FDA Perspectives on Novel Kidney Biomarker Tests

2013 Arnold O. Beckman Conference
Novel Biomarkers of Kidney Disease: False Dawn or New Horizon?
November 5-6, 2013

Courtney H. Lias, Ph.D.
Office of In Vitro Diagnostics and Radiological Health
Food and Drug Administration

Kidney Disease

• Millions of people and their families in the U.S. are affected by kidney disease
• New tools for the diagnosis and treatment of renal diseases are sorely needed
• Novel biomarkers may provide for better management of these patients

Biomarker ≠ Test

• Biomarker – A characteristic
  (e.g., increase in creatinine due to renal damage)
• Biomarker test – The method of detecting the characteristic.
  (e.g., can the creatinine test accurately and reliably measure creatinine?)
• When biomarker tests are applied for a purpose, safety and effectiveness for that purpose (e.g., for clinical use, for drug development) must be evaluated
  (e.g., can the creatinine test be used to diagnose renal failure? How sensitive and specific is it?)
Biomarker Tests

- Drug Discovery
  - Identify Drug Targets
- Pre-clinical Research
  - PK, PD
  - Toxicity
  - Effect
- Clinical Uses
  - Screen for disease
  - Classify disease
  - Monitor a disease
  - Stratify population by severity/risk
  - Select patients for a specific therapy
  - Identify drug responders/non-responders
  - Identify patients at high risk for adverse events
  - Select drug dose for safety/efficacy
  - Act as a surrogate endpoint

Each context has its own requirements for studies or trials to demonstrate analytical and clinical validity of the test.

Device Regulation

FDA Regulates IVDs by the intended use and the risk of an incorrect result:

**Class I** – Low risk – Usually exempt from Premarket FDA review

**Class II** – Moderate risk – requires a predicate device - requires 510(k) clearance

**Class III** – High risk and novel intended uses - requires premarket approval (PMA)
Intended Use

The risk of an IVD is based on the consequences of a false result

Examples:
High risk – HIV, tuberculosis
Lower risk – Calcium, pregnancy

Premarket Review

All IVDs must establish adequate:

Analytical performance
• How accurately does the test measure the analyte?
• How reliably?

Clinical performance
• How reliably does the test measure the clinical condition?

Labeling (21 CFR 809.10)
• Adequate instructions for use
• Intended use, directions for use, warnings, limitations, interpretation of results, performance summary

Analytical Performance

• Repeatability/Reproducibility
  • Will I get the same result in repeated tests over time?
  • Will I get the same result as someone else testing the same sample?

• Accuracy
  • Will I get results that are the same as “Truth”?
  • “Truth” – may be a reference method, clinical endpoint, predicate device, etc…

• Limit of Detection

• Potential Interferences / Cross-Reactivity

• Cross-contamination / Carry-over
• etc…
Clinical Claims

- Diagnostic Markers
- Prognostic Markers

Diagnostic Markers

- Identify/classify disease
  - e.g., diagnosis of renal failure
- Screen for disease in asymptomatic people
  - e.g., screen for kidney cancer
- Monitor for recurrence/status
  - e.g., monitor for allograft rejection
- Select patients who will respond to a specific drug
  - e.g., select responders for a certain chemotherapy drug

Prognostic Markers

- Stratify disease by severity/risk
  - e.g., Assess CKD patients at high risk of complications
- Predict disease development/recurrence/progression
  - e.g., detect “sub-clinical” AKI
- Identify patients at risk for adverse event/drug reaction
  - e.g., Identify patients who might suffer adverse event from a certain drug

Note: Look at relative vs. absolute risk?
Endpoints

- Surrogate endpoints
  - Currently few acceptable surrogate endpoints
  - Difficult to develop/validate
    - “Holy Grail”
- Clinical diagnosis
  - e.g., compare new test to diagnosis by current clinical practice/guidelines
- Longitudinal outcome
  - e.g., compare prognostic marker to outcomes (death, progression)
- Special attention to cut-offs/decision points
  - e.g., “positive” above 45 ng/mL

Test Effectiveness

- Sensitivity
  \[ \text{Sens} = \frac{TP}{TP + FN} \]
  How likely is the test to detect the presence of a disease in someone with the disease?

- Specificity
  \[ \text{Spec} = \frac{TN}{TN + FP} \]
  How likely is the test to detect the absence of a disease in someone without the disease?

- Positive Predictive Value (PPV)
  \[ \text{PPV} = \frac{TP}{TP + FP} \]
  How likely is someone with a positive test result to actually have the disease?

- Negative predictive value (NPV)
  \[ \text{NPV} = \frac{TN}{TN + FN} \]
  How likely is someone with a negative test result to actually not have the disease?

<table>
<thead>
<tr>
<th>Disease Prevalence in the Intended Test Population</th>
<th>Probably of having the Disease if you have a Positive Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1%</td>
<td>1.9%</td>
</tr>
<tr>
<td>1%</td>
<td>16%</td>
</tr>
<tr>
<td>10%</td>
<td>68%</td>
</tr>
<tr>
<td>20%</td>
<td>83%</td>
</tr>
<tr>
<td>50%</td>
<td>95%</td>
</tr>
</tbody>
</table>

- Predictive Value is not intrinsic to the test - it depends on the prevalence of disease.
- The results of a study may not apply to all situations if there are different prevalence rates between the study population and the test population.
- If prevalence is very low, even if sensitivity and specificity are high, many positive test results will be false positives.
Study Population Matters

"Symptomatic" Patients

<table>
<thead>
<tr>
<th>Disease +</th>
<th>Disease -</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test +</td>
<td>PPV = 0.63</td>
</tr>
<tr>
<td>Test -</td>
<td>NPV = 0.97</td>
</tr>
</tbody>
</table>

Sens = 0.71 Spec = 0.91

Study Population Matters

Non-representative sample

<table>
<thead>
<tr>
<th>Disease +</th>
<th>Disease -</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test +</td>
<td>PPV = 0.57</td>
</tr>
<tr>
<td>Test -</td>
<td>NPV = 0.97</td>
</tr>
</tbody>
</table>

Sens = 0.71 Spec = 0.84

Example 1

Novel biomarker test for diagnosis of acute kidney injury (AKI)
- Study compares test results in healthy patients to those with known AKI
- Positive and negative predictive value are excellent, everyone is excited
Example 1
Novel biomarker test for diagnosis of acute kidney injury (AKI)
• Study compares test results in healthy patients to those with known AKI
• Positive and negative predictive value are excellent, everyone is excited

CAUTION:
• PPV and NPV will be overestimated in this study design
• Different performance observed when used in the intended use population (testing all patients suspected of having AKI)
  • This population will include those who do and those who don’t have AKI
  • Mimic conditions could cause false positives
  • Undiagnosed patients (e.g., earlier stages, ambiguous symptoms) may give false negatives
  • Will never know this by that study design

Example 2
Novel biomarker test for further stratification of RIFLE “R” patients for risk of progression to severe kidney injury
• Retrospective study compares predictive value of biomarker test to outcome for each patient
• Positive and negative predictive value are excellent, everyone is excited

CAUTION:
• Study used “convenience samples” from a study intended to assess general outcomes in an ICU environment
• Sub-study only included patients that had 2 creatinine measurements in their record, excluded all others
• Different performance observed when used in the intended use population
  • Selection bias not controlled for – this design favors patients more likely to have 2 creatinine tests ordered in this time frame
  • Baseline risk of included vs. excluded patients may not be the same
  • Inclusion/exclusion criteria for this ICU study may not be appropriate for this test
  • PPV and NPV (or even relative risk) cannot be accurately calculated
Study Population Matters

• To demonstrate the effectiveness of a biomarker test, it should be studied in the population in which the test will be used
• Test result is compared to “Truth” standard (e.g., clinical diagnosis, outcome, etc.)
• The entire intended test population (including mimic and confounding conditions) should be represented (ideally proportionally) in the study
• Studies in wrong population provide biased estimates of test performance

Challenges to Innovation

• Manufacturer has not submitted the device to FDA
• Manufacturer has not started the clinical study
  • Studies approved but Manufacturer decides not to move forward or to delay
  • Poor study design
• Bad data
  • Cutoffs not independently validated
  • Device doesn’t work
  • “Fixing” the reference method
  • Data exclusion

FDA Pre-Submissions

• Mechanism to get FDA feedback on clinical and analytical studies
• Informal and flexible
• Can help to make test development more efficient by focusing efforts of industry

Draft Guidance for Industry and FDA Staff Medical Devices: The Pre-Submission Program and Meetings with FDA Staff
(http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm310375.htm)
Summary

• New tests for diagnosis and prognosis of renal disease will be beneficial for patients

• Proper validation of a biomarker test for a particular purpose is critical

• Study design matters – need unbiased estimates of effectiveness and reliability

• Come talk to us!

Thank you

courtney.lias@fda.hhs.gov