



# Molecular Diagnostics and the Quest for Clinical Utility

Steven Gutman, M.D.

# Fryback and Thornbury --1991

- Level 1: Technical efficacy
- Level 2: Diagnostic accuracy efficacy
- Level 3: Diagnostic thinking efficacy
- Level 4: Therapeutic efficacy
- Level 5: Patient outcome efficacy
- Level 6: Societal efficacy

# Fryback and Thornbury --1991

- Level 1: Technical efficacy - *analytical validity*
- Level 2: Diagnostic accuracy efficacy - *clinical validity*
- Level 3: Diagnostic thinking efficacy
- Level 4: Therapeutic efficacy
- Level 5: Patient outcome efficacy - *clinical utility*
- Level 6: Societal efficacy

# Current System -- CLIA

- Analytical validity and quality assured for all
- Non transparent system based on on-site sampling by an operations team

# Current System -- FDA

- Analytical and clinical validity assured for some (test systems commercially marketed as kits)
- Highly transparent system based on rigorous data review

# Current System - Other Controls

- Medical necessity, clinical utility, other evidence that the test works
- Non standardized, non coordinated, leaky system with variable transparency

# Broken System

- SACGHS report (2008): “There are inadequate data on which to base utility assessment and only a few studies have been done of the clinical utility of specific genetic tests. More fundamentally, there has been insufficient analysis of the standard of evidence on which the clinical utility of genetic tests should be evaluated....”

# Broken System

- EGAPP experience (2012): “Test applications are being proposed and marketed based on descriptive evidence and pathophysiologic reasoning, often lacking well-designed clinical trials or observational studies to establish validity and utility, but advocated by industry and patient interest groups.”

# Hayes et al. -- 1996

- Proposed it was appropriate to “standardize the tumor marker information for clinical utility”
- Tumor Marker Utility Grading System (TMUGS) to evaluate clinical utility
- Three critical observations

# Intended Use Matters

- Determine risk
- Screening
- Differential diagnostics
- Prognosis - predict relapse/progression
- Monitor course of disease
- Prognosis - predict response to therapy \*

# The Assay Matters

- “One cannot assume that two assays for the same alteration of the same molecule provide identical results.”
- “Each claim should be based on independent studies that demonstrate the utility of that marker in the manner in which it was tested, rather than on assumptions that one method provides the same correlation with endpoints and outcomes as another”

# Hierarchy of Evidence Matters

- Level 1: Evidence from a single, high-powered, prospective, controlled study that is designed to test marker or evidence from meta-analysis and or overview of level II or III studies.
- Level 5: Evidence from small pilot studies designed to determine or estimate distribution of marker levels in sample populations.

# Study Designs Matter

- Prospective
  - Biomarker stratified (all comers) - MARVEL study
  - Enrichment design (randomization of treatment in biomarker positive only) -- HER2, BRAF, ALK
  - Biomarker strategy (random assignment to an experimental arm that uses biomarker to determine therapy versus a control which does not) - ERCC1 gene expression for NSCLC

# Study Designs Matter

- Prospective - retrospective
  - Archived samples from completed RCT identified (KRAS, Oncotype DX)
  - Study plan developed
  - Repurposed study used to analyze one or more biomarker assays

# Study Designs Matter

- Adaptive designs (I-SPY; the BATTLE study; Jiang, Friedlin, Simon 2007)

# Burgeoning Literature

- Putszai and Hess (2004)
- Sargent et al (2005)
- Simon and Wang (2005)
- Berry (2007)
- Pepe (2008)
- Zhou et al (2008)

# Burgeoning Literature

- Simon, Paik and Hayes (2009)
- Freidlin, McShane, Korn (2009)
- Freidlin, Jiang, Simon (2010)
- Simon (2010)
- Scher, Nasso, Rubin (2011)
- Jonas et al (2012)

# Who Is Helping

- FDA - Guidance documents
- <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM332181.pdf>
- <http://www.fda.gov/downloads/drugs/scienceresearch/researchareas/pharmacogenetics/ucm116689.pdf>
- <http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm262292.htm>

# Who Is Helping

- CDC - Laboratory Practice Program
  - <http://www.cdc.gov/mmwr/pdf/rr/rr5806.pdf>
- CDC - EGAPP
  - <http://www.egappreviews.org/>

# Who Is Helping

- AHRQ in collaboration with Evidence Practice Centers and the Journal of General Internal Medicine
  - Developed and published a Methods Guide for Medical Test Reviews
  - <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=1088>

# Who Is Helping

## National Comprehensive Cancer Network

- NCCN Molecular Testing White Paper; NCCN Molecular Testing White Paper: Effectiveness, Efficiency, and Reimbursement (Engstrom et al. 2011)
  - <http://www.ncbi.nlm.nih.gov/pubmed/19755046>
- NCCN Task Force Report: Evaluating the Clinical Utility of Tumor Markers in Oncology (Febbo et al. 2011)
  - [http://www.nccn.org/JNCCN/supplements/PDF/TumorMarkers\\_Task\\_Force\\_Report.full.pdf](http://www.nccn.org/JNCCN/supplements/PDF/TumorMarkers_Task_Force_Report.full.pdf)

# Who Is Helping

- Institute of Medicine
- Workshops
- Reports

# Who Is Helping

- Center for Medical Technology Policy - Effective Guidance Document for Demonstration of Clinical Validity and Clinical Utility of Molecular Diagnostic Tests in Oncology

# Who Is Helping

- NCI - today's initiative
- Cancer Steering Committee
- Some work in other programs: EDRN, SPORE, etc.

# Who Is Helping

- Grips
- PRoBE
- ReMARK
- STARD

# Who's On First

# No Rule Book

- Woodcock rule: “Given the fact that resources for clinical trials are limited, it is rational to apply a risk-based approach to evidence generation.”
- Take into account where in life cycle the test is discovered: before final clinical study of the drug, as a rescue diagnostic, or as a retrofit
- Recognize and manage stakeholder differences

# Decision Elements

- Mechanistic understanding of biomarker
- Prevalence of conditions of interest (sizing of study)
- Access to archived samples
- Estimates on test performance if known
- Ethics of drug administration
- Costs

# Trade Offs

- Cost
- Time
- Size
- Quality of results
- Return

# CDC National Office of Public Health Genomics

- 2004 Program: Evaluation of Genomic Applications in Practice and Prevention (EGAPP)
- 2005 Working Group

# Goal

- Nonfederal panel
- Establish a systematic process for evidence-based assessment that is specifically focused on genetic tests and other applications of genomic technology

# Work Plan

- Develop high priority recommendations to direct test use
- Create a scientific and administrative process that was transparent, publically accountable, would minimize conflicts of interest and optimize existing evidence review methods

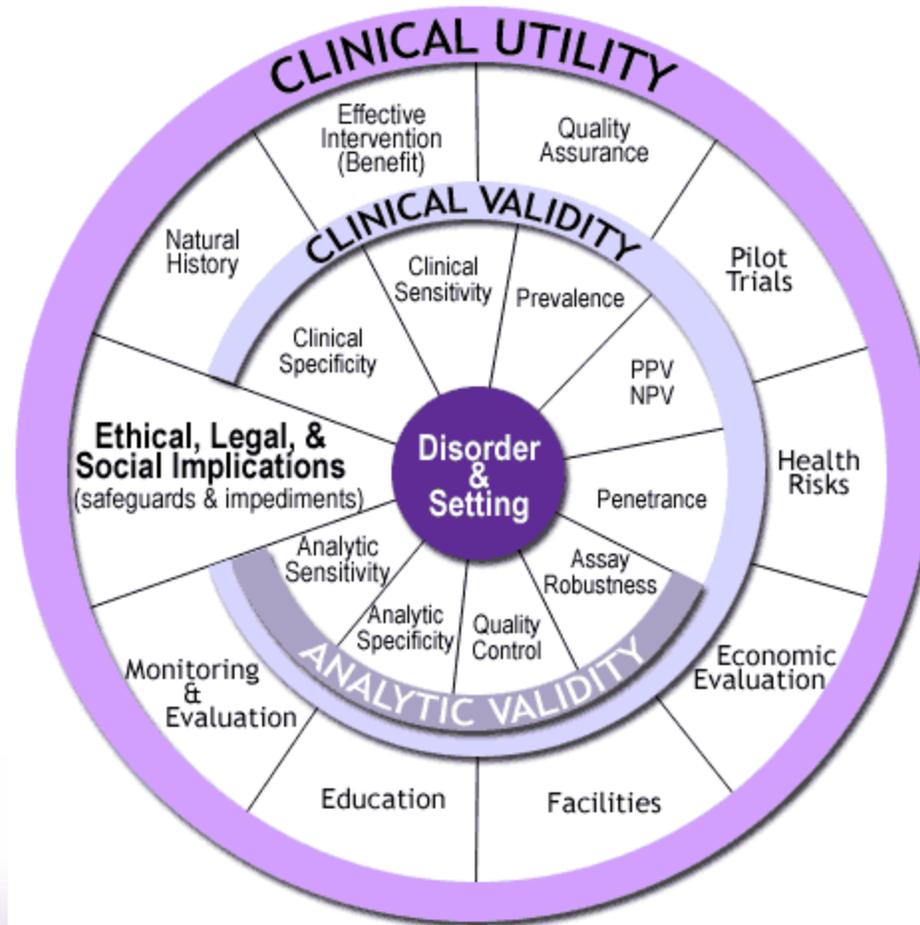
# Collaborators

- AHRQ and its Evidence Practice Centers
- Institute for Preventive Medicine and Medical Screening (IPMMS)

# EGAPP Work Products

- Teutsch et al. - The EGAPP initiative: methods of the EGAPP Working Group - 2009
- Botkin et al. - Outcomes of Interest - 2010
- Veenstra et al. - Improving the efficiency and relevance of evidence-based recommendations - 2012

# ACCE Model



# Clinical Utility -- definition

- The evidence of improved measurable clinical outcomes - is the test worth doing?
- More nuanced
  - Test adds value to patient management decision making compared with current management without the test
  - Benefits of testing outweigh harms

# Familiar Work Process

- Use an analytical framework (causal pathway) with key questions to frame the evidence review
- Identify outcomes of interest
- Develop explicit search strategies
- Perform formal assessment of quality (internal and external validity)
- Evaluate the resulting chain of evidence to reach conclusions about how test effects health care outcomes

# EGAPP Evidence Reports

Title	Author	Conclusions
Genomic profiling and cardiac disease	CDC/IPMMS	Insufficient evidence
Genetic testing in venous thromboembolism	Hopkins EPC	Against routine use
UGT1A1 testing	CDC/IPMMS	Insufficient evidence
Lynch Syndrome	CDC/IPMMS	For use in relatives of patients with disease
Gene expression and breast cancer outcomes	Hopkins EPC	Insufficient evidence
Hereditary nonpolyposis colorectal cancer	Tufts EPC	Limited evidence
P450 and non-psychotic depression	Duke EPC	Insufficient evidence
Genomic tests for ovarian cancer	Duke EPC	Archived

# Challenges Observed (among others)

- Lack of direct evidence of clinical utility
- No formal framework for evaluating indirect evidence of clinical utility
- Limited consensus among stakeholders about the types of evidence needed, outcomes to be assessed and thresholds to be set before recommending a new test

# Lessons Learned

- Quick and early assessment of quality (to R/O topics)
- Search and use existing reviews
- Evaluate clinical validity before proceeding (quite unique)
- Use decision modeling as a component of full evidence review

# A Bible for Test Review

- AHRQ Methods Guide for Medical Test Review
- <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=1088>

# AHRQ Methods Guide

- Developing the topic and structuring the review - a discussion of the use of **analytical frameworks**, **PICOTS** (population, interventions, comparators, outcomes, timing, setting), and **key questions**
- Assessing risk of bias - the **QUADAS** system (patient selection, index tests, reference standard, flow and timing)
- **GRADE** approach to assessing the strength of a body of evidence (risk of bias, consistency, directness, precision)
- Deciding when to use **decision modeling**
- **Chasm**

# Observation # 1: Companion diagnostics are not your father's Oldsmobile

- When drug use becomes tied to test use, the safety and efficacy of drug becomes subservient to the test
- Clinical utility of the two products begin to converge
- Test has potential to have direct impact on drug pipeline

## Observation #2: If evidence based medicine is a goal why not ask for evidence

- We should be setting the bar higher; no shortage of studies (12,500 studies on cancer biomarkers in 2012); it shouldn't be such a struggle

## Observation # 3: The goose that lay the golden egg may be working for Medicare dollars

- We should put costs squarely on the table (should not be a dirty word)
- How research is funded a matter of discussion
- Truthful labeling of products best characterized as investigational should not be a matter of discussion

# Observation # 4: To have value based medicine we need a better understanding of what patients value

- Should include patient values and the ability to accommodate to adversity (resiliency)
- As Harzband and Groopman suggest (2012) There is more to life than death, there is also the “vital dimensions of life that are not easily quantified” .

## Observation # 5:

Patients may not always know what they are getting

- According to Weeks et al (2012) 81 % of patients with advanced colorectal cancer and 69% of patients with advanced lung cancer receiving chemotherapy believe they are receiving curative rather than palliative care
- Those who knew the truth generally liked their doctors less

## Observation # 6:

Physicians may not always know what they are getting

- Cacophony of terms: positive and negative predictive values, ROC curve and C-statistic, likelihood ratio, relative risk, odds ratio, net reclassification improvement, others)
- Problem of misuse: both over and underuse

# Observation # 7: In the pursuit of biomarker success don't forget

Bad News: "If wishes were horses, beggars would ride."

# Observation # 7: In the pursuit of biomarker success don't forget

Good News: "The tortoise won the race"

# Stopping by Woods on a Snowy Evening - Frost

My little horse must think it queer  
To stop without a farmhouse near  
Between the woods and frozen lake  
The darkest evening of the year.

He gives his harness halts a shake  
To ask if there is some mistake  
The only other sounds' the sweep  
Of easy wind and downy flake.

The woods are lovely, dark and deep.  
But I have promises to keep,  
And miles to go before I sleep,  
And miles to go before I sleep.