The Biomarker Evaluation Framework: Recommendations 1 and 2

John Wagner
Committee Member

The findings and conclusions in this slide deck do not represent the endorsement of Merck & Co., Inc.
Outline

- Definitions
- Evaluation Framework with examples
- Discussion of Surrogate Endpoints
- Summary
Definitions - Biomarker

Biomarker: “a characteristic that is objectively* measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a[n]…intervention.” Example: cholesterol level.

*The committee defines “objectively” to mean “reliably and accurately.”

Definitions – Risk Biomarker

- Committee’s Definition: a biomarker that indicates a risk factor for a disease (this includes genetic biomarkers)
- CFSAN Definition: biological indicators that signal a changed physiological state that is associated with the risk of a disease
- The committee uses the term biomarker instead of risk biomarker in order to clearly delineate between biomarkers and risk factors, both of which are precisely defined terms
Definitions – Surrogate Endpoint

Surrogate Endpoint: “a biomarker that is intended to substitute for a clinical endpoint. A surrogate endpoint is expected to predict clinical benefit (or harm or lack of benefit or harm) based on epidemiologic, therapeutic, pathophysiologic, or other scientific evidence.” Example: blood pressure for trials of several classes of antihypertensive drugs.

Definitions – Clinical Endpoint

Clinical Endpoint: “a characteristic or variable that reflects how a patient [or consumer] feels, functions, or survives.” Example: death.

Recommendation: Biomarker Evaluation Framework

The biomarker evaluation process should consist of the following three steps:

- **Analytical validation**: analyses of available evidence on the analytical performance of an assay;
- **Qualification**: assessment of available evidence on associations between the biomarker and disease states, including data showing effects of interventions on both the biomarker and clinical outcomes; and
- **Utilization**: contextual analysis based on the specific use proposed and the applicability of available evidence to this use. This includes a determination of whether the validation and qualification conducted provide sufficient support for the use proposed.
Biomarker Evaluation Framework

- Discovery Development
- Validation
- Qualification: Evidentiary Assessment
- Utilization
# Biomarker Evaluation Framework: Analytical Validation

## TABLE 3-1 Sources of Variability in Biomarker Measurements

<table>
<thead>
<tr>
<th>Preanalytical Sources of Variability</th>
<th>Analytical Sources of Variability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Biological</strong></td>
<td><strong>Sample Collection</strong></td>
</tr>
<tr>
<td>Sociodemographics (including age and gender)</td>
<td>Mislabeling</td>
</tr>
<tr>
<td>Posture</td>
<td>Duration of tourniquet application</td>
</tr>
<tr>
<td>Exercise</td>
<td>Strength of collection vacuum</td>
</tr>
<tr>
<td>Meals/fasting status</td>
<td>Size of needle gauge</td>
</tr>
<tr>
<td>Diet</td>
<td>Dead volume in catheters/collection tubes</td>
</tr>
<tr>
<td>Diurnal biohythm</td>
<td>Anticoagulants</td>
</tr>
<tr>
<td>Seasonal biohythm</td>
<td></td>
</tr>
<tr>
<td>Concurrent diseases</td>
<td></td>
</tr>
<tr>
<td>Concurrent medications</td>
<td></td>
</tr>
<tr>
<td>Overall health/preexisting disease</td>
<td></td>
</tr>
<tr>
<td>Gastrointestinal motility</td>
<td></td>
</tr>
<tr>
<td>Anesthesia/surgical intervention</td>
<td></td>
</tr>
<tr>
<td>Stress</td>
<td></td>
</tr>
<tr>
<td>Pregnancy</td>
<td></td>
</tr>
<tr>
<td>Menstrual cycle</td>
<td></td>
</tr>
<tr>
<td>Dehydration</td>
<td></td>
</tr>
<tr>
<td>Kidney function</td>
<td></td>
</tr>
<tr>
<td>Body composition (obesity)</td>
<td></td>
</tr>
<tr>
<td><strong>Note:</strong> Biomarker tests need to be reliable, reproducible across multiple laboratories and clinical settings, and maintain adequate sensitivity and specificity before data based on them can be used in subsequent evaluation steps.</td>
<td></td>
</tr>
</tbody>
</table>

---

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

Advising the nation / Improving health
Case Study: Tumor Size

- Tumor size is a variously defined biomarker of efficacy of cancer therapeutics using tumor diameter, tumor volume, or tumor mass, as measured by a variety of platforms and techniques, including magnetic resonance (MR), computed tomography (CT), and positron emission tomography (PET).

- Different contrast agents and different protocols may be used, all of which affect the precision of measurement.

- Analytical validation is complicated by multiple imaging platforms and other assay performance issues.

http://imaging.cancer.gov/imaginginformation/cancerimaging/page4#
Biomarker Evaluation Framework: Qualification

Qualification requires:
(1) Evaluation of the nature and strength of evidence regarding whether a biomarker is on a causal pathway in the disease, and
(2) Assembly of available evidence demonstrating that interventions targeting the biomarker impact the clinical endpoints of interest.
Case Study: C-Reactive Protein

- In observational studies, CRP is an independent predictor of future vascular events, including myocardial infarction, ischemic stroke, peripheral vascular disease, and vascular death.
- In the qualification step of biomarker evaluation, evidence is found for CRP’s prognostic value, but not for evidence of use as a surrogate endpoint.

http://www.nist.gov/cstl/analytical/index.cfm
Biomarker Evaluation Framework: Utilization

Utilization is a contextual analysis based on the specific use proposed and the applicability of available evidence to this use.

This includes a determination of whether the validation and qualification conducted provide sufficient support for the use proposed.

Strong evidence and a compelling context are needed for the utilization of a biomarker as a surrogate endpoint.
Case Study: Troponin

- Use of troponin as a biomarker in acute settings is ubiquitous as a method to diagnose myocardial infarction (MI)
- Troponin can be elevated due to a variety of chronic heart conditions, inflammatory conditions, side effects from drugs, or organ failures
- Use of troponin levels as a surrogate endpoint for interventions in these situations is not yet possible due to a dearth of evidence
- Decisions made in the utilization step of biomarker evaluation will reflect the state of the evidence.

http://www.cdc.gov/heartdisease/heart_attack.htm
Separation of Utilization from Qualification

- Subjective decisions may be needed in the utilization step.
- The objective, data-gathering and analysis aims of qualification should be conceptually separate from the subjective, context-dependent decisions in the utilization step.
- “The committee’s framework has three distinct yet interrelated steps; they are not necessarily separated in time (i.e., some of the steps may occur concurrently) and conclusions in one step may require revisions or additional work in other steps”
Further Biomarker Evaluation Recommendations

- For biomarkers with regulatory impact, the FDA should convene expert panels to evaluate biomarkers and biomarker tests.
- Initial evaluation of analytical validation and qualification should be conducted separately from a particular context of use.
- The expert panels should reevaluate analytical validation, qualification, and utilization on a continual and a case-by-case basis.
Case Study: Blood Levels of Beta-Carotene

β Studies have consistently shown that diets rich in fruit and vegetables are associated with a reduced risk of chronic diseases such as heart disease and cancer.

β Years of epidemiological studies suggested that blood levels of β-carotene were associated with lower incidence of cardiovascular disease and cancer, leading many to believe that β-carotene was responsible for the lower risk.

β However, definitive clinical trials showed that this hypothesis was incorrect and that supplementation with β-carotene did not lower risk for cancer or cardiovascular disease.
Relevance to Drugs, Biologics, and Devices – Surrogate Endpoints

**Successes**
- Blood pressure as a surrogate endpoint for cardiovascular disease clinical endpoints
- HIV-1 RNA levels as an indicator of complete viral suppression for HIV interventions

**Failures**
- Arrhythmia suppression as a surrogate endpoint for reduction of cardiac sudden death
- LDL reduction through hormone replacement therapy as a surrogate endpoint for cardiovascular disease clinical endpoints
Recommendation: Biomarker Evaluation Framework

The biomarker evaluation process should consist of the following three steps:

- **Analytical validation**: analyses of available evidence on the analytical performance of an assay;
- **Qualification**: assessment of available evidence on associations between the biomarker and disease states, including data showing effects of interventions on both the biomarker and clinical outcomes; and
- **Utilization**: contextual analysis based on the specific use proposed and the applicability of available evidence to this use. This includes a determination of whether the validation and qualification conducted provide sufficient support for the use proposed.
Further Biomarker Evaluation Recommendations

- For biomarkers with regulatory impact, the FDA should convene expert panels to evaluate biomarkers and biomarker tests.
- Initial evaluation of analytical validation and qualification should be conducted separately from a particular context of use.
- The expert panels should reevaluate analytical validation, qualification, and utilization on a continual and a case-by-case basis.
Backups
Blood Pressure

“Even as one of the most well-established surrogate endpoints, an effect on blood pressure may not fully capture the benefit—or risk—of an intervention.”

For example:
- Blood pressure is a good surrogate endpoint for primary cardiovascular disease endpoints (mortality and MI) for diuretics and alpha-adrenergic blocker drugs
- Blood pressure is a poor surrogate for secondary cardiovascular disease endpoints for alpha-adrenergic blockers; for example, these drugs raise risk for congestive heart failure
HIV-1 RNA as a surrogate endpoint for clinical endpoints in HIV infection

“Complete viral suppression [tracked using measurements of HIV-1 RNA] often leads to durable suppression, perhaps because of the lower risk of development of viral resistance mutations in patients without replicating virus.”

“The value of HIV-1 RNA as a surrogate in settings where suppression of HIV-1 RNA is partial is much more problematic…”
Arrhythmia Suppression

Researchers hypothesized that suppression of ventricular arrhythmias after myocardial infarction would reduce the rate of death.

Studies were initiated using arrhythmia suppression as the surrogate endpoint.

Scientists were so confident in their hypothesis that many of them believed that randomizing patients to a study drug or a placebo would be unethical.

More than 200,000 people eventually took these drugs each year after FDA approval.

The Cardiac Arrhythmia Suppression Trial (CAST) later found that these drugs resulted in higher incidence of sudden cardiac death: 56 patients in the encainide and flecainide groups died, compared to 22 patients in the placebo group. Later data confirmed that patients taking moricizine were also at increased risk for death.
LDL is not a Generic Surrogate Endpoint: Lesson from Hormone Replacement Therapy

“Postmenopausal hormone replacement therapy (HRT) was thought to protect women from cardiovascular disease based on both observational, epidemiologic data and the apparent beneficial effects of estrogen on lipoproteins and other cardiovascular disease biomarkers.”

“After several clinical trials, HRT was discovered to raise mortality from cardiovascular events and have other adverse unexpected effects.”

LDL and HDL cholesterol will be discussed in the next presentation
“Even in the best of circumstances, it is possible for surrogate endpoints to be misleading by either overestimating or underestimating an intervention’s effect on clinical outcomes.”

Failures of Surrogate Endpoints

Biological Complexity Leads to Many Opportunities for Error

- Pure or multicomponent substance or intervention
- Component 1
- Component 2
- Component 3
- Component n

HEALTH STATUS

Biological Pathways 1, 2, 3, ...n

Surrogate Endpoint

Outcome 1
Outcome 2
Outcome 3
Outcome n