

# Comparative Effectiveness Research for diagnostic tests and biomarkers

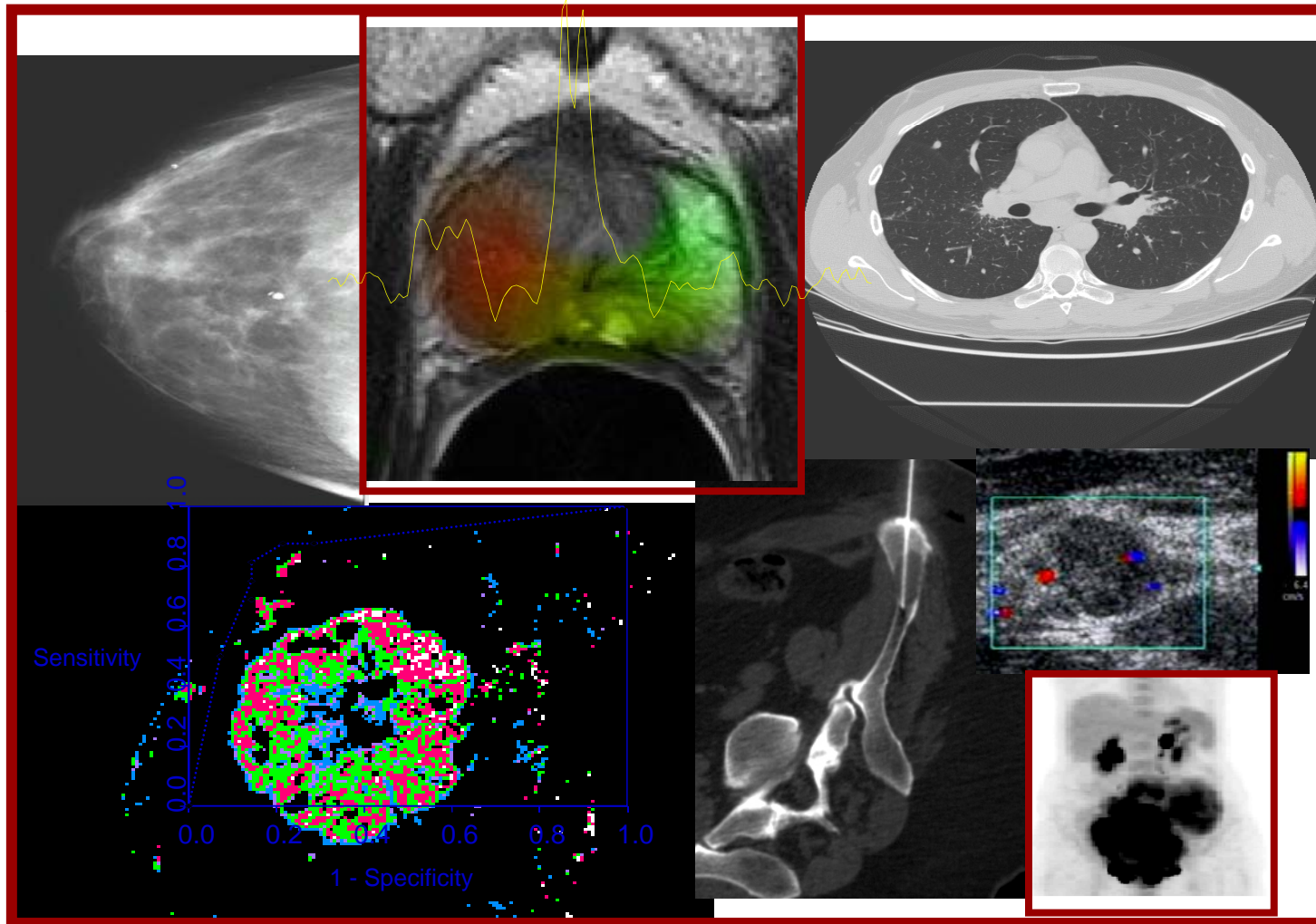
**Constantine Gatsonis**

*Center for Statistical Sciences*

*Program in Public Health*

*Brown University*

# Testing modalities are receiving much attention in CER



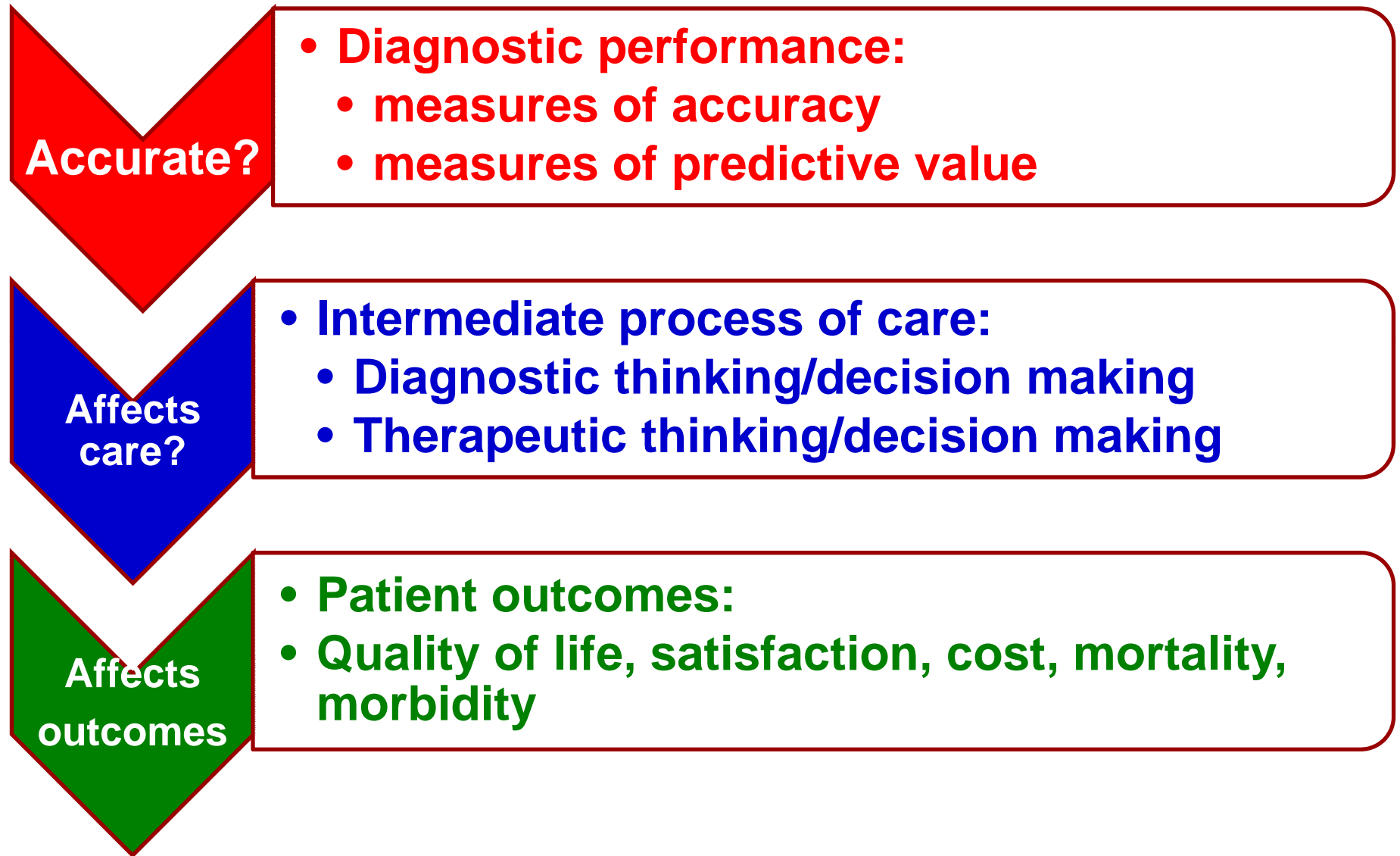
NIH-CER Dec 2010

Constantine Gatsonis, PhD

# First Quartile in IOM report

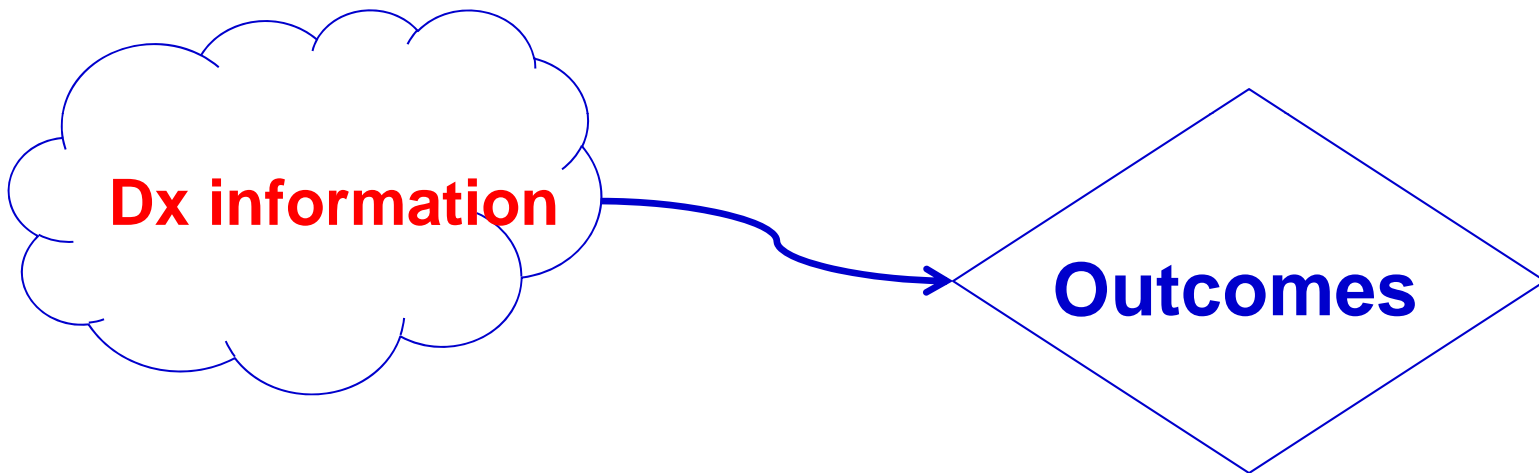
- *Compare the effectiveness of imaging technologies in **diagnosing, staging, and monitoring patients with cancer** including positron emission tomography (PET), magnetic resonance imaging (MRI), and computed tomography (CT).*
- *Compare the effectiveness of genetic and biomarker testing and usual care in **preventing and treating breast, colorectal, prostate, lung, and ovarian cancer**, and possibly other clinical conditions for which promising biomarkers exist.*

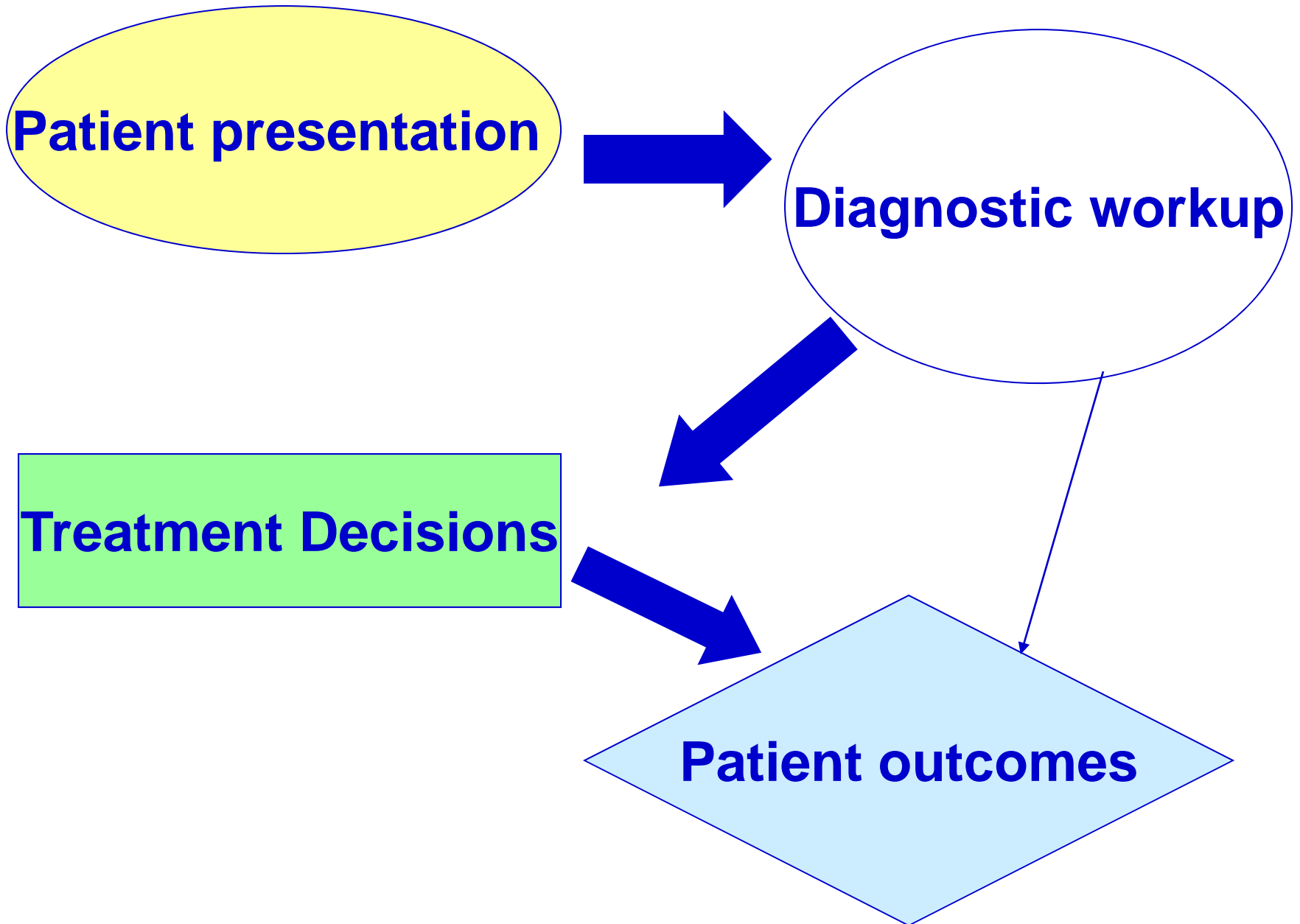
# Endpoints for diagnostic test evaluation



# Fundamental challenge for diagnostic tests

CER calls for illuminating the path:





# Tests and therapeutic interventions

- Fundamentally tests provide information for use in selecting course of care.
- Both long- and short-term effects of tests materialize in context of available health care options, including therapeutic interventions.
- Not possible to define and measure test effects outside the particular health care context in which the test will be used.
- However, oftentimes diagnosis may be ahead of therapy: DCIS is a good example

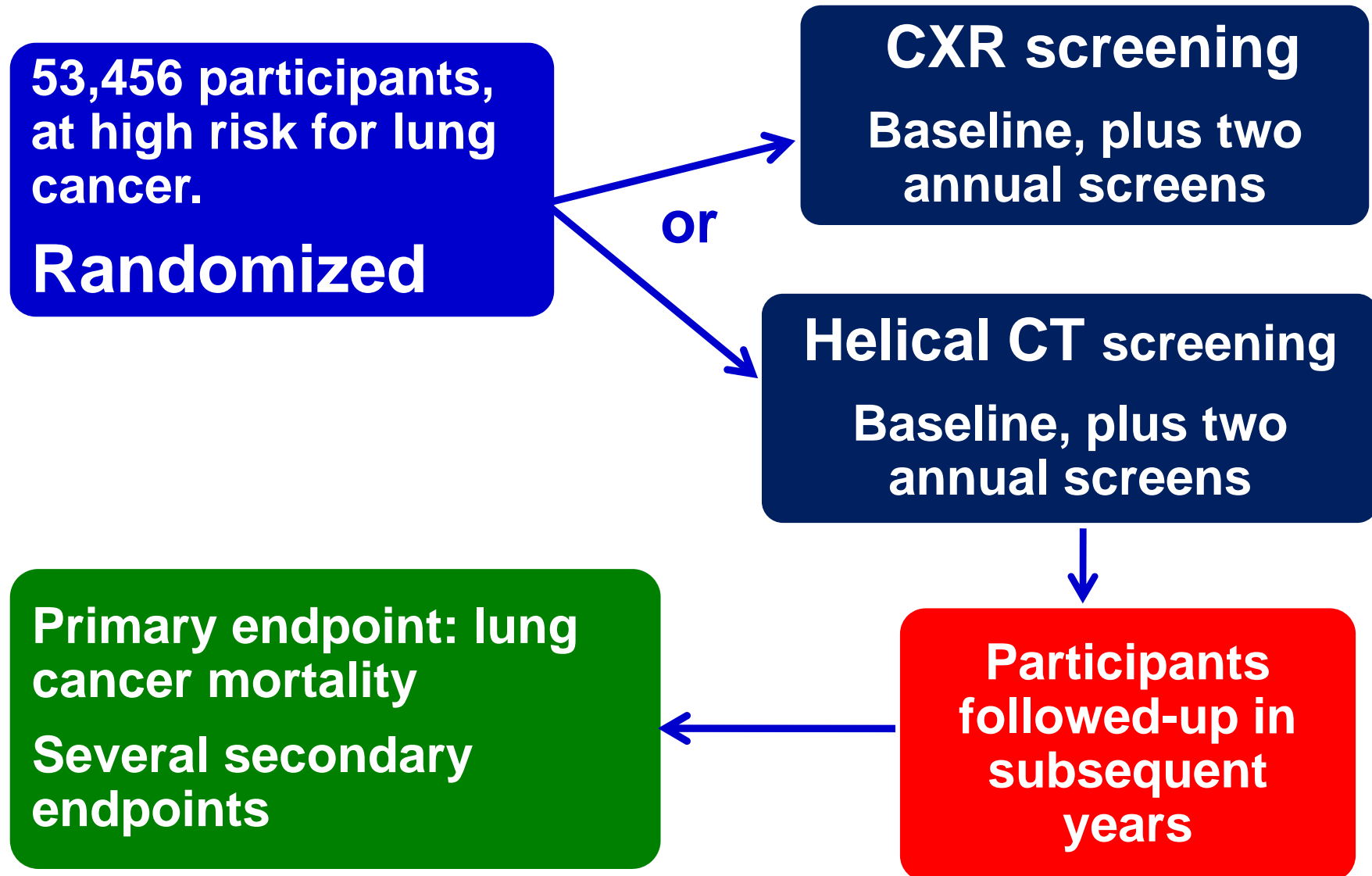
**CER thinking has had a long history in  
the some parts of diagnostic test  
evaluation:**

**The case of cancer screening research**

## **CER in cancer screening**

- **Methodologic challenges in extrapolating from accuracy and intermediate outcomes to patient-level outcomes have been studied.**
- **Fallacies arising from length and lead time bias have been documented and addressed.**
- **Long debates about benefits **and** harms of testing.**
- **Randomized studies of outcomes have become the “gold standard” .**

# National Lung Screening Trial (LSS/ACRIN 6654)



**The traditional randomized studies of how cancer screening affects mortality are**

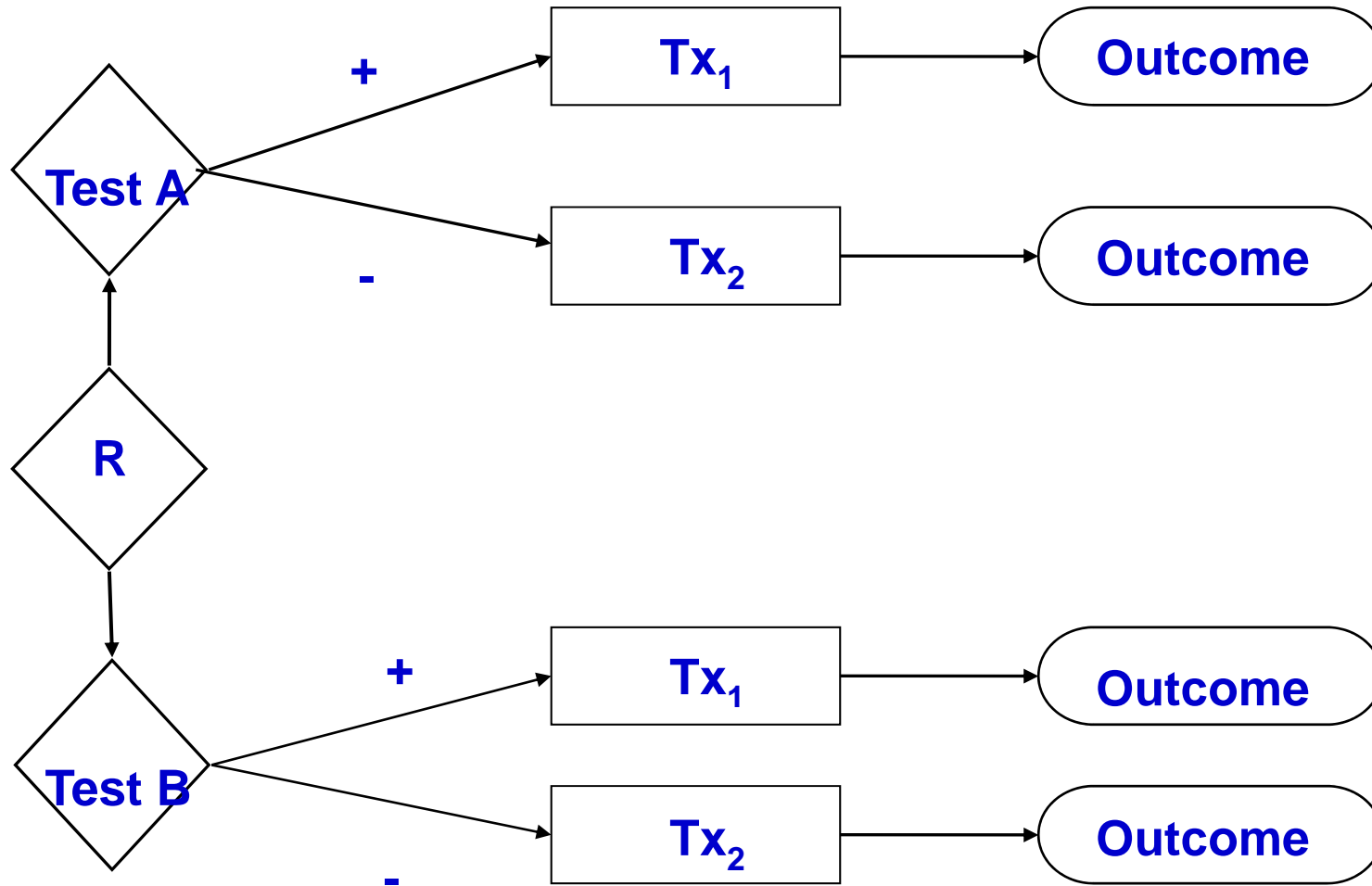
***blunt instruments***

- **Typically used to study broad populations**
- **Typically framed as evaluations of public health interventions, without tight connections to specific diagnostic and intervention algorithms**

## **Simple example of outcomes study**

- **Goal: Compare tests A and B on the basis of a down-stream binary outcome.**
- **(Simplified) therapy setting:**
  - **A positive test result leads to treatment Tx1**
  - **A negative test result leads to treatment Tx2**
  - **Tx2 may be “usual care” (e.g. screening) or active therapy (e.g. determining whether cancer is operable or not).**
- **Prior studies provide estimates of success rates for Tx1 and Tx2, performed on appropriate cases.**

# Simple randomized design, comparing two tests



**Difference in success rates between two arms:**

$$D=(r_1-r_2)p(\text{Sens}_A - \text{Sens}_B)$$

$r_1$  and  $r_2$  = *success rates* for therapeutic interventions Tx<sub>1</sub> and Tx<sub>2</sub>, when performed on cases that have the clinical condition (irrespective of which test detected them)

$p$  = prevalence of the clinical condition

*Sens* = test sensitivity

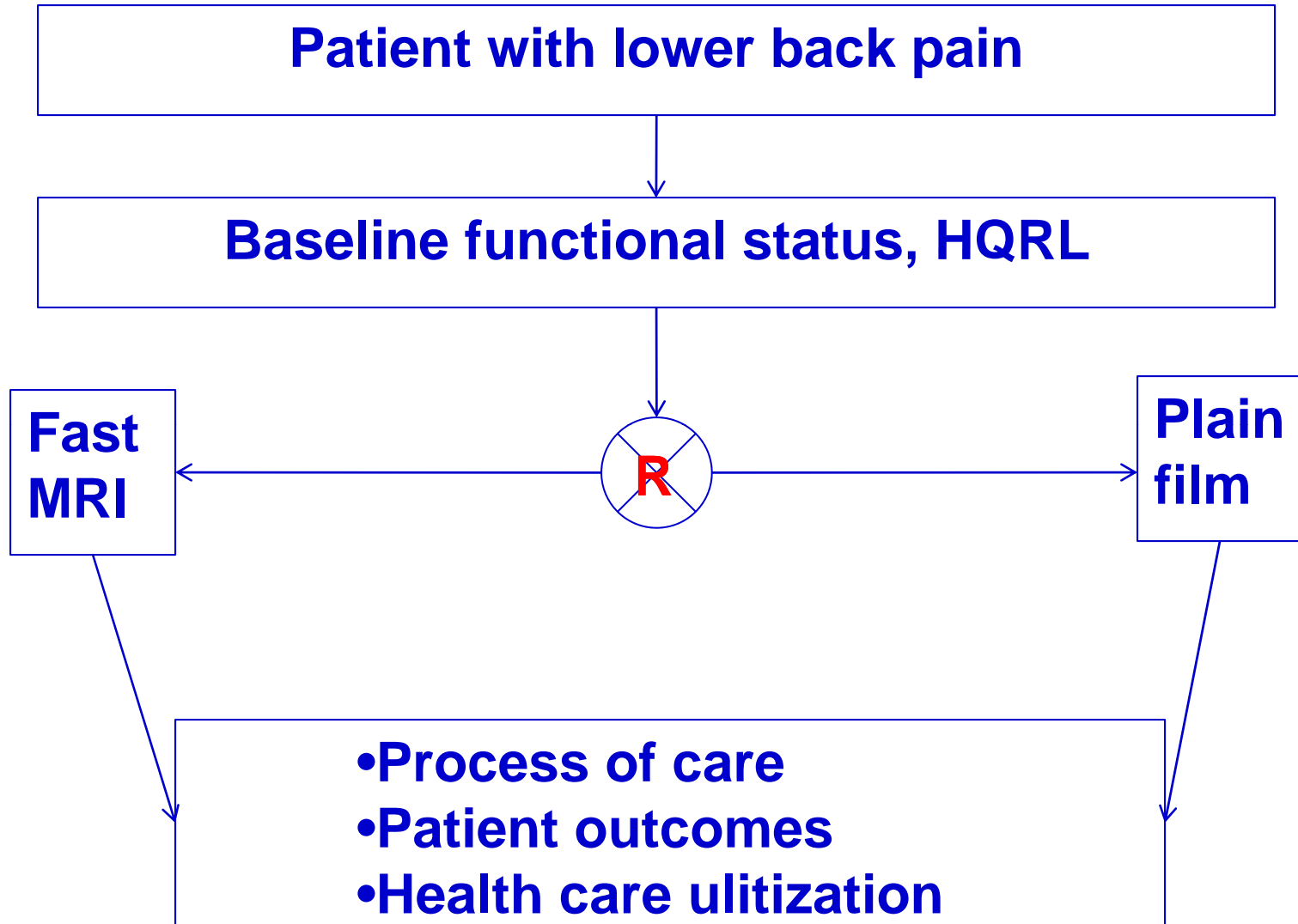
Specificities are assumed equal.

*Typically,  $D$  will be much smaller than  $r_1-r_2$*

*If  $r_1-r_2 = 0$  or  $\text{Sens}_A = \text{Sens}_B$  then  $D=0$ .*

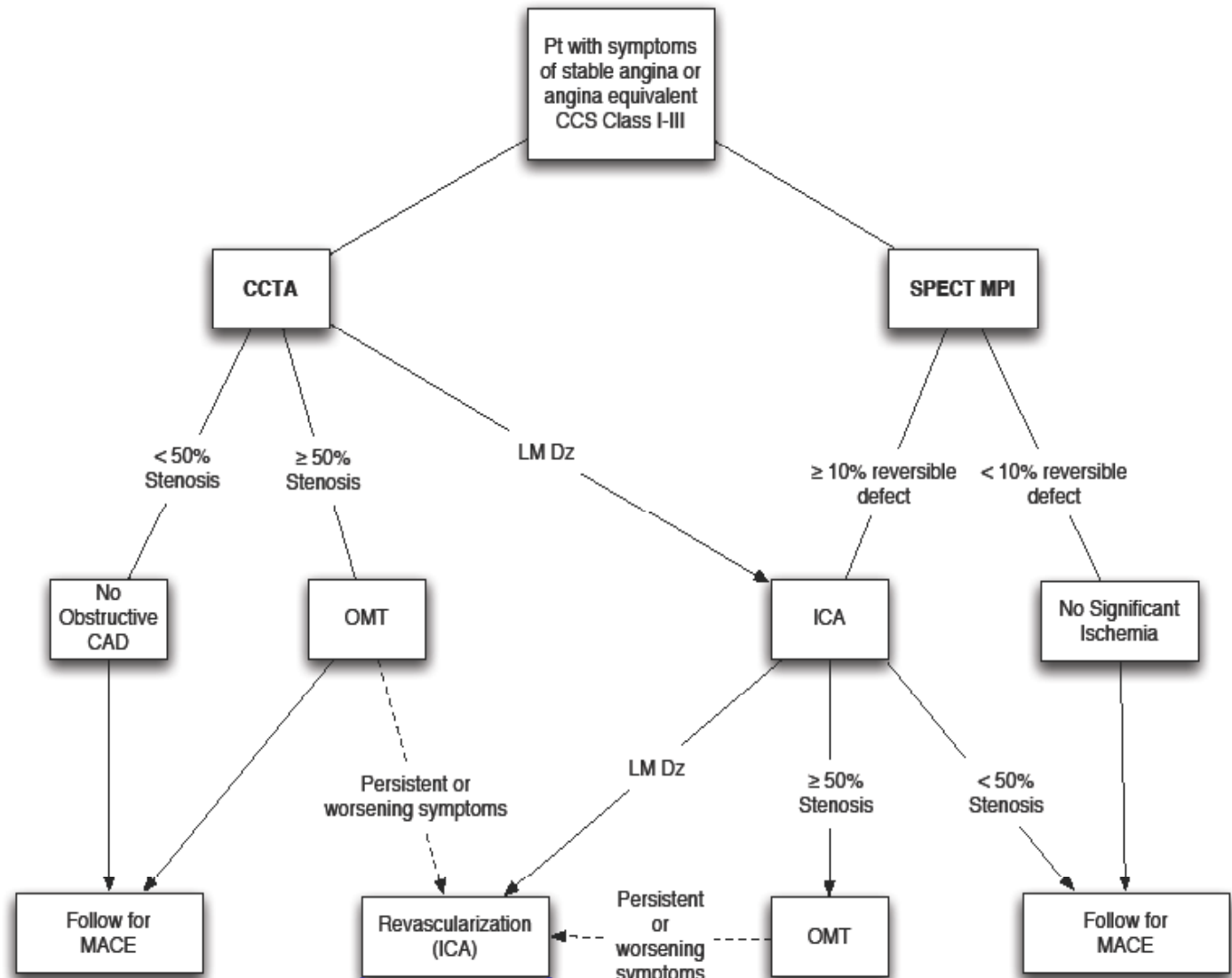
**Still, don't count out RCTs from the evaluation of diagnostic tests !**

## A randomized study of outcomes



Jarvik et al Radiology 1997, JAMA 2003

## A current example



# Other options for CER of diagnostic tests

- **Observational studies**
  - **Registry studies**
  - **Electronic Medical Records**
  - **Claims datasets**
  - **Cohort studies**
- **Research synthesis**
  - **Systematic reviews and meta-analysis**
  - **Modeling**

## Example of potential for observational studies

### Biennial Versus Annual Mammography and the Risk of Late-Stage Breast Cancer

Tumor characteristics	N†	All (n = 7155)
All cancers (DCIs and invasive)		
Invasive cancer	7155	1.31 (1.13 to 1.51)
Late stage‡	6962	1.03 (0.89 to 1.19)
Invasive cancers only		
Late stage‡	5594	0.97 (0.84 to 1.13)
Tumor size > 20 mm	5387	1.07 (0.92 to 1.24)
Grade III or IV	4720	1.01 (0.87 to 1.17)
ER-negative	4164	1.22 (1.01 to 1.48)

OR  
2yr vs 1yr

**White et al, JNCI 2004**

## **Observational studies can answer:**

- **Screening performance in real world settings**
- **Examine factors influencing variability in performance and outcomes**
  - **patient (risk, subtype of disease)**
  - **facility, region, health care system**
- **Assess generalizability of efficacy results to broader cohorts.**
- **Examine the components of screening effectiveness and program effectiveness.**
- **Assess technology diffusion and utilization patterns.**

# National Oncologic PET Registry: A Nationwide Collaborative Program

Sponsored by 

Advisor 

Managed by  

## **Goal:**

***Assess the effect of PET on referring physicians' plans of intended patient management.***

- across a wide spectrum of cancer indications for PET, not covered currently by the Medicare program,
- in relation to cancer-type, indication, performance status, physician's role in management, and scan type

## PET Changed Intended Management in 36.5% of Cases

		Clinical Indication for PET Study (Percent)				
Pre-Pet Plan	Post-PET Plan	Dx n=5,616	Staging n=6,464	Restaging n=5,607	Recurrence n=5,388	All n=22,975
Treat	Same	16.0	46.5	15.8	20.4	25.5
Non-Treat	Same	52.9	14.0	48.0	40.7	37.9

Non-Treat	Treat	23.2	31.6	28.6	29.2	28.3
Treat	Non-Treat	7.9	7.9	7.5	9.7	8.2
Patients with change post-PET (%)		31.1	39.5	36.1	39.0	<b>36.5</b>

## **Data needs for observational studies**

**Registry and claims data (at present!) are deficient in information about**

- **Context of the test (indication, filtering)**
- **Test interpretation and findings**
- **Clinical characteristics of patient**
- **Subsequent diagnostic and therapeutic decisions .**

## **Brighter future**

- **More comprehensive and reliable data systems from health care payers.**
- **Widespread use of EMR.**
- **Advanced informatics for sharing data while preserving confidentiality.**
- **Great strides in statistical methodology for observational data analysis.**

# **CER and new diagnostic technology**

- **Emphasis on patient outcomes can be**
  - **Very helpful OR**
  - **Simplistic and counterproductive**
- **Helpful:**
  - **Emphasize consideration of potential relevance from early development (e.g. “challenge ROC”) to maturity**
- **Simplistic and counterproductive:**
  - **Ask for ultimate proof on Day 1**

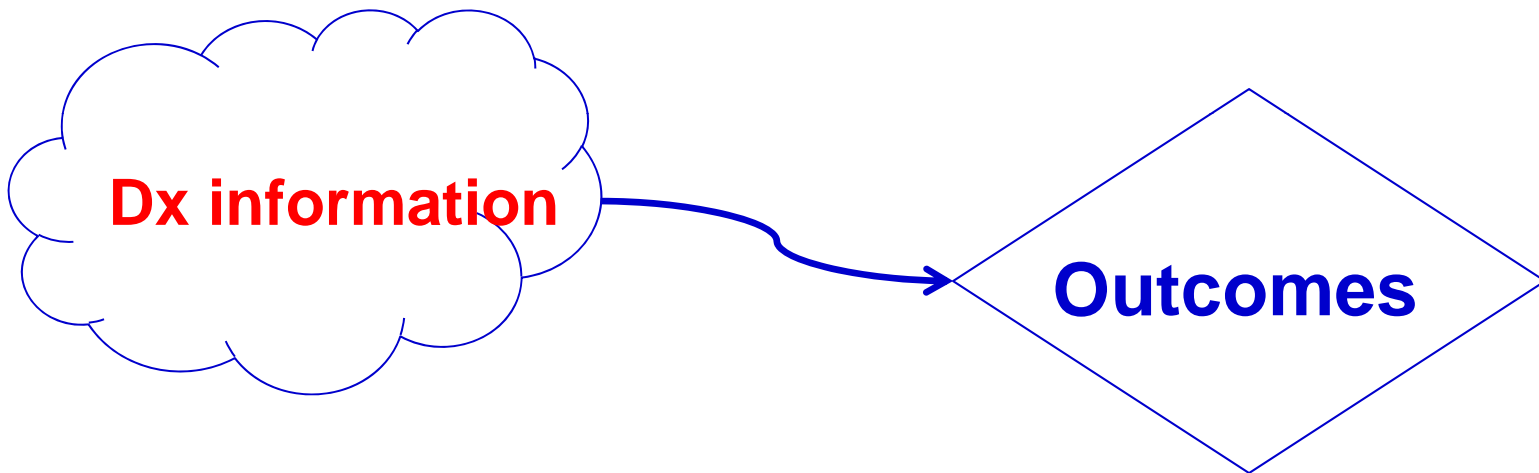
# **New and more accurate technologies**

- **May also intensify the potential for**
  - **Overdiagnosis (e.g. breast MRI)**
  - **Incidental findings (e.g. Helical CT for lung cancer, CTC for colon cancer)**

**To conclude**

**For diagnostic tests**

**CER calls for illuminating the path:**



## **For diagnostic tests**

**achievable and highly relevant CER studies can be developed if**

- The patient focus is maintained**
- the full range of CER methodologies is embraced,**
- the proper role of randomized studies is understood,**
- the potential of sophisticated modeling is fully developed.**

# **CER requires new thinking about evidence**

- i. What evidence will be considered definitive?**
- ii. What statistical precision will be required?**
- iii. What level of bias can be tolerated?**
- iv. Will the answers depend on the particular decision context or will they apply across the board?**

# Acknowledgment

The ideas presented here benefited from the discussions at *RSNA/NIBIB Workshop on CER for Diagnostic Imaging* held in Chicago, 22-23 April 2010.

