National Cancer Institute

# Building the Evidence Base for Decision Making in Cancer Genomic Medicine using Comparative Effectiveness Research (CER)

Andrew N. Freedman
Clinical and Translational Epidemiology Branch
National Cancer Institute

Evidence for Clinical Utility of Molecular Diagnostics in Oncology: A Workshop
May 24, 2012

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
National Institutes of Health

# Outline

- Background
- Identification of the key questions for CER applications in cancer genomics
- General methodological approaches to addressing these key questions using examples
- Future research needs

# Background

- Analytic validity and clinical validity is now available for an increasing number of genomic applications
- Clinical Utility is largely unknown for most genomic applications
- Uncertain clinical utility potentially wastes health-care resources and decreases quality through inconsistent or unnecessary use of those tests
- It is important to ensure that clinically valid tests also have high clinical utility before they become widely used

# Genomic Predictive Markers of Cancer Treatment Efficacy and Safety

| Test/Markers | Drugs | Cancer Outcomes |
|---|---|---|
| | ***In Clinical Use*** | |
| HER2/neu | Trastuzumab | Breast Cancer-recurrence/survival |
| Oncotype Dx | Treatment regimen | Breast Cancer-recurrence/survival |
| EGFR Mutation | Erlotinib | Lung Cancer-Recurrence/survival |
| K-ras | Cetuximab, Panitumumab | Colorectal Cancer - recurrence/survival |
| EML4-ALK mutation | Crizotinib | Lung Cancer-Recurrence/survival |
| BRAF V600E | Vemurafenib | Melanoma Recurrance/survival |
| BCR-ABL | Imatinib, Dasatinib, Nilotinib | CML-Response |
| C-Kit | Imatinib | GIST-Response |
| *TPMT* | *6-MP, 6-TG* | *ALL, AML-Toxicity* |

Febbo et al. J Natl Compr Canc Netw 2011;9:S-1-32

# Genomic Predictive Markers of Cancer Treatment Efficacy

| Test/Markers | Drugs | Cancer Outcomes |
|---|---|---|
| ***Emerging Evidence*** | | |
| MSI and/or MMR | 5-FU | Colorectal Cancer-Recurrence/survival |
| Mammaprint | Treatment regimen | Breast Cancer-recurrence/survival |
| Oncotype Dx Colon ColoPrint | Treatment Regimen | Colorectal Cancer Recurrence/survival |
| K-ras mutation | Anti-EGFR therapy | Lung Cancer-recurrence/survival |
| ERCC1 | Cisplatin-based Therapy | Lung Cancer-Recurrence/survival |

Febbo et al. J Natl Compr Canc Netw 2011;9:S-1-32

# Comparative Effectiveness Research (CER)

- CER is intended to create evidence for decision making, and to find out “what works” in health care.

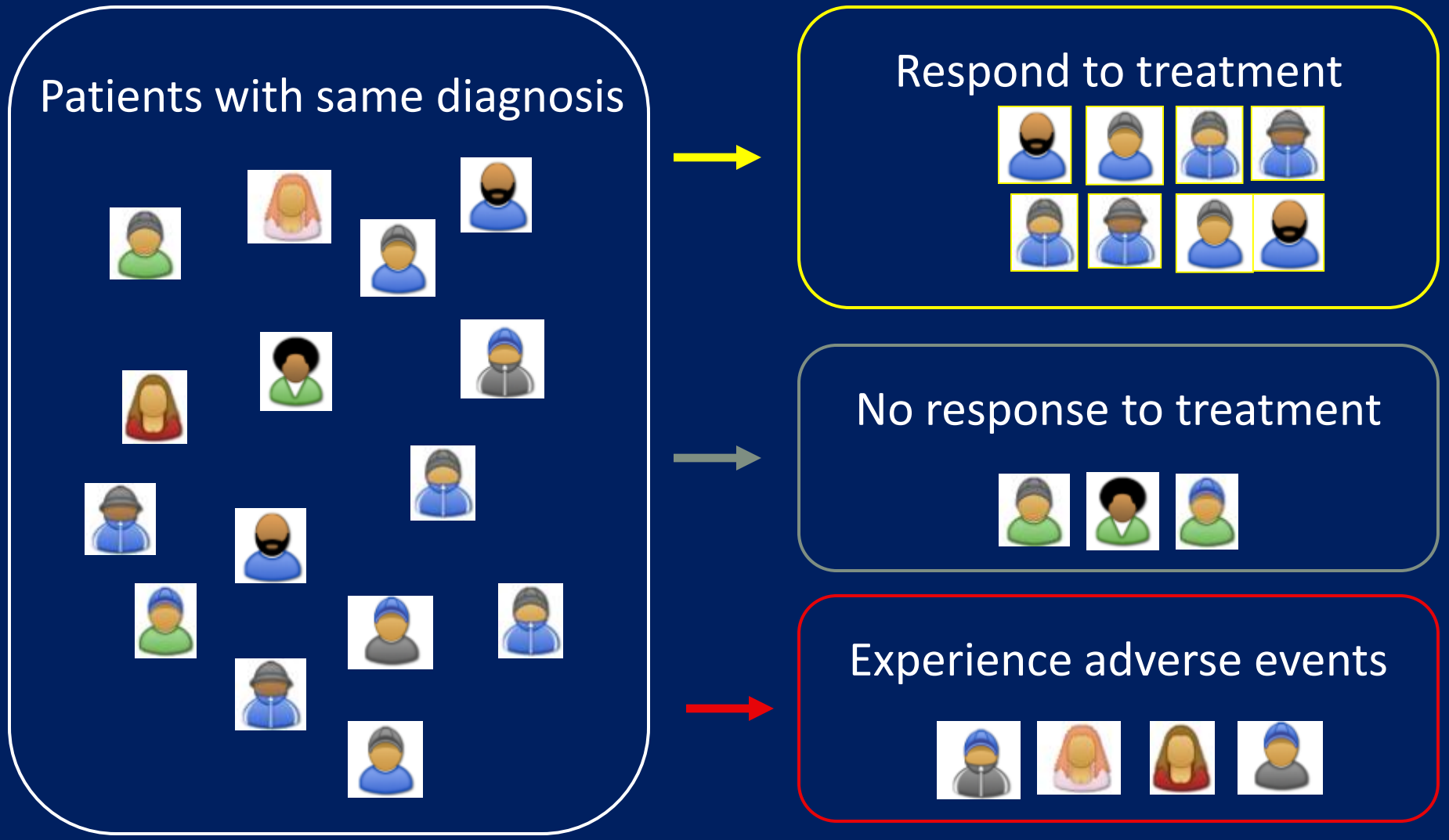
- IOM definition: “CER is the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers, and policy makers to make informed decisions that will improve health care at both the individual and population levels.”

- “Patient-Centered Outcomes Research”

# "The Trouble With Averages"

- CER typically focuses on average treatment effects, but…
  - Interventions that yield a significant treatment effect across a study population may be ineffective for some patients and harmful for others
  - Conversely, interventions that may be dismissed as ineffective actually may work for certain subgroups of the population

# Cancer Genomic Medicine



# Decision-Makers Key Questions for Cancer Genomic Medicine

1. Does the genomic application provide correct information? (analytic validity)

2. Is there a significant association between the results of the genomic application and clinical phenotype? (clinical validity)

3. Does the genomic application provide clinically significant information? (clinical utility)

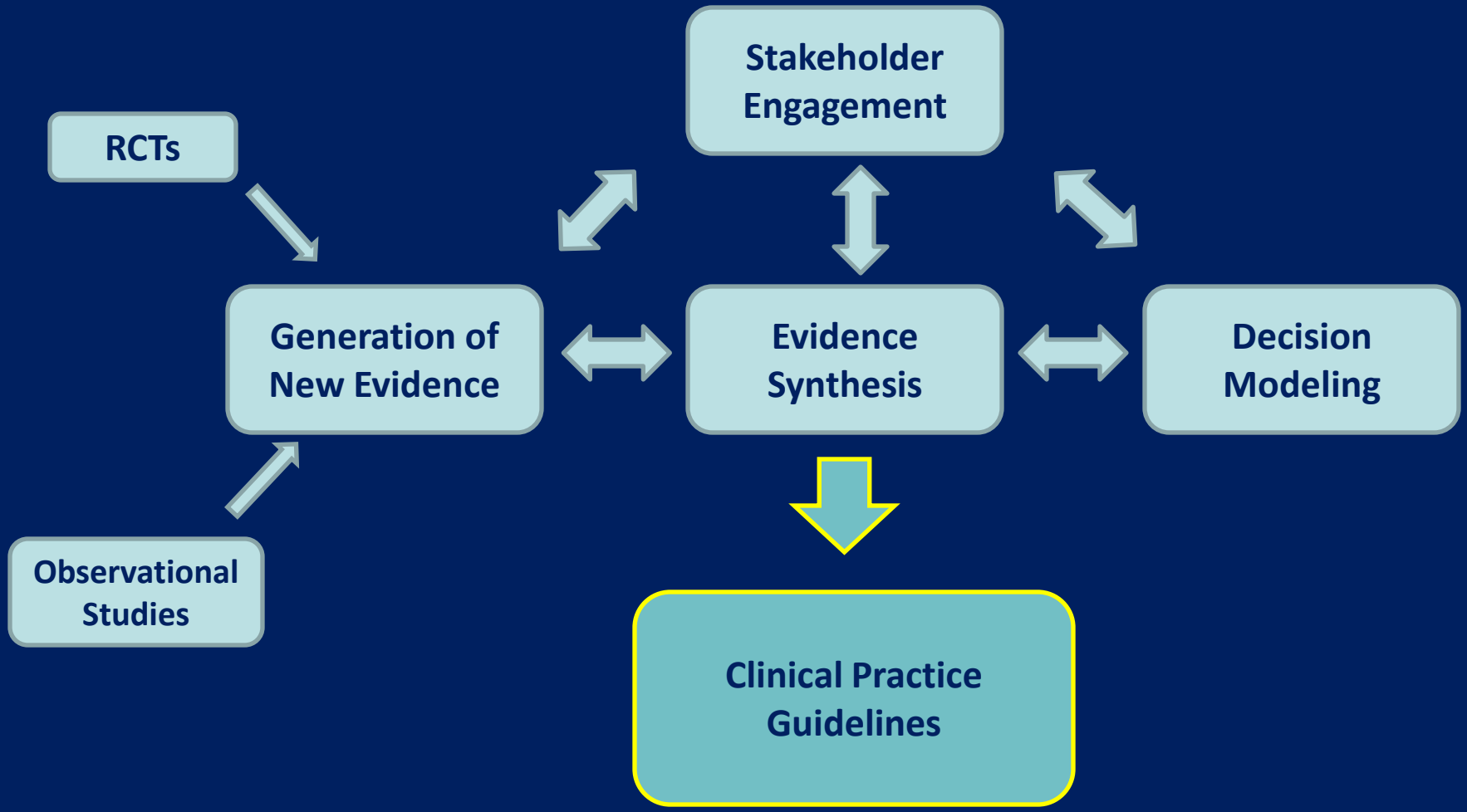
4. Does the genomic application lead to improved patient outcomes as compared with the alternative? (comparison or added clinical value)

# Comparative Effectiveness Research (CER) vs. Traditional Studies Of Genomic Tests For Cancer

| Feature of research | Comparative effectiveness research | Traditional studies |
|---|---|---|
| **Priority of study among alternatives** | Determined by multiple stakeholders, using criteria such as disease burden or cost, lack of information, variability in care | Opportunity as dictated by expert assessment of emerging technology |
| **Study design** | Retrospective or prospective analysis | Retrospective analysis of existing tumor specimens; occasional prospective analysis of observational data |
| **Comparisons** | Direct comparisons of new therapy with usual care | Direct comparisons of competing therapies, often not considering usual care |
| **Topics** | Prevention, treatment, monitoring, and other broad topics | In most cases, prediction of narrow effects such as serious drug interactions, response to treatment, tumor recurrence |
| **Perspectives** | Multiple, including clinician, patient, purchaser, and policy maker | Clinician and patient |
| **Study populations and settings** | Representative of clinical practice | Highly selected |
| **Data elements** | Patient characteristics, quality of life, safety of treatment, resource use and costs, patients' preferences | Patient characteristics, clinical endpoints |
| **Funding** | "Coverage with evidence development" programs, public-private partnerships | Private investors, research grants from federal sources such as the National Institutes of Health |

Ramsey SD et al. How comparative effectiveness research can help advance 'personalized medicine' in cancer treatment. *Health Aff (Millwood).* Dec 2011; 30(12):2259-2268.

# General Methods for Comparative Effectiveness Research

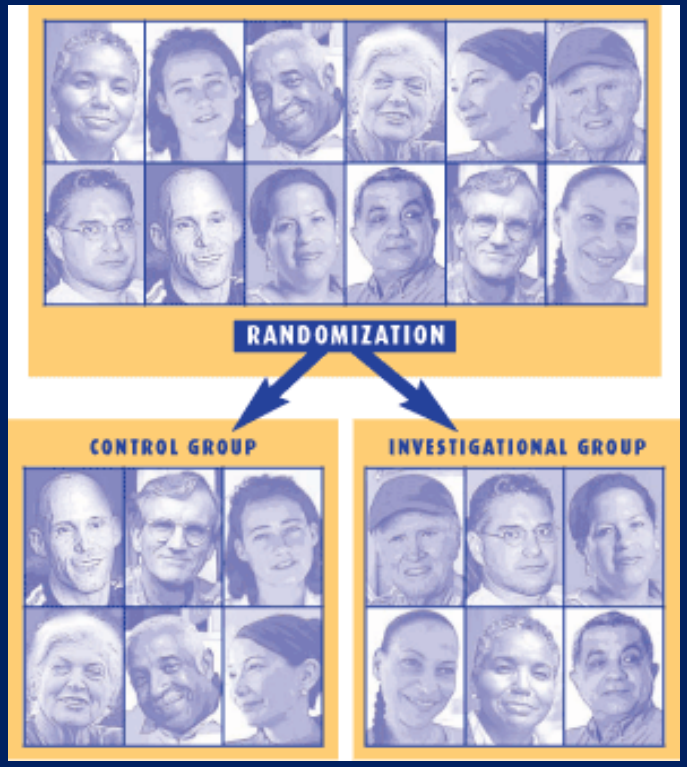


# Evidence Synthesis

- Horizon scanning
  - Searching published and gray literature
  - existing curated databases for emerging genomic applications
- State-of-the-science reviews
  - The NIH State of the Science Consensus Development Conferences
- Identification and prioritization of research gaps
  - Institute of Medicine Consensus Study (e.g. Comparative Effectiveness Research Prioritization)
- Systematic Reviews
  - Evaluation of Genomic Applications in Practice and Prevention Working Group
  - AHRQ's Evidence-based Practice Centers
  - ASCO, NCCN
  - The Cochrane Collaboration , U.S. Preventive Services Task Force
- Health technology assessment
  - Blue Cross and Blue Shield Association Technology Evaluation Center (TEC)

# Randomized Clinical Trial (RCT)

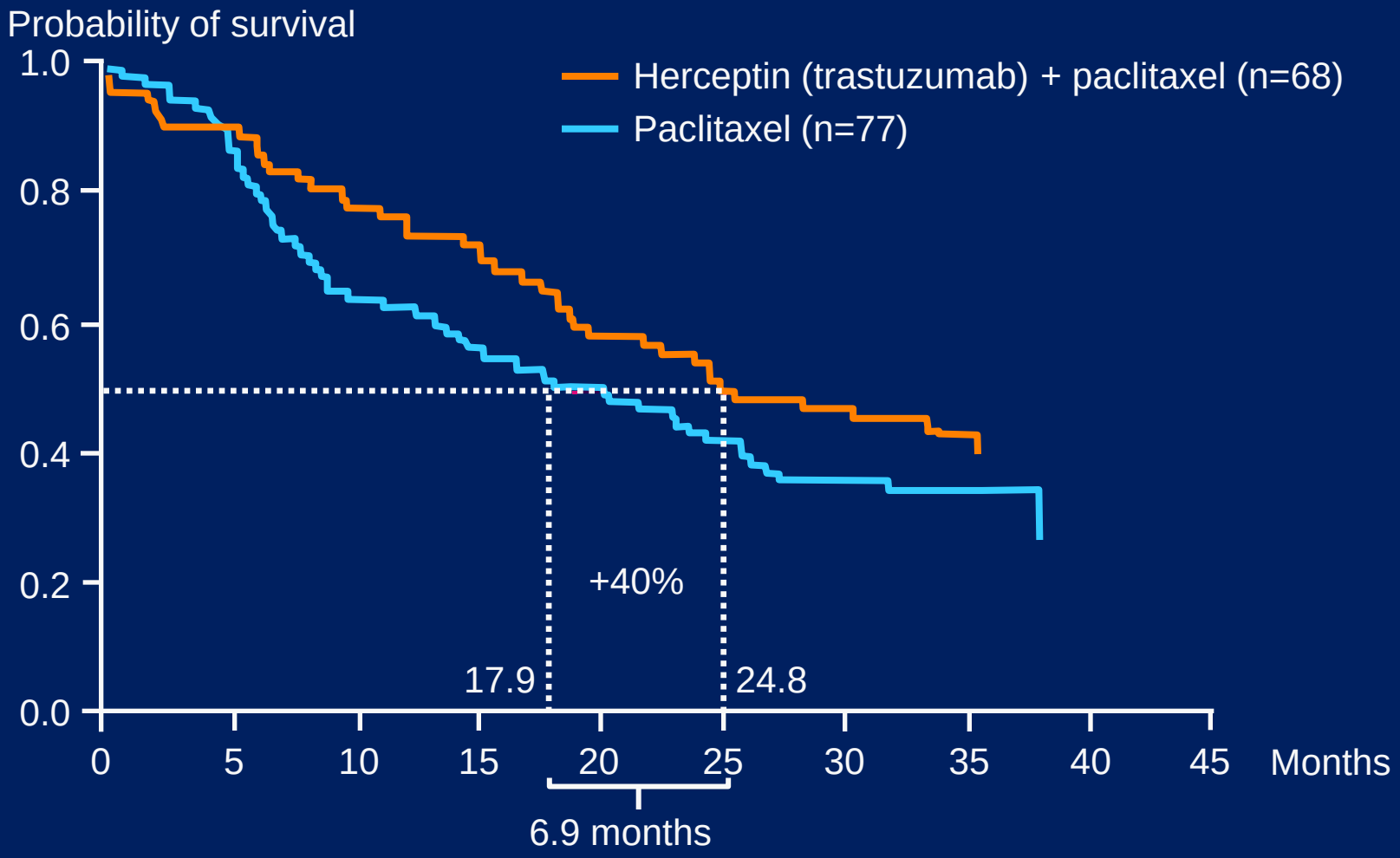
- A planned experiment designed to assess the efficacy of a treatment/marker by comparing the outcomes in a group of patients randomized to the treatment/marker with those observed in a comparable group of patients randomized to a control treatment/marker, where patients in both groups are enrolled, treated, and followed over the same time period.
  - Explanatory RCTs
  - Adaptive Clinical Trials
  - Pragmatic clinical trials
  - Cluster randomized trials




In a randomized clinical trial, you will be assigned by chance to either a control group or an investigational group.

*Meinert, CL. Clinical Trials: Design, Conduct, Analysis. 1986

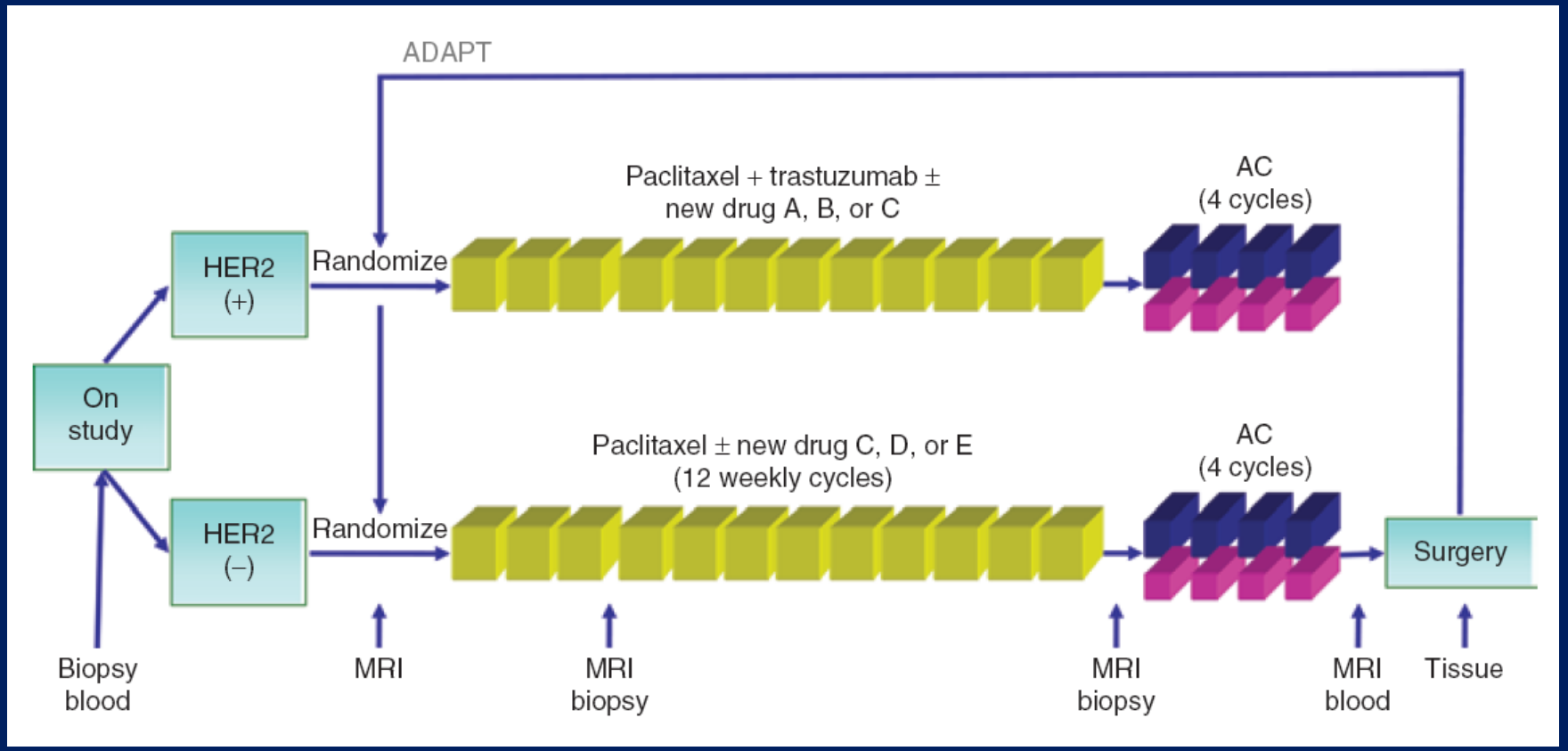
# H0648g: overall survival (IHC 3+ and taxane subgroup)




IHC 2+ tumours were not confirmed for HER2-positivity by *in situ* hybridisation
Subgroup analysis, p value not reported;
OS, overall survival

Smith 2001

Adaptive Clinical Trial- “Learn as you go”: One of more decision points are built into trial design for analysis of outcomes and associated patient or disease characteristics to Identify subgroups who are responding favorably.




I-SPY 2 Trial Design

Barker AD, et al. *Clin Pharmaco Ther*. 2009.

# Pragmatic Clinical Trial (PCT)

- Help decision-makers choose between options for care in routine clinical practice
- Have also been called “practical clinical trials,” “effectiveness trials,” or “large simple trials”
- Provides high quality scientific evidence to support clinical and health policy decisions on a broad range of health outcomes
  - Include morbidity/mortality endpoints, QOL, symptom severity, satisfaction, costs, etc.

Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *Journal of chronic diseases.* Aug 1967;20(8):637-648.

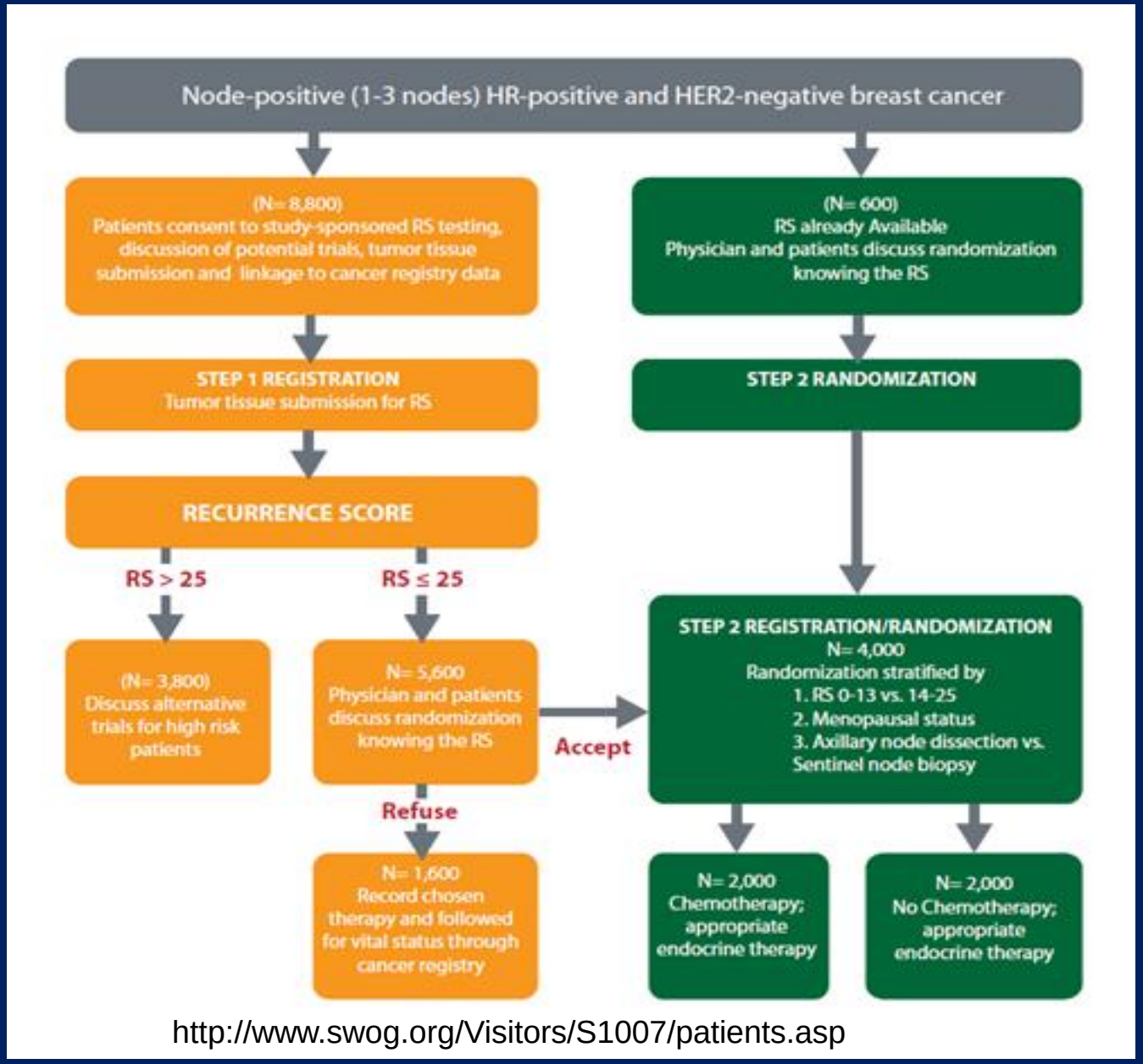
Tunis et al. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *JAMA : the journal of the American Medical Association.* Sep 24 2003;290(12):1624-1632.

# Explanatory vs. Pragmatic Trials

| | Explanatory Trial | Pragmatic Trial |
|---|---|---|
| **Motivation** | Regulatory approval for efficacy – can the intervention work? | Formulary approval for effectiveness – does the clinically relevant intervention work in the 'real world' |
| **Setting** | 'Ideal' (i.e., experimental) setting: usually academic specialists | Normal practice; usually community-based with multiple physician specialties |
| **Participants** | Highly restrictive eligibility criteria with high compliance | Broad eligibility criteria that reflect diversity of patients seen in clinical practice; low compliance |
| **Intervention** | Strictly enforced fixed regimen with forced titration | Applied flexibly as it would be in normal practice |
| **Comparator** | Usually placebo or arbitrarily chosen comparator | Usual care or least expensive/most efficacious |
| **Outcomes** | Condition specific, often short tem surrogates or process measures | Comprehensive and relevant; more often long-term to reflect disease natural history and broad range of functional outcomes |

Zwarenstein M, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ.* 2008;337:a2390.
Bombardier C, Maetzel A. Pharmacoeconomic evaluation of new treatments: efficacy versus effectiveness studies? Annals of the rheumatic diseases. Nov 1999;58 Suppl 1:I82-85.

# RxPONDER (SWOG S1007)



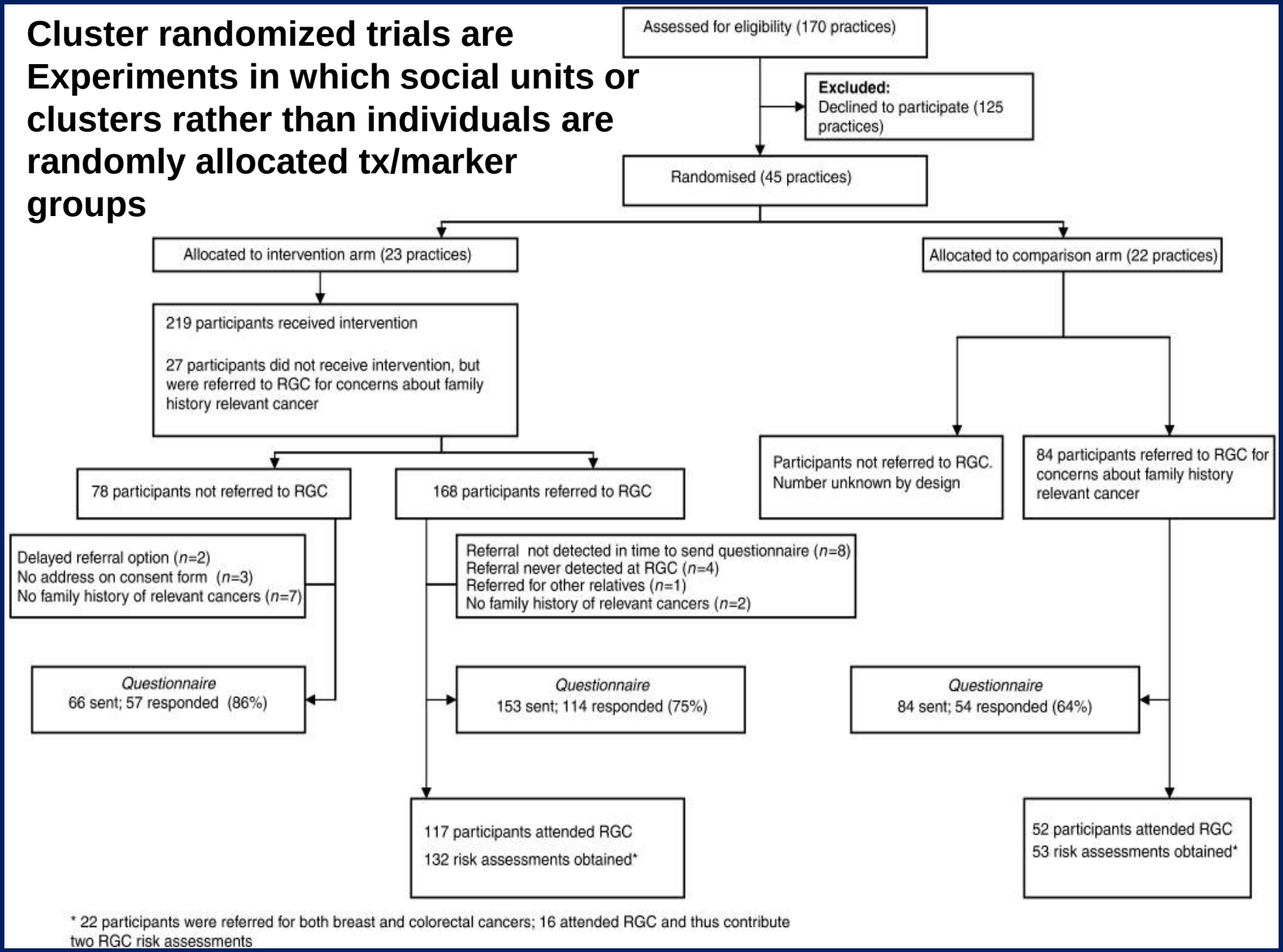

http://www.swog.org/Visitors/S1007/patients.asp

| | Explanatory Trial | Pragmatic Trial | RxPONDER |
|---|---|---|---|
| **Motivation** | Regulatory approval for efficacy – can the intervention work? | Formulary approval for effectiveness – does the clinically relevant intervention work in the 'real world' | **Formulary approval for effectiveness – does the clinically relevant intervention work in the 'real world'** |
| **Setting** | 'Ideal' setting: academic specialists | "Normal" practice setting: community-based with multiple physician specialties | **CCOPs, which are reasonably representative of general clinical practice** |
| **Participants** | Highly restrictive eligibility criteria with high compliance | Broad eligibility criteria that reflect diversity of patients seen in clinical practice | **Restrictive: breast cancer [Node-positive (1-3 nodes) HR-positive and HER2-negative] Broad: recruit minorities , all ages, very large study** |
| **Intervention** | Strictly enforced fixed regimen | Applied flexibly as it would be in normal practice | **Simple intervention: chemo decision based on RS** |
| **Monitoring** | High Intensity | Low Intensity | **Moderate to high intensity** |
| **Comparator** | Usually placebo or chosen comparator | usual care or least expensive/most efficacious | **usual care or least expensive/most efficacious** |
| **Outcomes** | Condition specific, often short tem surrogates | Comprehensive and relevant; more often long-term to reflect disease natural history and broad range of functional outcomes | **cut point for RS, disease free survival, QOL, decision making, cost-effectiveness** |

Zwarenstein M, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ.* 2008;337:a2390.
Bombardier C, Maetzel A. Pharmacoeconomic evaluation of new treatments: efficacy versus effectiveness studies? Annals of the rheumatic diseases. Nov 1999;58 Suppl 1:I82-85.

http://www.swog.org/Visitors/S1007/patients.asp

**Cluster randomized trials are Experiments in which social units or clusters rather than individuals are randomly allocated tx/marker groups**




Emery J et al. The GRAIDS Trial: a cluster randomised controlled trial of computer decision support for the management of familial cancer risk in primary care. *British journal of cancer.* Aug 20 2007;97(4):486-493

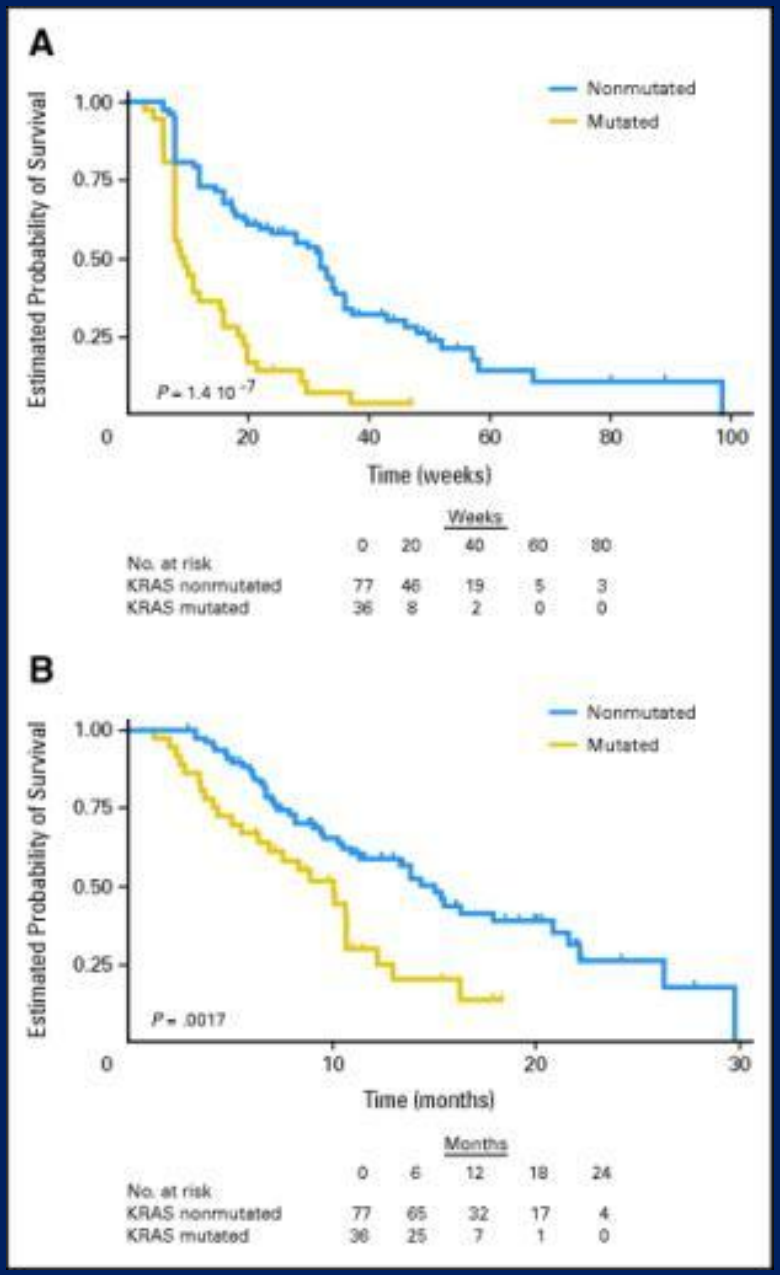
# Observational Studies

- Retrospective and prospective cohort design
- Registries
- Case-control design
  - Administrative databases and EMRs
- Retrospective analysis of biospecimens from RCTs

# 113 CRC Patients According to KRAS Mutation

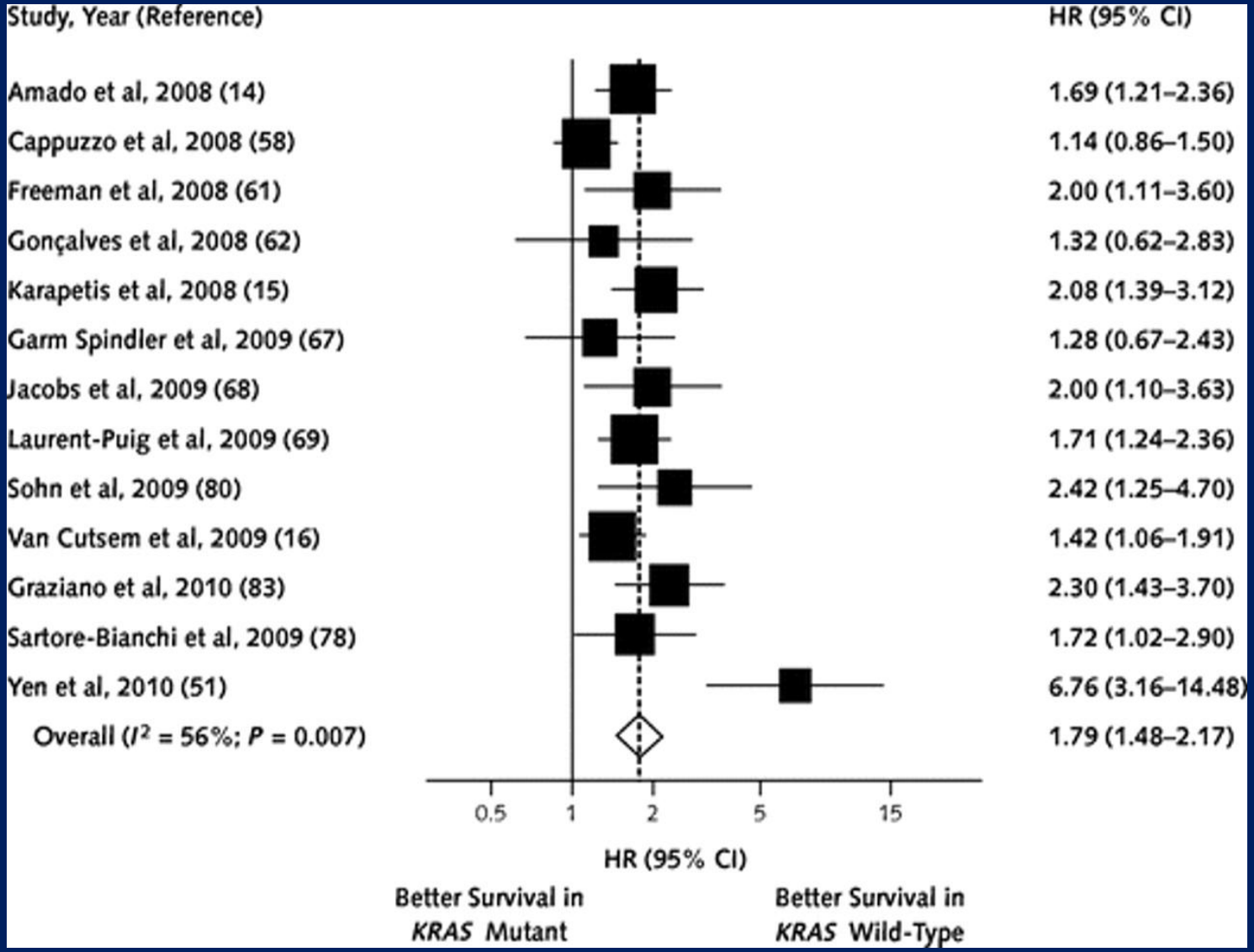
**Progression-free survival**

**Overall survival**



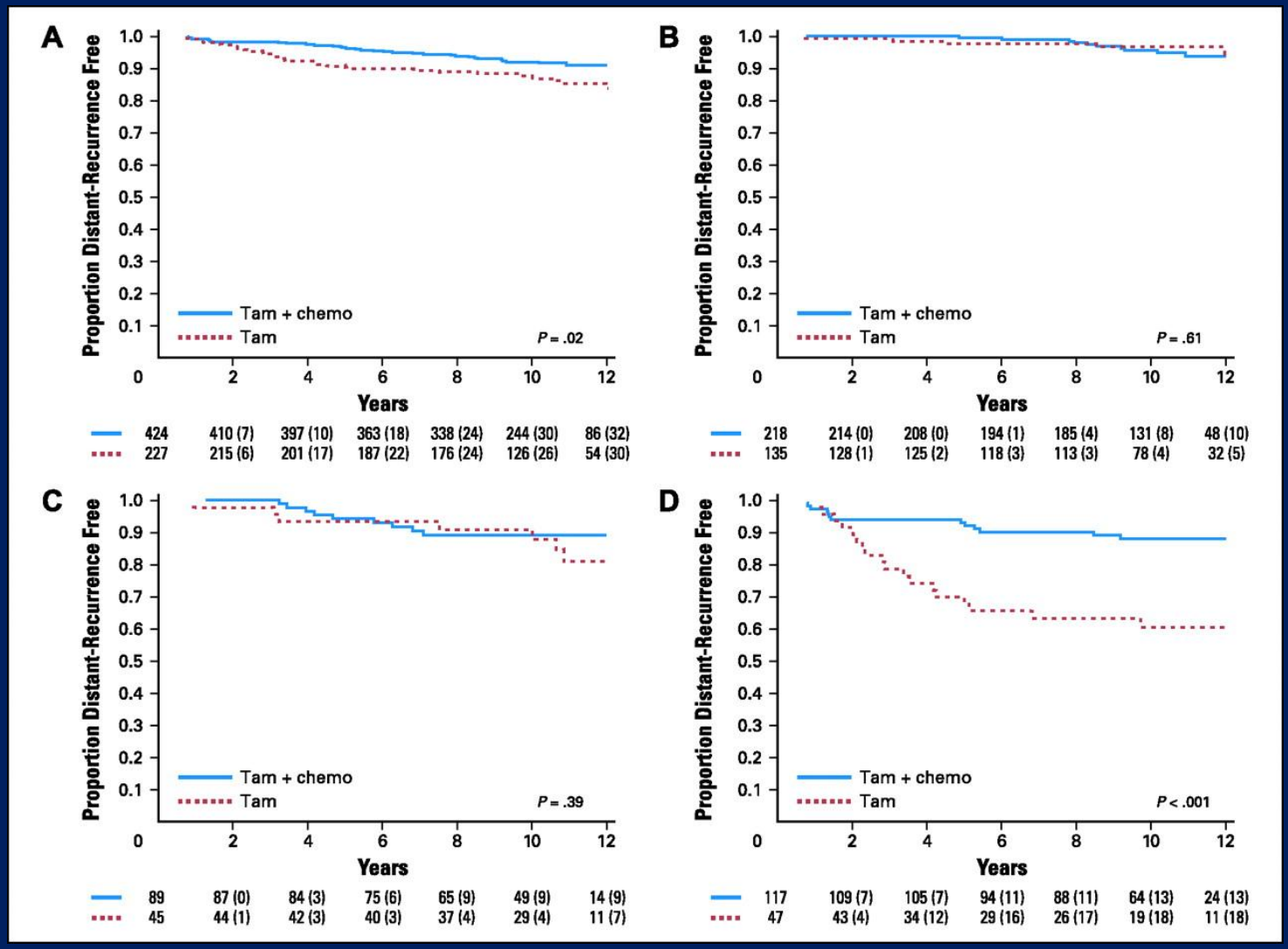

Lievre A, et al. *J Clin Oncol*. 2008

# Forest plot of HRs for overall survival comparing KRAS-mutated and wild-type




Dahabreh I J et al. Ann Intern Med 2011;154:37-49

# Kaplan-Meier plots for distant recurrence comparing treatment with tamoxifen (Tam) alone vs. treatment with tamoxifen plus chemotherapy (Tam + chemo)




Paik S, et al. *J Clin Oncol.* 2006

# Tumor Gene Expression and Risk of BC Death

**Table 5**

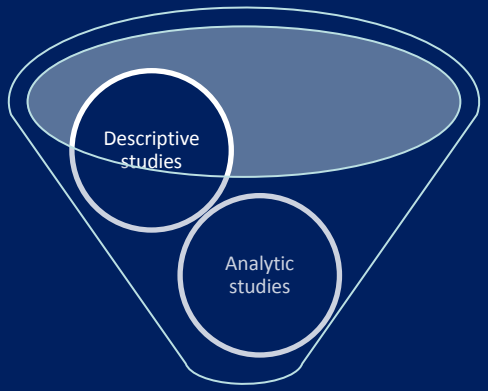
**Ten-year risk of death in relation to Recurrence Score and tumor size and grade among ER-positive patients, stratified by treatment with tamoxifen**

| Risk classifier | Cases | Controls | 10-Year risk | |
|---|---|---|---|---|
| | | | % | 95% CI |
| Tamoxifen treated | | | | |
| Recurrence Score (55 cases and 150 controls) | | | | |
| Low (<18) | 29% | 63% | 2.8 | (1.7–3.9) |
| Intermediate (18–30) | 40% | 23% | 10.7 | (6.3–14.9) |
| High (≥31) | 31% | 13% | 15.5 | (7.6–22.8) |

Habel LA, et al. *Breast Cancer Res*. 2006.

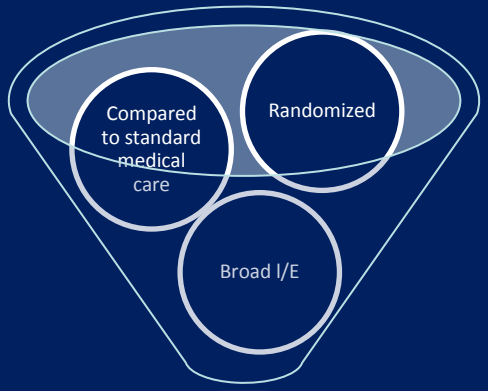
# Causality vs. External Validity

Strength of Evidence: Weaker --> Strong



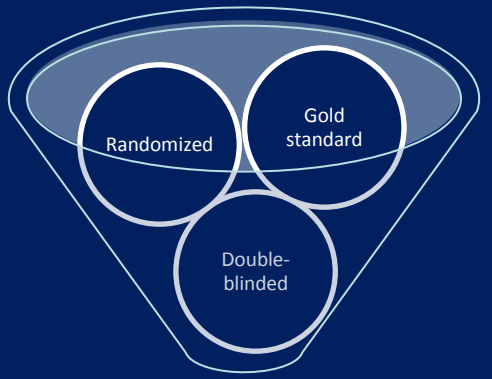

**Observational Design**

Is the unassigned tx/marker strongly associated with the outcome?




**Pragmatic RCT**

Is the assigned tx/marker caused the outcome in the 'real world'?




**Explanatory RCT**

Is the assigned tx/marker caused the outcome in an experimental setting?

Validity: External ← Internal

# When are RCTs more suitable to address CER GPM questions?

- Decisions require the highest level of certainty
- Detecting small or modest differences in the results of treatment/testing
- Ensure high level of internal validity
  - Control for selection bias, patient compliance and other confounding factors
- Accessible biospecimens are required for all participants
- Require detailed information on outcomes
- Incorporate genomic markers in design
- Complex testing or multi-therapy treatment

# When are observational studies more suitable to address CER GPM questions?

- Study populations not represented in clinical trials
  - Comorbidities, age, medication
- Larger studies and diverse populations are needed
  - rare outcomes or analysis of subgroups
- Need long-term follow-up
- a RCT is not ethical or feasible
- Treatments/testing are used off-label
- Compare outcomes from multiple treatment regimens

# When are observational studies more suitable to address CER GPM questions? (Continued)

- Detect larger differences in the results of treatment/testing
- Confirm result from RCTs
- Generate hypotheses to be tested in RCTs
- Study results need to be generalizable
- Study results are needed quickly
- Treatment adherence differs

# What are the major concerns about using observational studies to inform clinical utility?

- Poorly designed studies
- Difficulty replicating findings
- Difficulty in obtaining reliable outcome measures
- Bias and confounding
  - Selection, response, adherence, attrition, misclassification

# How can we improve our observational studies to inform clinical utility?

- Poorly designed studies
  - http://www.graceprinciples.org/
  - http://www.ispor.org/workpaper/practices_index.asp
  - ENCePP The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP®)

- Difficulty to replicate findings
  - Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)
  - Strengthening the Reporting of Genetic Associations (STREGA)
  - Genetic Risk Prediction Studies (GRIPS)

# How can we improve our observational studies to inform clinical utility? (Continued)

- Difficulty in obtaining reliable outcome measures
  - Validation studies (e.g. chart review)

- Bias and confounding
  - Instrumental variables
  - Propensity score matching
  - Matching and stratification
  - Prior event rate ratio
  - Sensitivity analysis

# Decision modeling

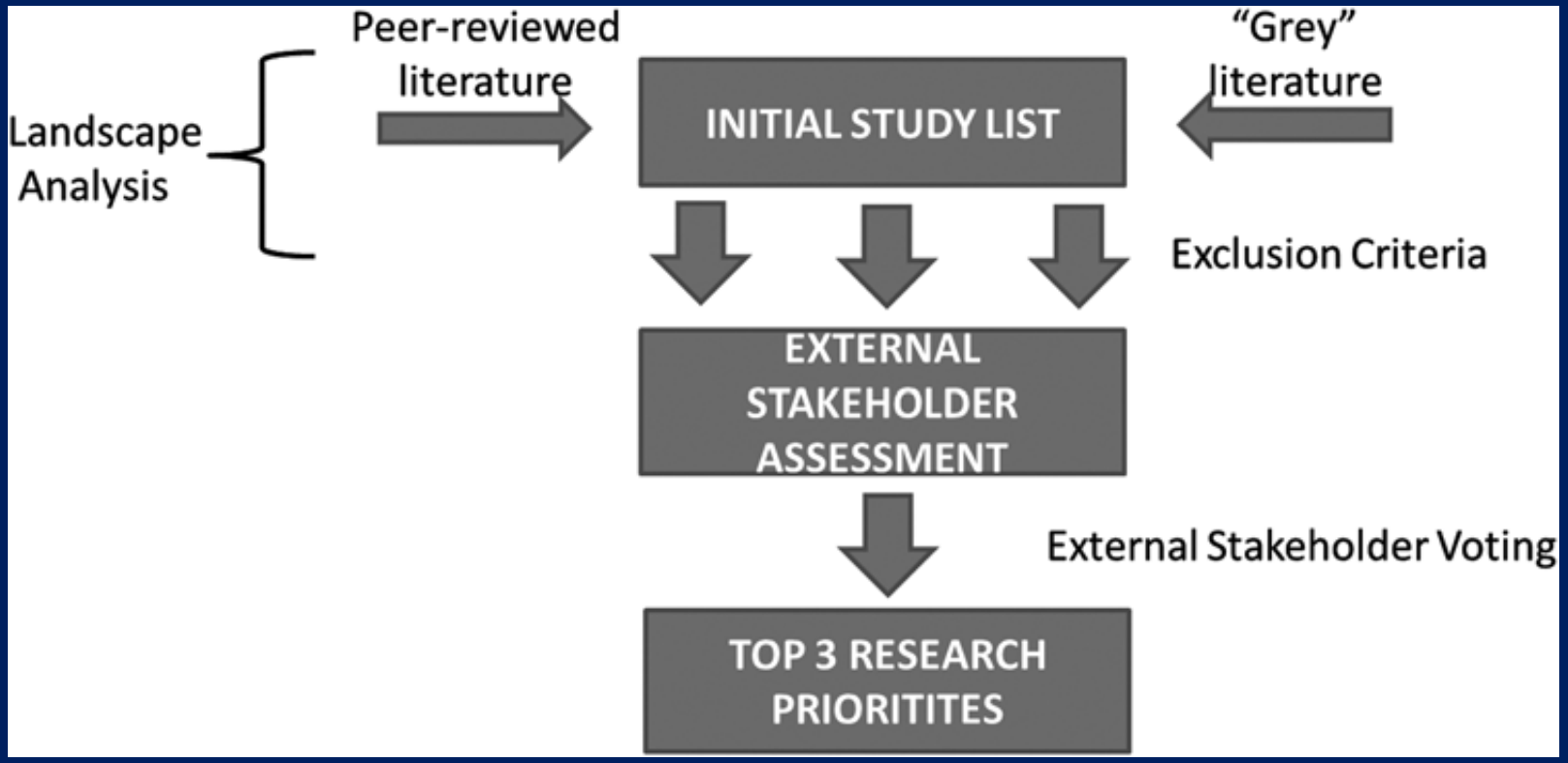
- Value of Information analysis/scenerio modeling
- Cost-effectiveness analysis
- Benefit-Risk modeling to facilitate evaluation of indirect evidence

# Risk-Benefit Policy Matrix

| Risk-Benefit Profile | Uncertainty: High | Uncertainty: Moderate | Uncertainty: Low |
|---|---|---|---|
| Favorable | Use with evidence-development | Consider use in clinical practice | Appropriate for use in clinical practice |
| Neutral | Do not use, conduct additional research | Use with evidence-development | Consider use in clinical practice |
| Unfavorable | Do not use, conduct additional research | Do not use | Do not use |

Veenstra DL et al. A formal risk-benefit framework for genomic tests: facilitating the appropriate translation of genomics into clinical practice. *Genetics in medicine : official journal of the American College of Medical Genetics.* Nov 2010;12(11):686-693

# Stakeholder Engagement




Thariani et al. Prioritization in Comparative Effectiveness Research: The CANCERGEN Exprerience. Med Care 2012

# Decision-Makers Key Questions for Cancer Genomic Medicine (Revisited)

1. Does the genomic application provide correct information? (analytic validity)

2. Is there a significant association between the results of the genomic application and clinical phenotype? (clinical validity)

3. Does the genomic application provide clinically significant information? (clinical utility)

4. Does the genomic application lead to improved patient outcomes as compared with the alternative? (comparison or added clinical value)

# KRAS mutation and anti-EGFR therapy treatment response in mCRC

## Analytic Validity

| | |
|---|---|
| **Question** | How well can we measure KRAS mutation? |
| **Problem** | Test characteristics depend on tumor heterogeneity, sample handling, slide preparation, techniques for tumor enrichment, DNA preparation, assay design and sensitivity |
| **Study Approach** | Compare KRAS test results in large multi-center observational study |
| **Results** | 90% concordance across 5 labs. 10% discordance due to tumor heterogeneity or contamination of the tumor sample with normal tissue |

Feigelson HS et al. BMC Research Notes 2012

# KRAS mutation and anti-EGFR therapy treatment response in mCRC

## Clinical Validity

| | |
|---|---|
| **Question** | Is KRAS mutation associated with treatment response? |
| **Problem** | Limited evidence from randomized studies exists. Lack of patient-level data to assess modifiers (i.e. pathologic and prognostic tumor characteristics) of the mutation-by-treatment interaction. Publication bias could be a concern |
| **Study Approach** | Systematic review of cohort studies and retrospective analysis of trials |
| **Results** | KRAS mutations are consistently associated with reduced overall and progression survival and increased failure rates among patients with mCRC treated with anti-EGFR therapies |

Dahabreh IJ et al. Annals of Internal Medicine April 2011

# KRAS mutation and anti-EGFR therapy treatment response in mCRC

## Clinical Utility

| | |
|---|---|
| **Question** | What are the benefits and harms of KRAS mutation testing for treatment decisions? |
| **Harms** | Potential for anti-EGFR therapy to be effective for some (small %) of patients whose tumors are KRAS mutant |
| **Benefits** | Avoiding serious adverse side effects in patients who will not respond to anti-EGFR therapies<br><br>Avoiding use of an expensive treatment for patients who will not benefit from treatment |

# KRAS mutation and anti-EGFR therapy treatment response in mCRC

## Added Clinical Value

| | |
|---|---|
| **Question** | Is the use KRAS mutation testing better than the alternative? |
| **Comparator** | Anti-EGFR treatment decisions in the absence of knowledge of KRAS mutation status |
| **Study Approach** | Observational cohort study pre and post KRAS testing, Prospective RCT? |

# Genomic Predictive Markers of Cancer Treatment Efficacy and Safety

| Markers | Drugs | Studies informing Clinical Utility |
|---|---|---|
| ***In Clinical Use*** | | |
| HER2/neu | Trastuzumab | **Genome-guided RCTs** |
| Oncotype Dx, | Treatment regimen | **Retrospective analysis of biospecimens from RCTs, Retrospective Cohort Study** |
| EGFR Mutation | Erlotinib | **Prospective Cohort Studies** |
| K-ras | Cetuximab, Panitumumab | **Retrospective analysis of biospecimens from RCTs** |
| EML4-ALK mutation | Crizotinib | **Genome-guided RCT** |
| BRAF V600E | Vemurafenib | **Genome-guided RCT** |
| BCR-ABL | Imatinib, Dasatinib, Nilotinib | **RCT** |
| C-Kit | Imatinib | **Genome-guided RCT** |
| *TPMT* | *6-MP, 6-TG* | **Prospective and Retrospective Cohort studies, Case-Control** |

# Need a comprehensive approach to resolve questions about the clinical utility of genomic applications

- **Future research must consider more outcome measures, conducted in settings that are relevant to more real-world clinical decisions**
- **A multitude of stakeholders should have a role in evidence generation**
- **Consider new strategies involving transformation of the research infrastructure to "learning systems" that allow continual addition to the evidence base.**
- **Clear priorities for CER must be identified to ensure that limited resources are used to resolve the most compelling questions**

# Need for an evidentiary framework to clearly define evidence standards for clinical utility

- **Recognize that an RCT is not desirable or feasible in every circumstance and high quality observational study designs and evidence of underlying biological mechanisms can contribute to the evidentiary framework.**

- **Any reforms of the evidentiary framework should uphold rigorous standards existing best research practices**

- **An evidentiary framework needs to articulate the minimal evidence necessary before genomic clinical application is warranted**

# Summary

**Health policy decisions must take into consideration the clinical context, the type of genomic application, the quality and availability of evidence to assess a *marker's* benefits and risks, and the risk to patients that a wrong decision could pose.**

# Collaborators

Kaiser Permanente Northwest
Katrina A.B. Goddard PhD and Evelyn Whitlock MD

Kaiser Permanente Colorado
Heather Spencer Feigelson PhD

University of Virginia
William A. Knaus MD

Duke University and the Duke Cancer Institute
Gary H. Lyman MD, MPH

National Cancer Institute
Sheri D. Schully PhD

National Cancer Institute and Centers for Disease Control
Muin J. Khoury MD, PhD

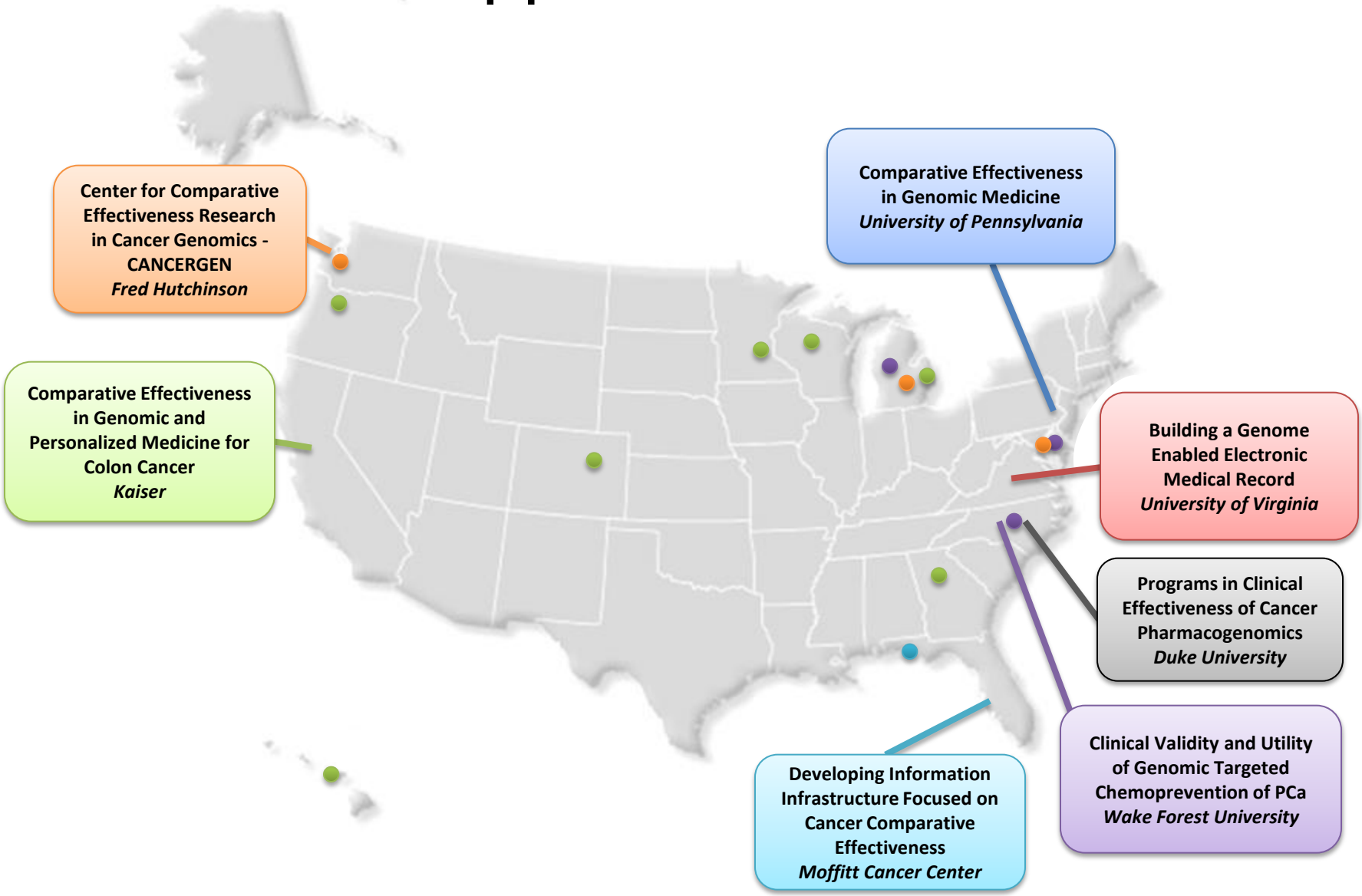
Fred Hutchinson Cancer Research Center
Scott Ramsey MD, PhD

Center for Medical Technology Policy
Sean Tunis MD, MSc

University of Washington
David L. Veenstra PharmD, PhD

Genet Med. 2012 Apr 19

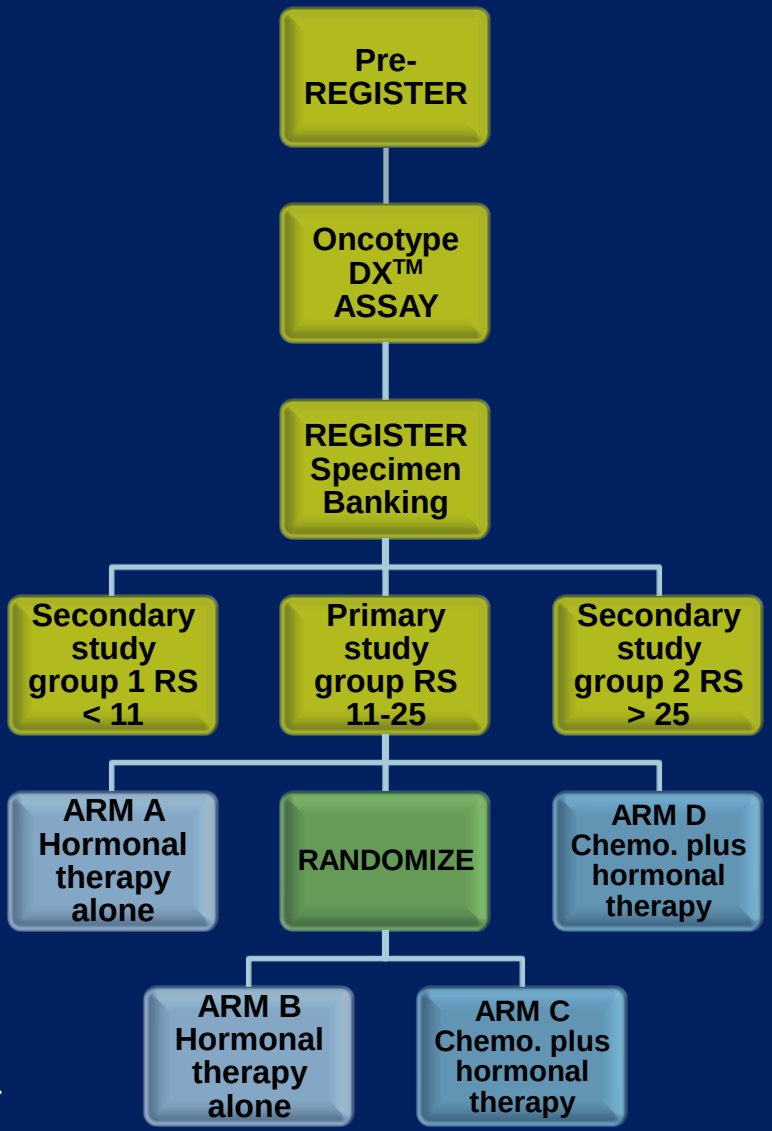
# NCI-Supported CER in GPM





NATIONAL® CANCER INSTITUTE

# Trial Assessing Individualized Options for Treatment for Breast Cancer (TAILORx)




RS: Recurrence score

Zujewski JA, Kamin L. *Future Oncol*. 2008.