Navigating NASH biomarkers: From discovery to clinical practice

May 6, 2017

John J. Sninsky, PhD
Chief Scientific Officer
CareDx

2nd International Workshop on NASH Biomarkers
Meeting
Washington, D.C.
May 5-6, 2017
Disclosure

- Employee and stock with CareDx
Outline

• Background
• Research-grade versus clinical-grade biomarker assays
• Diagnostic metrics
• Evidence development
• Diagnostic development pathway
• Context of Use
• Evolving regulatory landscape
• Reimbursement
• Summary
Functional definition of a diagnostic test:

Translation of the complex and fuzzy spectrum of pathophysiology of disease into practical actionable next steps for clinicians to improve patient outcomes
Sea Change in Clinical Diagnostics

- Increased complexity of our understanding of disease
  - Multiple underlying etiologies
- Formal phased development of diagnostic tests similar to drug development has been adopted
- High quality evidence needs to be provided by test service (LDT) and test kit providers
- Clinical utility now required for reimbursement instead of only clinical validity as in past
Biomarker ‘Discovery’ and ‘Translation’ Have Different and Discrete Objectives: equally valuable

• **Biomarker Discovery (Exploratory Walk)**
  – Biomarker-centric
  – Promises key insights into fundamental underlying pathophysiology
  – Plethora of biologically plausible biomarkers
  – Benefits from deep understanding of biology
  – Correlations and group diagnostic metrics suffice

• **Biomarker Translation for clinical practice (Directed Path)**
  – Clinical question-centric
  – Promises improved patient management
  – Few biomarkers that merit prioritization
  – Benefits from translation and diagnostic development path knowledge
  – Predictive values are most important for individual patients
Clinical-grade vs Research-grade Assays

- Biomarkers are not validated, biomarker assays are validated
- Clinical-grade assays are much more than just testing clinical samples; require rigorous analytical validation
- Clinical-grade assays have to be of highest quality because they inform critical patient management decisions

<table>
<thead>
<tr>
<th>Research-grade</th>
<th>Clinical-grade</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference materials</td>
<td>Internal specimens / External specimens</td>
<td>External standards; orthogonal technology validation</td>
</tr>
<tr>
<td>Methods-based proficiency</td>
<td>Rarely used</td>
<td>Performed regularly</td>
</tr>
<tr>
<td>Information tracking systems</td>
<td>Sometimes used</td>
<td>Always use version control LIMS; some integration with EMRs</td>
</tr>
<tr>
<td>Bioinformatic analysis</td>
<td>Open source combined with subscription/license; frequently changing; &amp; early adoption of new software/algorithms</td>
<td>Open source combined with subscription/license; use mature software and CDS Locked and change requires re-validation</td>
</tr>
<tr>
<td>Validation of steps in process</td>
<td>Sometimes</td>
<td>Always</td>
</tr>
<tr>
<td>Documentation</td>
<td>No design control Little documentation</td>
<td>Yes Extensive</td>
</tr>
</tbody>
</table>
Statistical Metrics for Test Performance

• No single statistical measure provides necessary and sufficient insight
• Predictive values (NPV and PPV) are more important than sensitivity and specificity
• ROC curves have a qualified contribution
• Methods based on risk stratification have recently been proposed to compare models
  – reclassification statistic
• Bayesian models for diagnostic test performance provide key insights (conditional probabilities; likelihood ratios)
• Explore integration of clinical variables and molecular biomarkers in multivariate analysis
• Risk prediction analysis is an active area in multiple disciplines especially cardiovascular disease and includes proposed standardization
  – Overlap of cardiovascular disease and NASH risk presents an opportunity to discuss best practices

ROC is a Qualified Diagnostic Metric

- Threshold independent technique to visualize dichotomous diagnostic test performance
  - ROC plots false positives (1-specificity) versus true positives (sensitivity) for every possible cutoff including regions not clinically relevant
  - Requires highly accurate and related reference method to be informative
  - A test with high sensitivity may have an identical or similar AUC to a test with high specificity
  - Weights false positives and false negatives equally
  - Does not address predictive values critical to ruling-in and ruling-out a diagnosis
  -Insensitive to changes in absolute risk of tests compared
Diagnostic Metrics

- **Sensitivity** is defined as the probability of getting a positive test result in subjects with the disease.
- **Specificity** represents the probability of a negative test result in a subject without the disease.
- Neither sensitivity nor specificity are influenced by disease prevalence.

- **Positive predictive value (PPV)** defines the probability of having the state/disease of interest in a subject with positive result.
- **Negative predictive value (NPV)** describes the probability of not having a disease in a subject with a negative test result.
- Unlike sensitivity and specificity, predictive values are largely dependent on disease prevalence in tested population.
- PPV increases while NPV decreases with the increase of prevalence of the disease in a population.
- Predictive values from one study should not be transferred to another setting with a different prevalence of the disease in the population.
Impact of Prevalence on Predictive Value

- 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 discovery and validation studies or development and clinical practice populations
- If prevalence is very low even if sensitivity and specificity are high, many positive test results will be false positives
- Context of use determines whether PPV or NPV is critical

### Disease Prevalence in the Intended Test Population vs. Probability of having the Disease if you have a Positive Result

<table>
<thead>
<tr>
<th>Disease Prevalence</th>
<th>Probability 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>

Assumes a 95% sensitive and 95% specific test
Hierarchy of Evidence

- Meta-analysis of randomized control trials
  - Highest level of evidence
- Randomized control trial
  - Prospective in design
  - High level of evidence
  - Post hoc analysis possible (e.g. pre-specified, avoid subgroups, use primary endpoint)
  - Adaptive designs show promise for precision medicine
- Observational cohort
  - Prospective in design
  - Less likely to have masked bias
- Case control
  - Retrospective in design
  - Susceptible to masked bias (e.g. survivorship, selection, ascertainment, drug treatment)
  - Most GWAS studies use this design
- Anecdotal study
  - Replication rarely reported

Where are ‘adaptive’ trials, registries, and EMR data positioned in this hierarchy?
Evolution of Thinking About Evidence


The Randomized Registry Trial — The Next Disruptive Technology in Clinical Research?
Michael S. Lauer, M.D., and Ralph B. D’Agostino, Sr., Ph.D.

The randomized trial is one of the most powerful tools clinical researchers possess, a tool that enables them to evaluate the effectiveness of new (or established) therapies while accounting for the effects of unmeasured confounders and selection bias by indication. Randomized trials, especially huge megatrials, have transformed medical practice. Thanks to randomized trials, we no longer, for example, treat acute myocardial infarction with thrombolytics and streptokinase. Instead we use rapid revascularization, angioplasty, and antithrombotic agents, and during long-term follow-up we routinely prescribe statins, beta-blockers, and angiotensin-converting-enzyme inhibitors. But the reputation of randomized trials has suffered of late, owing to reasonable concern about excess complexity, expense, and time required to recruit study participants, as well as inadequate representativeness. What good are trials if the results aren’t applicable to real-world patients and if, because of excessive expense they can be used to answer only a tiny fraction of our important clinical questions?

One possible solution is to look to observational registries for answers. Over the past 20 to 30 years, a number of professional societies, government agencies, private corporations, and independent researchers have established high-quality registries that collect standardized data from patients seen in a variety of settings. In cardiovascular medicine, for example, registries in the United States and abroad have collected vast amounts of data from patients with acute coronary syndromes, stable coronary disease, and heart failure, as well as from patients with rare diseases such as hypertrophic cardiomyopathy and patients referred for surgery, percutaneous invasive procedures, and device implantation. Investigators and public health officials use registries to describe practice patterns and trends, to identify outliers, and to detect safety signals. They often use registries to assess comparative effectiveness, too, but are forced to admit that purely observational findings may not be internally valid owing to the absence of randomization.

As debates about comparative-effectiveness research have intensified over the past few years, we find ourselves in a kind of intellectual trap: yes, in theory we would like to conduct more randomized trials, but in practice they are too complex and difficult to apply to many clinical questions. And, yes, in theory we could answer many questions at

Evolution of Thinking About Evidence

The Randomized Registry Trial — The Next Disruptive Technology in Clinical Research?

Michael S. Lauer, M.D., and Ralph B. D’Agostino, Sr., Ph.D.

The randomized trial is one of the most powerful tools clinical researchers possess, a tool that enables them to evaluate the effectiveness of new (or established) therapies while accounting for the effects of unmeasured confounders and selection bias by randomization. Randomized trials, especially large trials, have transformed medical practice. Thanks to randomized trials, we do longer, for example, treat acute myocardial infarction with “make and not etalons. Instead we use rapid revascularization, anticoagulants, and antplatelet agents, and during long-term follow-up we statistically prescribe statins, beta-blockers, angiotensin-converting enzyme inhibitors, the aggregation of randomized trials as suffered of late, owing to questionable concern about excess complexity, expense, and time required to recruit study participants, as well as inadequate representativeness. What good are trials if the results aren't applicable to real-world patients and if, because of excessive expense they can be used to answer only a tiny fraction of our important clinical questions?

One possible solution is to look to observational registries for answers. Over the past 20 to 30 years, a number of professional societies, government agencies, private corporations, and independent researchers have established high-quality registries that collect standardized data from patients seen in a variety of settings. In cardiovascular medicine, for example, registries in the United States and abroad have collected vast amounts of data from patients with acute coronary syndromes, stable coronary disease, and heart failure, as well as from patients with rare diseases such as hypertrophic cardiomyopathy and patients referred for surgery, percutaneous invasive procedures, and device implantation. Investigators and public health officials use registries to describe practice patterns and trends, to identify outliers, and to detect safety signals. They often use registries to assure comparative effectiveness, too, but are forced to admit that purely observational findings may not be internally valid owing to the absence of randomization.

As debates about comparative-effectiveness research have intensified over the past few years, we find ourselves in a kind of intellectual trap yes, in theory we would like to conduct more randomized trials, but in practice they are too complex and difficult to apply to many clinical questions. And, yes, in theory we could answer many questions at


Evolution of Thinking About Evidence

Lauer 

Evolution of Thinking About Evidence
Psaty & Breckenridge NEJM 370, 2165 (2014)

The Randomized Registry Trial — The Next Disruptive Technology in Clinical Research?
Michael S. Lauer, M.D., and Ralph B. D’Agostino, Sr., Ph.D.

Thank you. The comments I am about to make are my personal views. They are not the views of the National Heart, Lung, and Blood Institute or the US Department of Health and Human Services. I have no conflicts of interest disclosure.

There once was a king who had 2 servants. One day, the king asked the servants to take buckets down to the well, which was located just outside the palace. He told them to draw water into the buckets and bring them back to him. The 2 servants did exactly as they were told, but when they reached the well, they noticed something rather strange. The bucket they had been using was porous, full of holes. The king said, “The king must have made a mistake. It may be a bad idea to use a bucket with holes.” The second servant said, “I agree that it may not make sense, but nonetheless I will do the king asked.” The servant drew water into the bucket, sure enough within a few minutes all the water had drained out. When the servants returned to the king, he asked them whether they had carried out his orders. The first servant declared that he discovered the bucket had holes, but he


Staged Diagnostic Test Development

- **Analytical validity** refers to how well the test predicts the presence or absence of a biomarker. In other words, can the test accurately detect whether a specific biomarker is present or absent?

- **Clinical validity** refers to how well the biomarker being analyzed is related to the presence, absence, or risk of a specific disease.

- **Clinical utility** refers to whether the biomarker can provide information about diagnosis, treatment, management, or prevention of a disease that will be helpful to a patient, healthcare provider, or family member.

With recent coding changes by Medicare and decisions by most private payors, reimbursement requires strong evidence of clinical utility.

Different Kinds of Diagnostic Tests (context of use)

Prognostic biomarker
• A biomarker that predicts a clinical outcome regardless of treatment and includes element of time

Predictive biomarker
• A biomarker that changes in response to treatment, and predicts a clinically relevant event or process, and could be used to identify subsets of patients who are most likely to respond to treatment

Clinical end point
• A characteristic or variable that reflects how a patient feels, functions, or survives

Surrogate end point
• A biomarker that can substitute for a clinical end point based on biological rationale; accurately predicts a clinical end point and the effect of a given treatment on the clinical end point

Pharmacogenomic
• A biomarker that provides information on drug metabolism

Context of use drives intended use
Why is Understanding Regulatory Oversight and Reimbursement Essential?

• Even though diagnostics only makes up about 4% of healthcare expenditure, diagnostics informs how 50% of spend directed
• Concern that important medical insights are not being translated in a timely manner to patient care
• Accelerated translation of discoveries into practice of medicine requires ‘directed path’ instead of ‘exploratory walk’
• High quality, evidence-supported clinical-grade biomarker assays will not be used unless reimbursed at equitable level
• If clinical-grade assays are not value priced, innovation from government and private industry will be stifled
Two Paths for Regulatory Oversight

The FDA device classification for a regulated diagnostic device will depend on the perceived risk associated with the diagnostic device.


1 Clinical Laboratory Improvement Amendments of 1988 (CLIA)
FDA Medical Device Process

Long, costly and arduous path without assurance of reimbursement

- Concept & Design
- Pre-Clinical Development
- Clinical Trials
- FDA Review
- Reimbursement Assignment

- Pre-Submission Process

- 510(k): 3-9 mon
- IDE (PMA): 9-36 mon

- FDA submission

- 510(k) 3-5 mon

- 121 day average (2011)

- 0-90 day average (2011)

- Limited Patient Access

- Broad Patient Access

- 0-150 day average (2010)

- 0-90 day average (2010)
## In Vitro Diagnostic Test Regulation

<table>
<thead>
<tr>
<th></th>
<th>CLIA</th>
<th>CLIA-CAP</th>
<th>CLIA-CAP-NYDOH</th>
<th>FDA ¹</th>
<th>Payor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Analytical Validation</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Clinical Validation</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Clinical Utility</strong></td>
<td>NR³</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Comparative-effectiveness</strong></td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Required</td>
</tr>
</tbody>
</table>

- Analytical Validation
- Clinical Validation
- Clinical Utility
- Comparative-effectiveness

**Even though a multi-path regulated system has developed, reimbursement alone determines what diagnostic tests are used**

1. FDA clears and approves IVD kits and single laboratory LDTs if submitted
2. Complementary and Companion
3. Not required

---

¹ FDA clears and approves IVD kits and single laboratory LDTs if submitted
² Complementary and Companion
³ Not required
Proposal for Phased and Timely Introduction Based on Benefit/Risk Ratio of Patients: Adaptive licensing

<table>
<thead>
<tr>
<th>Conventional</th>
<th>‘Adaptive’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single gating licensing decision</td>
<td>Phased licensing decisions</td>
</tr>
<tr>
<td>Risk of expanded use</td>
<td>Risk addressed by surveillance and monitoring</td>
</tr>
<tr>
<td>RCT only</td>
<td>Adopt new evidence (EMR)</td>
</tr>
<tr>
<td>Broad population</td>
<td>Targeted subset of population</td>
</tr>
<tr>
<td>Focus on licensing</td>
<td>Focus on patient access</td>
</tr>
<tr>
<td>Open utilization</td>
<td>Specified utilization</td>
</tr>
</tbody>
</table>

Though proposed for drug approvals, equally applicable to diagnostics and reimbursement


European Medicines Agency, MIT Center for Biomedical Innovation with some FDA co-authors.
Biomarkers Support Drug Development & Approval

Biomarkers can be used in clinical development to:

- Better balance subjects in arms of trial
- Identify subjects at differential risk of disease
- Identify subjects at differential risk of adverse events (safety)
- Identify patients with sub-clinical disease
- Predict treatment efficacy
- Encourage increased adherence to medication\(^1\)
- Monitor response to treatment

Prototype Drug and Diagnostic Co-Development

- Idealized co-development alignment strategies have been proposed; however, co-development pipelines vary
- Different ‘points of entry’ of diagnostic
Rich Legacy of Paired Diagnostic Tests and Drugs

Evolving Regulatory Science

• 30 diagnostic assays paired with 17 drugs starting in 2014 (1-10 assays for each drug)
  – Clinical utility matched with drug (PMAs)
  – Different technologies used (e.g. IHC, FISH, PCR, Sanger sequencing, NGS sequencing, etc.)
  – 28 Companion assays
    • 26 IVD kits
    • 2 single laboratory (CLIA) PMA assays (Dec 2014 and Dec 2016)
  – 2 Complementary assays

https://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/InVitroDiagnostics/ucm301431.htm
Types of Predictive Diagnostic Tests

Companion
- Essential for safe and effective use of drug
- Directs practice of medicine
- Examples include (KRAS (cetuximab and panitumumab); EGFR (gefitinib, osimertinib, etc.); BRAF (vemurafenib); PD-L1 IHC 22C3 pharmDx (pembrolizumab (Keytruda) Merck)

Complementary
- Category introduced by FDA August, 2015
- Recognizes complexity of biomarker-drug correlation
- Recognizes that significant fraction of patients without marker may respond
- Perceived to ‘define’ the use rather than ‘restrict’ the use of drug
- Informs practice of medicine decisions
- PD-L1 IHC 28-8 pharmDx (nivolumab (Opdivo) BMS); first test in this category
LDT Bridging Strategy Accelerates Drug & IVD Approvals

<table>
<thead>
<tr>
<th>LDT Key Benefits</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduce Development Timeline and Costs</td>
<td>• Facilitates rigorous and flexible biomarker testing in development</td>
</tr>
<tr>
<td></td>
<td>• Parallel LDT and IVD strategy can bridge biomarker and IVD timelines</td>
</tr>
<tr>
<td></td>
<td>• Option to defer or postpone high IVD development costs to later</td>
</tr>
<tr>
<td>Optimize Clinical Trial Utility</td>
<td>• LDT approach provides test availability earlier in clinical trials than IVDs</td>
</tr>
<tr>
<td></td>
<td>• Early insight can further mitigate development and financial risks</td>
</tr>
<tr>
<td>Launch and Regulatory Risk Mitigation</td>
<td>• Regulatory approval of single laboratory LDT PMA may allow earlier adoption of the drug</td>
</tr>
<tr>
<td></td>
<td>• Provide data for FDA submissions for an approved LDT or an IVD kit</td>
</tr>
</tbody>
</table>
AlloMap® Reimbursement Chronology
FDA cleared IVD

- Strong peer-reviewed evidence of clinical utility and guideline inclusion were key to reimbursement
- 20 gene (11 functional) RT-PCR gene expression test
Take Home Messages

- Even though a multi-path regulated system has developed, reimbursement alone determines what diagnostic tests are used.
- Specific definition of clinical utility varies with stakeholder and remains undefined and undesignated.
- Timely diagnostic test regulatory approval and reimbursement requires ‘directed path’ rather than ‘exploratory walk’.
- Recognition of need and value of improved statistical tools to measure diagnostic test performance.
- Failure to reimburse appropriately will stifle precision medicine and translational science.

What does the future hold for regulated clinical diagnostics?

• Diagnostics determines how most of health care dollar spent and is keystone of precision medicine
• Diagnostic test development much more rigorous than in past and provider investment is substantially increased
• Neither CLIA nor IVD regulatory pathways satisfactorily address present clinical needs
  – IVD too slow to accommodate pace of accumulated evidence for patients with highest unmet need
  – LDTs perceived of in need of additional rigor of development
  – Neither sufficient for reimbursement decisions
• Just proposed oversight of LDT by FDA has sensitized FDA to critical issues in need of improvement and increased CLIA laboratory rigor for those that used lower standard
• Payors will increasingly look to professional organizations due to complexity and dynamic nature of evidence (need increased frequency of professional guidelines)
Summary

- Important to continue to consider that NASH is not one disease entity but multiple entities that have not yet been teased apart (NASHes, not NASH)
- Need to distinguish between equally valuable but different contributions of biomarker ‘discovery’ and ‘translation’ efforts
- Puzzle through performance evaluation when ‘reference method’ (not ‘gold standard’) is inaccurate—outcomes are final arbiter
- Established clear phased path for clinical-grade diagnostic development
- Value reimbursement of diagnostic tests is vital to test utilization
- Evolving landscape of evidence and use of ‘adaptive’ licensing to balance access with safety (different benefit-risk of patient groups)
- An exciting and evolving regulatory landscape is unfolding
- NASH biomarker community to be applauded for biobank and registry efforts
  - NIIMBLE, LITMUS, TARGET-NASH, The Liver Forum pooled placebo study, etc.
- NASH biomarker community efforts would benefit from increased input from diagnostic perspective
Acknowledgements

Thanks to all the patients who participated in randomized controlled trials and subjects who participated in observational studies in past and those in future who provide evidence for healthcare decisions.

Every number is a life™
CAP