renal Advisory Committee: 6/15/2005

- Effects on *Blood Pressure* predicting effects on each of the following, considered individually:
  
  ✓ *Stroke, MI, CVD, Mortality, Heart Failure*
• **Addressing Assay Performance**
  …analysis of analytical performance of an assay…
e.g., limit of quantitation, across lab reproducibility, etc

• **Evidentiary Assessment**
  …relationship between biomarker & disease state
  …data regarding effects of interventions on both biomarker and clinically meaningful outcomes…

• **Justifying the Proposed Use**
  …determining whether available evidence provides sufficient justification for the context of use proposed…
Some Uses of Biomarkers

As “Correlates”…

• Disease Diagnosis
• Assessing Prognosis
• In Patient-specific Therapeutic Strategies
• Primary Endpoints in Screening or Proof of Concept Trials
• Measures of Biologic Activity in Confirmatory (registrational) trials
Uses of Biological Markers: High Clinical Utility

• As Replacement or “Surrogate” Endpoints…
  …When one can fully capture effects on the principle causal mechanism of disease process (w treatment lacking key unintended mech)

• In Identifying Enriched Populations…
  …When the key mechanism(s) of Rx effect on the causal factor(s) of the disease process are specific to a targeted population (eg, gene) (w treatment possibly having unintended mech)

...EGFR Inhibitors: KRAS Wild Type vs. Mutation
Categorization of Nomenclature
Outcome Assessments

**Direct Measures of Patient “Functions, Feels, Survives”**
- Patient (symptoms: chest pain, dyspnea, fatigue, dizziness)
- Clinician (PANNS for schizophrenia syndrome, Clinician Global Measures)
- Observer (seizures, infant behavior, stroke, death)

**Indirect Measures**
- Measures depending on patient motivation or clinician judgment to perform the test

**Biomarkers**
- e.g. $H_bA_1c$, CD-4, PSA, PVRI, NT-proBNP, CO HR, Blood Pressure Pulm Arterial Pressure TIMI-III flow HDL, LDL, body temperature, urine GAG, urine KS cardiac rhythm, blood cultures, PCR, quantitative measures from radiology imaging.

---

John Powers, Dave DeMets, Marc Walton, Laurie Burke, Bob Temple...
Chest Pain
Dyspnea
Fatigue
Hospitalization
L.T., Death

Direct Measures

6-MWD
3-MSC
Exercise testing

Indirect Measures Continuum in PAH

PVRI
NT-proBNP
HR, BP
m-PAP
• **Addressing Assay Performance**
  ...analysis of analytical performance of an assay...
e.g., limit of quantitation, across lab reproducibility, etc

• **Evidentiary Assessment**
  ...relationship between biomarker & disease state
  ...data regarding effects of interventions on both biomarker and clinically meaningful outcomes...

• **Justifying the Proposed Use**
  ...determining whether available evidence provides sufficient justification for the context of use proposed...
Replacement Endpoints

- A replacement endpoint cannot be deemed to be a generic surrogate endpoint for a particular disease

Reasons why use needs setting-specific justification:
  - Multiple causal pathways of the disease process
  - *Magnitude* and *duration* of effect matters
  - Intended and *unintended* effects of interventions

- How does evaluating replacement endpoints impact the public?

Response: Need “*reliable*” as well as “*timely*” evaluation …not simply “*a choice*”; rather, “*an informed choice*”
• **Biological Activity:** Hemodynamic Measures: 
  \[ PVRI, \: mPAP, \: CO \]
  \[ NT-proBNP \]

• **Clinical Meaningful Benefit**

  ~ **Functions:** Ability to conduct normal activities
  ~ Ability to walk, Ability to engage in recreational activities,
  ~ Ability for self care, Risk of syncope
  ~ Time in hospital or missing school (overall, or cause specific)

  ~ **Feels:**
  ~ Chest pain, breathlessness, fatigue, dizziness

  ~ **Survives**
  ...Physician or Observer administered & PROs...
Direct Measures of ‘Feels, Functions, Survives’ in PAH

~ Overall survival ~ 6MWD @ 48 wks ~ Syncope (freq. & severity)
~ NYHA Functional Class (1-2 vs. 3-4) ~ Clinician Global Measures
~ Level of successful social interaction with peers (mod. CAMPHOR)
~ Days school missed for health-related reasons; Everyday living skills
~ Symptoms: SF-36, Borg Dyspnea Score, Pain Measures

Composites of measures of ‘Feels, Functions and Survives’:

~ (E.g. Acute Coronary Syndrome: CV Death, Stroke, MI)
  ✓ PAH: Death, L.T., PAH Hosp, (NYHA↑ & 6MWT↓)

~ (E.g. CABP: Cough, Pleuritic chest pain, Dyspnea, Sputum Prod)
  ✓ PAH: Chest pain, Dyspnea, Fatigue, Dizziness/Syncope
      .....scored as Absent, Mild, Moderate, and Severe....

The endpoint: a) one-point improvement in at least two symptoms
       &   b) no worsening of any other symptoms, at day TBD


* Fleming TR, Powers JH: Biomarkers and Surrogate Endpoints in Clinical Trials Statistics in Medicine 2012; 31: 2973-2984
**Principles & Insights**

“A Correlate does not Make A Surrogate”


* Fleming TR, Powers JH: Biomarkers and Surrogate Endpoints in Clinical Trials *Statistics in Medicine* 2012; 31: 2973-2984
Challenge Questions

~ If an outcome measure, such as a biomarker, is expected to be sensitive to differences between the effects of the control regimen and the experimental regimen, is this condition sufficient to justify its use as the primary endpoint in a registration trial?

…Hemodynamics, 6MWD in Pediatric PAH…

~ If responders live longer than non-responders, does this justify the conclusion that treatment effect on response rate reliably predicts treatment effect on patient survival?

~ Similarly, if biomarker status is strongly correlated with a direct measure of how a patient feels, functions or survives, does this justify a conclusion that treatment effect on biomarker status reliably predicts treatment effect on the direct measure of how a patient feels, functions or survives?
Challenge Questions

~ For proper rigorous validation of a biomarker as a replacement (or surrogate) endpoint, are the following of integral importance?

— An in depth clinical understanding of
  ✓ the causal pathways of the disease process; and
  ✓ intervention’s intended & unintended mechanisms of action;

— Meta-analyses of clinical trials showing the relationship between:
  ✓ the net effect of treatment on the biomarker, and
  ✓ the net effect of treatment on direct measures of how a patient feels, functions and survives?

~ Is it feasible that a biomarker could be validated in the early phases of a clinical development plan, and then be used as a replacement (or surrogate) endpoint in the phase 3 trial that is designed to provide a reliable and interpretable assessment of efficacy?