Understanding your biomarker: what this marker can do for you

Dr. John Bartlett/Dr. Harriet Feilotter

"As is your Pathology, so is your Medicine." Sir William Osler. 1849-1919.

Disclosure statement

John Bartlett discloses:

- Consulting/advisory roles with:
  - Insight Genetics
  - BioNTech GmbH
- Research funding (in kind) from:
  - NanoString
  - Stratifyer
  - Mammaprint
  - Genoptix
- Three pending patents relating to biomarkers in breast cancer

Harriet Feilotter discloses:

- CSO with Indoc Research
- Research partner with Life Technologies (ThermoFisher)

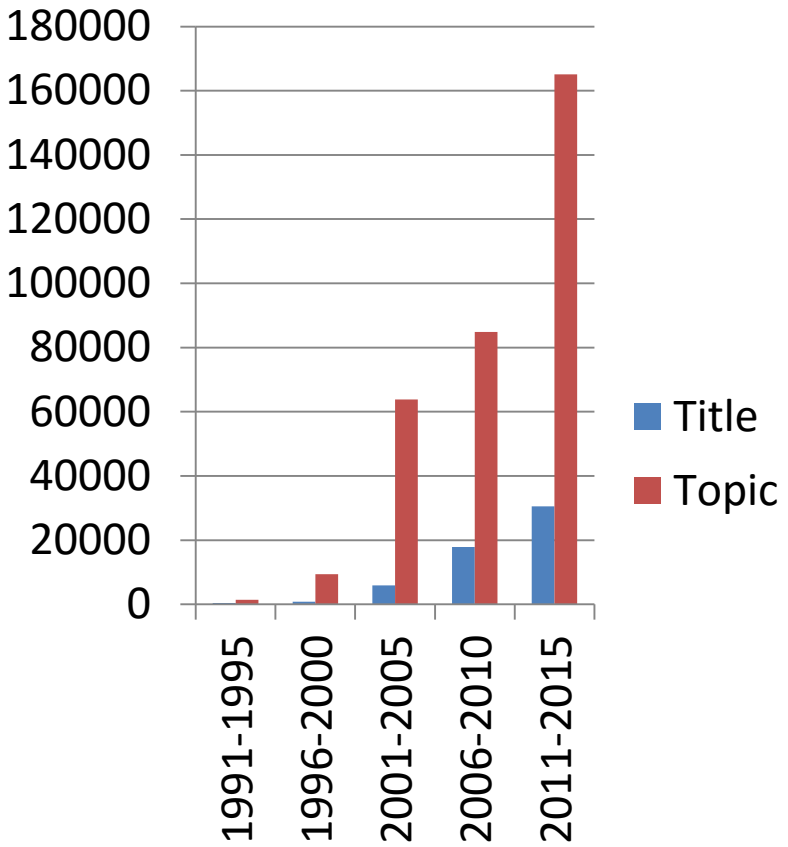
Objectives

- To understand major guidelines for biomarker used in the clinical or clinical trial setting
- To apply the concepts of clinical validity, clinical utility, analytical validity and cost assessment to a commonly used biomarker
- To appreciate issues that may impact the performance of a biomarker in the clinical setting

Biomarkers

a distinct biochemical, genetic, or molecular characteristic or substance that is an indicator of a particular biological condition or process.

- YAPI (Yet another prognostic indicator)



Simplistic WOS search “biomarker “ in Title/topic. See Hall PA, Going JJ “Predicting the future a critical appraisal of cancer prognosis studies” Histopathology 1999 35:489-494.

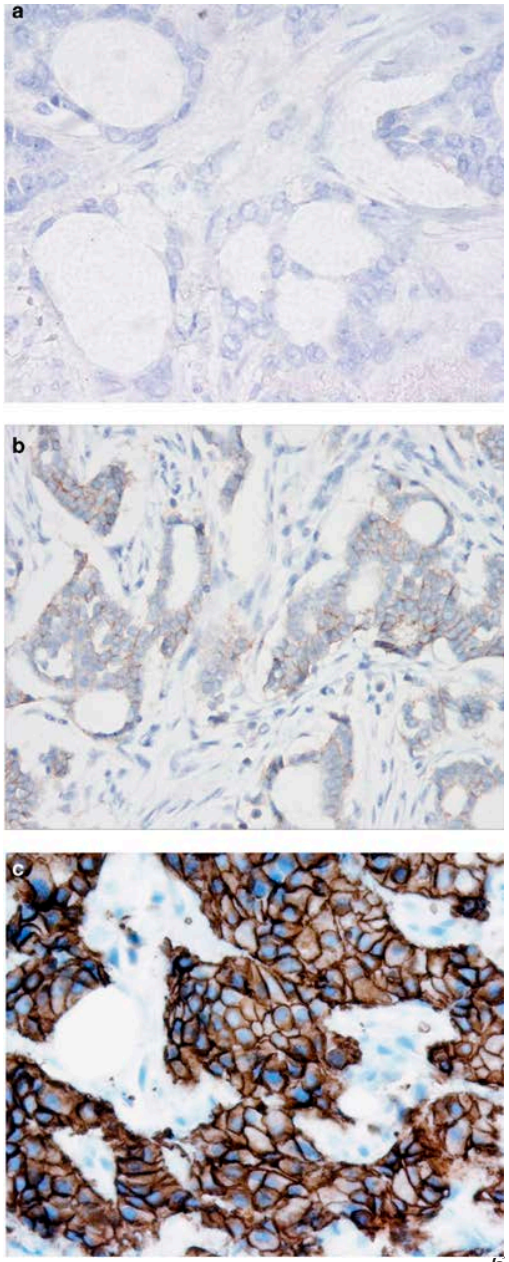
EGAPP Guidelines for diagnostic biomarkers

- **Clinical Utility**
  - Addresses key clinical question
  - Changes treatment plan.
- **Clinical Validity**
  - Level I Evidence of prognostic or predictive value
    - Prospective or retrospective
  - Precise, Reproducible
- **Analytical Validity**
  - Accurate: “A precise definition of accuracy is how close the measured values are to a supposed true value”
- **Cost effective**
  - Cheaper/more effective than current approaches

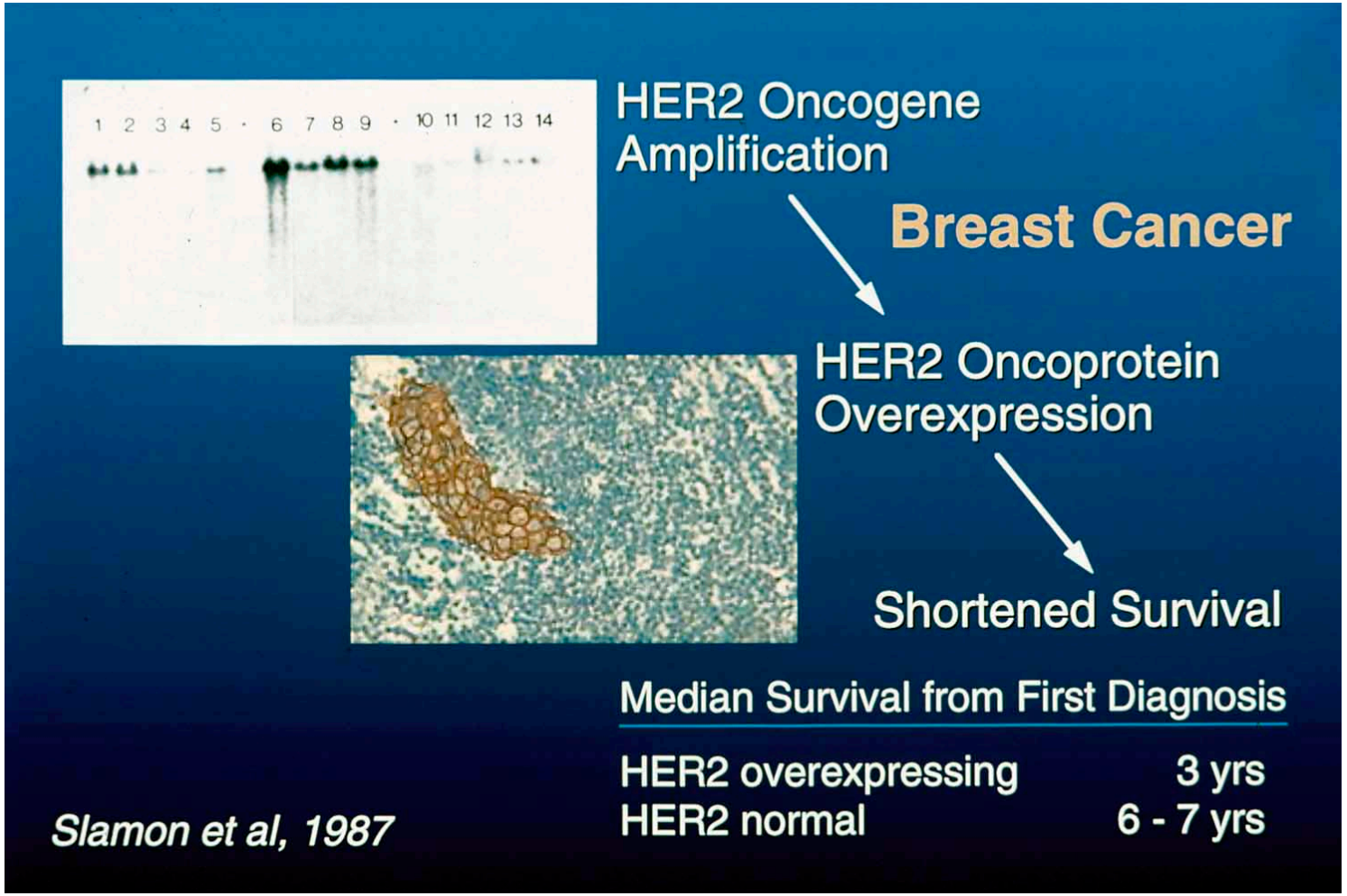
Teutsch et al, Genetics Medicine 2009 11:3-14
Wolff AC et al, Journal of Clinical Oncology 2007 25:118-145

HER2 or HER2/neu or ERBB2

- Human epidermal growth factor receptor 2
- Cell surface protein
- Amplified in some cancers, particularly breast cancer
- Amplification of locus can be seen by in situ assays, including IHC, FISH, CISH
- Tumours with amplified HER2 candidates for trastuzumab (Herceptin)



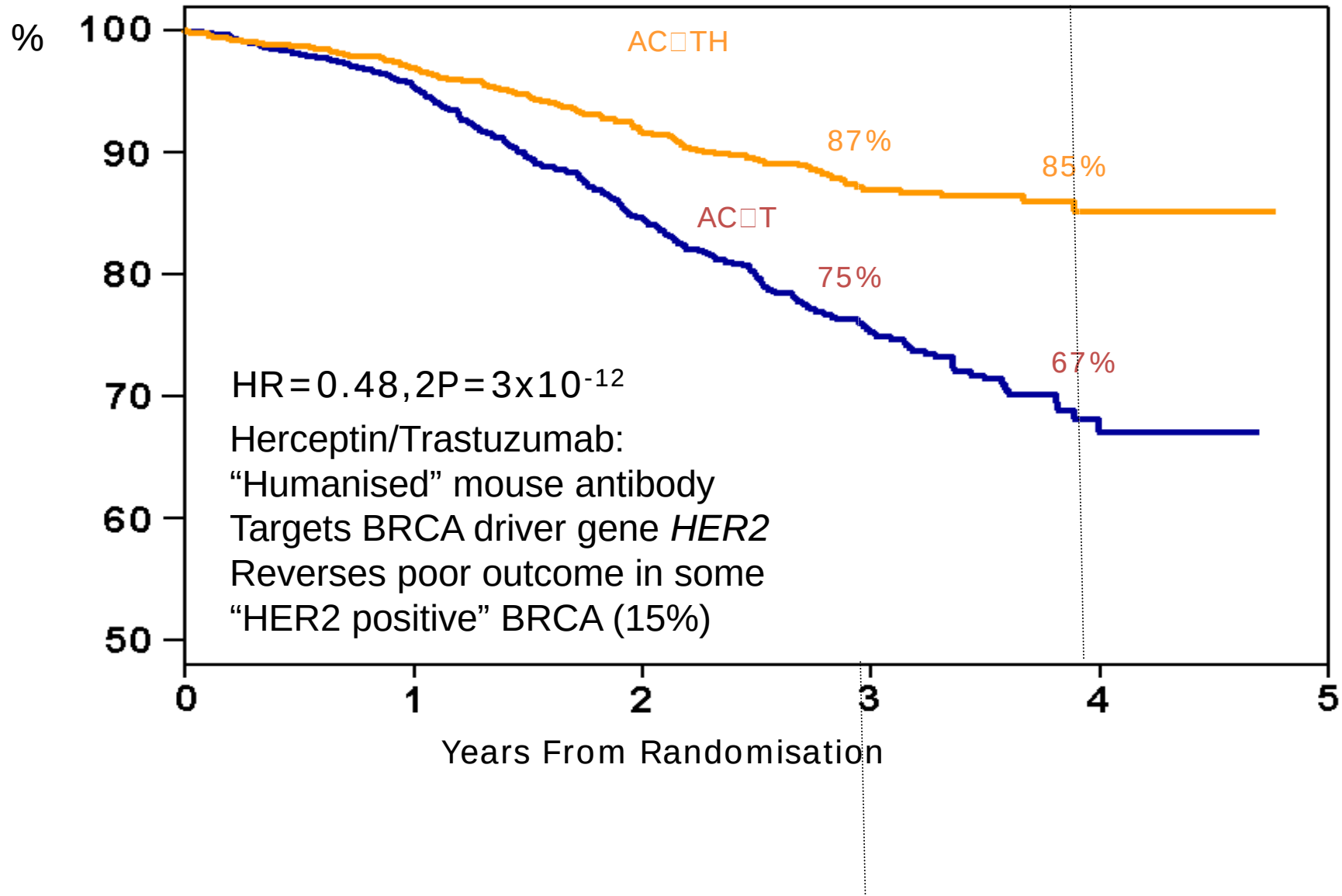
HER2 – the test that came in from the cold:



Clinical utility: If the answer is “biomarker test” – what’s the question?

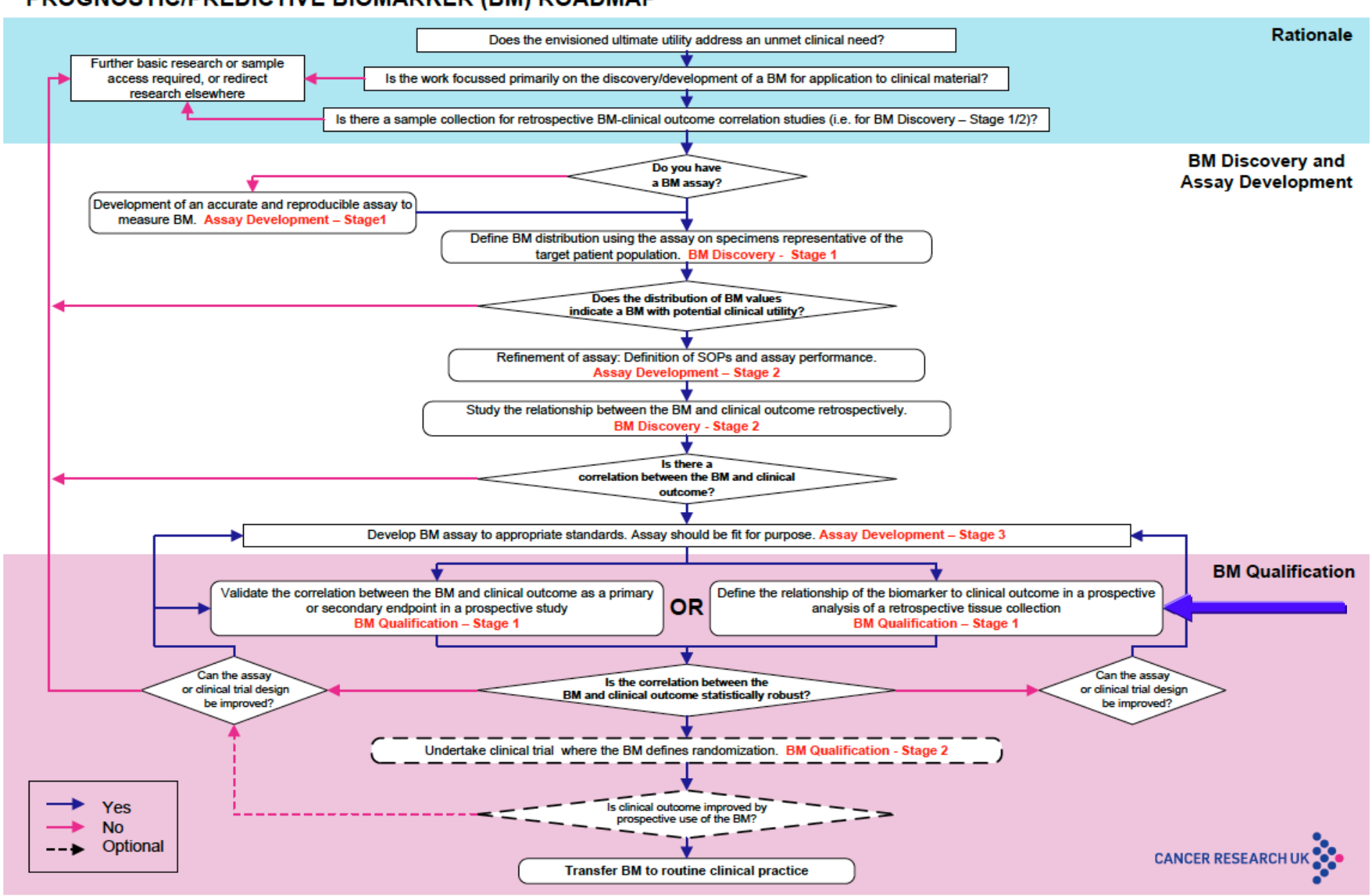
- Addresses the question "will it change the treatment approach?"
- How useful are test results to the person tested.
- May depend on effective access to appropriate interventions and an individual’s willingness to adopt the recommended interventions.

HER2 – selective for Herceptin Benefit (2005!)



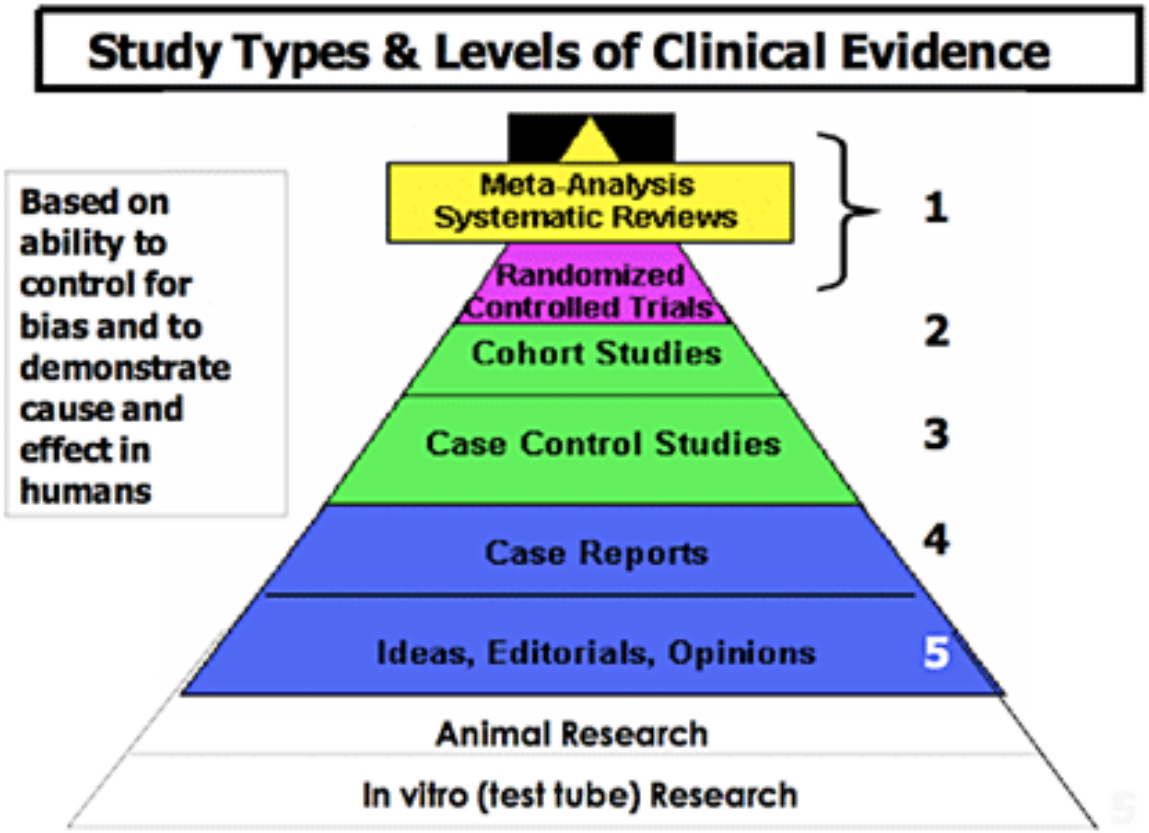
Clinical Validity:

- *Level 1 evidence of prognostic or predictive value*
- *Can be prospective or retrospective*



If you do not ask these questions there is a high probability your biomarker will fail!

Clinical Validity



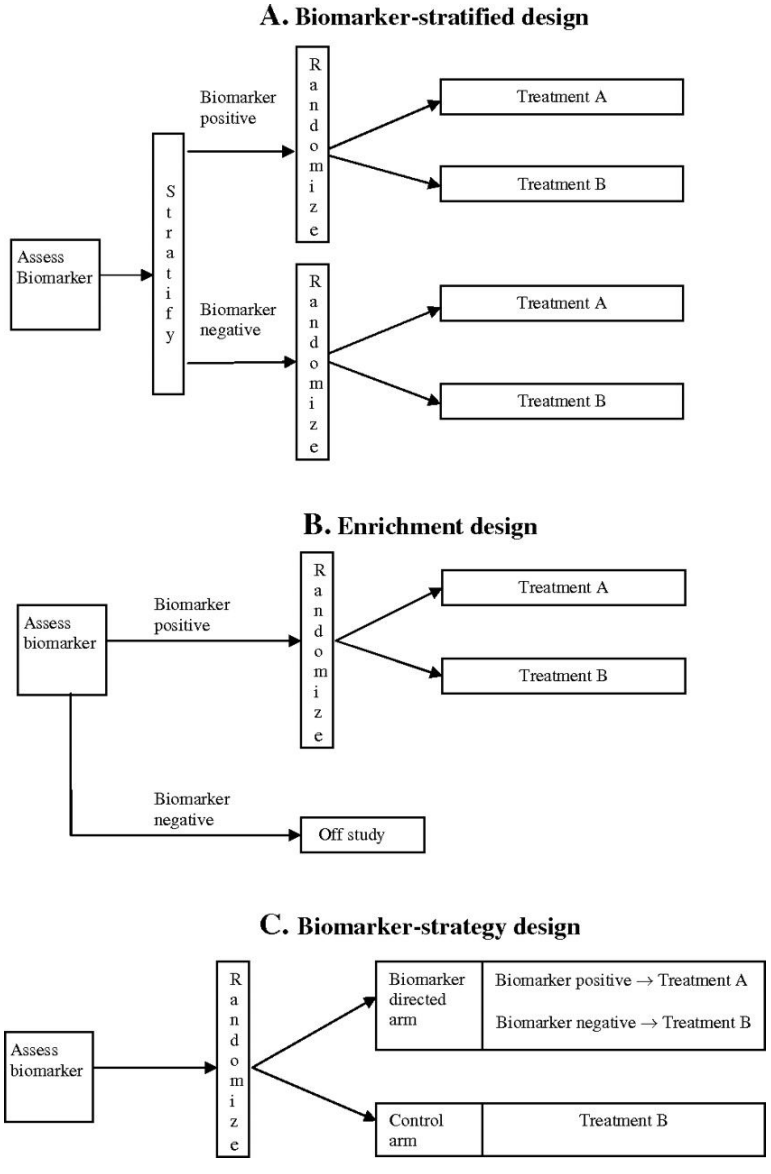
“Testing the test”: level 1 evidence:

1. *Prospective Clinical Trials: Marker is Primary Objective*
   - ***Sargent D.J., et al. J Clin Oncol. 23:2020-7, 2005***
   - ***Freidlin B., et al. J Natl Cancer Inst. 102:152-60, 2010***
2. *Analysis of RETROSPECITVE tissue banks*
   - *Is a Prospective Trial Always Necessary?*
     - ***NO! But use of archived tissue must be done with rigor***
       - *Simon R.M., Paik S, Hayes DF. J Natl Cancer Inst. 101:1446-52, 2009*

"Testing the test" reaching level 1:

1. *Prospective Clinical Trials: Marker is Primary Objective = Level 1(a)*
   - ***Sargent D.J., et al. J Clin Oncol. 23:2020-7, 2005***
   - ***Freidlin B., et al. J Natl Cancer Inst. 102:152-60, 2010***
2. *Analysis of RETROSPECITVE tissue banks*
   - *Is a Prospective Trial Always Necessary?*
     - ***NO! But use of archived tissue must be done with rigor***
       - *Simon R.M., Paik S, Hayes DF. J Natl Cancer Inst. 101:1446-52, 2009*

Biomarker trial designs.



A) Biomarker-stratified design. All patients are randomly assigned regardless of biomarker status with the random assignment and analysis plan stratified by the biomarker status.

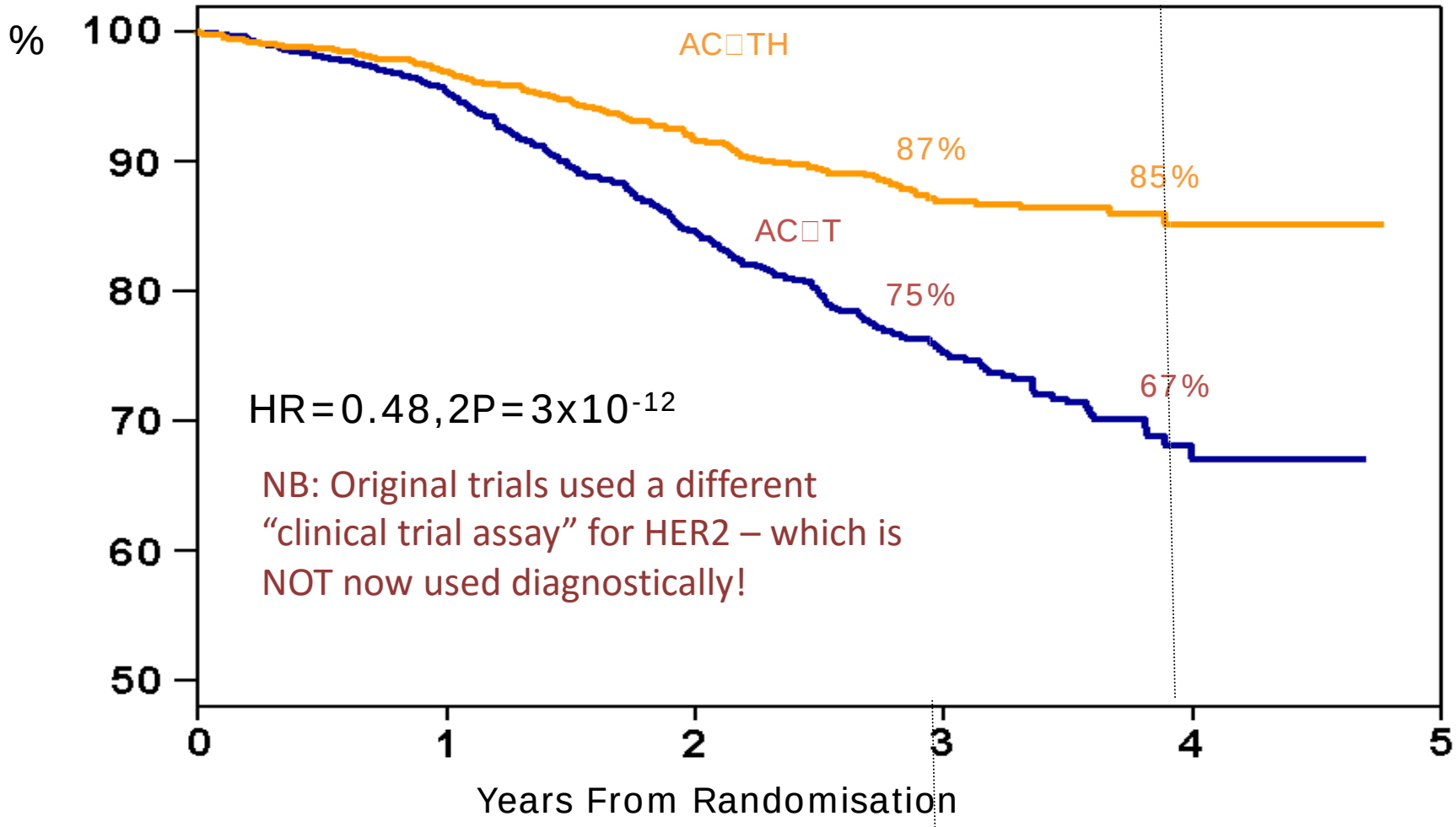
B) Enrichment design. The biomarker is evaluated on all patients, but random assignment is restricted to patients with specific biomarker values.

C) Biomarker-strategy design. Patients are randomly assigned to an experimental treatment arm that uses the biomarker to direct therapy or to a control arm that does not.

Freidlin B et al. JNCI J Natl Cancer Inst 2010;102:152-160

JNCI

HER2 – selective for Herceptin Benefit – enrichment design.

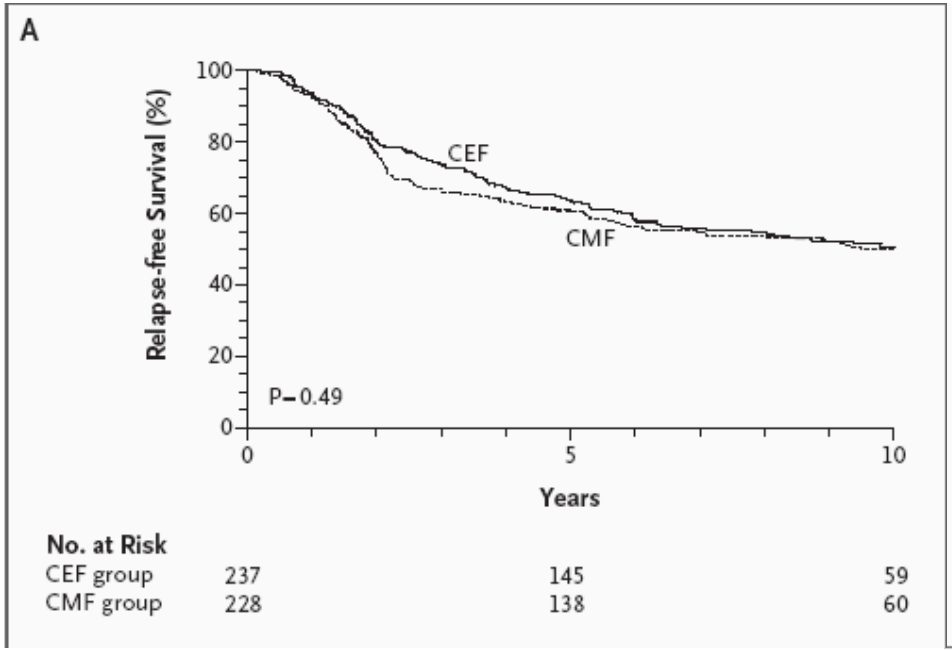


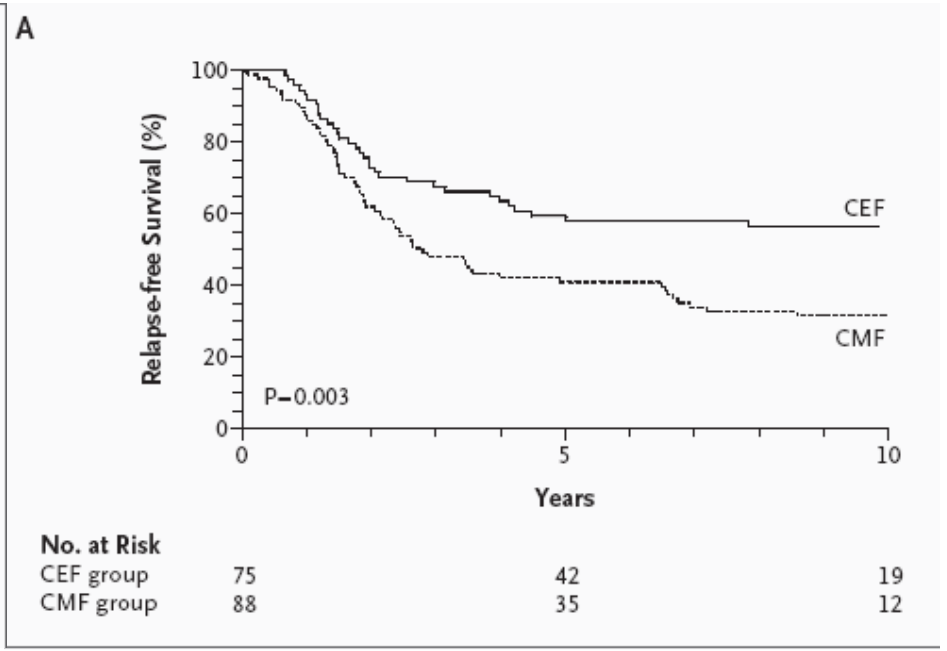
"Here we are 10 years into it, and we don't know how to test for it."

“Testing the test” reaching level 1:

1. *Prospective Clinical Trials: Marker is Primary Objective*
   - ***Sargent D.J., et al. J Clin Oncol. 23:2020-7, 2005***
   - ***Freidlin B., et al. J Natl Cancer Inst. 102:152-60, 2010***
2. *Analysis of RETROSPECITVE tissue banks*
   - *Is a Prospective Trial Always Necessary?*
     - ***NO:*** *But use of archived tissue must be done with rigor to reach level 1(b)*
       - *Simon R.M., Paik S, Hayes DF. J Natl Cancer Inst. 101:1446-52, 2009*

HER2 and anthracycline response in the Canadian MA5 study





HER2 negative HER2+ve (amplified)

Treatment by marker interaction p = 0.01, HR 1.96.

Pritchard KI et al, NEJM(!) 2006 354:2103-2111.

HER2 and anthracycline response:

- Prospectively planned treatment by marker analysis in a retrospective analysis of clinical trial.
- Prospectively defined biomarker cut point.
- Statistically significant result.
- Consistent with appropriate guidelines
  - REMARK: Reporting recommendations for tumor MARKer prognostic studies. McShane L et al, Nature Clinical Practice Oncology (2005) 2:416-422 (joint published in 6 journals!).
- What Level of evidence? Did it change practice?

Revised LOI Scale: Use of Archived Tissues

Level of Evidence	Category from Table 1	Validation Studies Available
I	A	None required
I	B	One or more with consistent results
II	B	None or Inconsistent results
II	C	2 or more with consistent results
III	C	None or 1 with consistent results or Inconsistent results
IV-V	D	NA

Simon R.M., Paik S, Hayes DF. J Natl Cancer Inst. 101:1446-52, 2009

Use of Archived Tissues To Determine Clinical Utility of Tumor Markers

Category	A	B	C	D
Trial Design	Prospective	Prospective using archived samples	Prospective /observational	Retrospective /observational
Clinical trial	PRCT designed to address tumor marker	Prospective trial not designed to address tumor marker, but design accommodates tumor marker utility. Accommodation of predictive marker requires PRCT	Prospective observational registry, treatment and follow up not dictated	No prospective aspect to study
Patients and patient data	Prospectively enrolled, treated, and followed in PRCT	Prospectively enrolled, treated, and followed in clinical trial and, especially if a predictive utility is considered, a PRCT addressing the treatment of interest	Prospectively enrolled in registry, but treatment and follow up standard of care	No prospective stipulation of treatment or follow up; patient data collected by retrospective chart review
Specimen collection, processing, and archival	Specimens collected, processed and assayed for specific marker in real time	Specimens collected, processed, and archived prospectively using generic SOPs. Assayed after trial completion.	Specimens collected, processed, and archived prospectively using generic SOPs. Assayed after trial completion.	Specimens collected, processed and archived with no prospective SOPs
Statistical Design and analysis	Study powered to address tumor marker question.	Study powered to address therapeutic question; underpowered to address tumor marker question. Focused analysis plan for marker question developed prior to doing assays	Study not prospectively powered at all. Retrospective study design confounded by selection of specimens for study. Focused analysis plan for marker question developed prior to doing assays	Study not prospectively powered at all. Retrospective study design confounded by selection of specimens for study. No focused analysis plan for marker question developed prior to doing assays
Validation	Result unlikely to be play of chance Although preferred, validation not required	Result more likely to be play of chance that A, but less likely than C. Requires one or more validation studies	Result very likely to be play of chance. Requires subsequent validation studies	Result very likely to be play of chance. Requires subsequent validation

Simon R.M., Paik S, Hayes DF. J Natl Cancer Inst. 101:1446-52, 2009

Use of Archived Tissues To Determine Clinical Utility of Tumor Markers

Category	A	B	C	D
Trial Design	Prospective	Prospective using archived samples	Prospective /observational	Retrospective /observational
Clinical trial	PRCT designed to address tumor marker	Prospective trial not designed to address tumor marker, but design accommodates tumor marker utility. Accommodation of predictive marker requires PRCT	Prospective observational registry, treatment and follow up not dictated	No prospective aspect to study
Patients and patient data	Prospectively enrolled, treated, and followed in PRCT	Prospectively enrolled, treated, and followed in clinical trial and, especially if a predictive utility is considered, a PRCT addressing the treatment of interest	Prospectively enrolled in registry, but treatment and follow up standard of care	No prospective stipulation of treatment or follow up; patient data collected by retrospective chart review
Specimen collection, processing, and archival	Specimens collected, processed and assayed for specific marker in real time	Specimens collected, processed, and archived prospectively using generic SOPs. Assayed after trial completion.	Specimens collected, processed, and archived prospectively using generic SOPs. Assayed after trial completion.	Specimens collected, processed and archived with no prospective SOPs
Statistical Design and analysis	Study powered to address tumor marker question.	Study powered to address therapeutic question; underpowered to address tumor marker question. Focused analysis plan for marker question developed prior to doing assays	Study not prospectively powered at all. Retrospective study design confounded by selection of specimens for study. Focused analysis plan for marker question developed prior to doing assays	Study not prospectively powered at all. Retrospective study design confounded by selection of specimens for study. No focused analysis plan for marker question developed prior to doing assays
Validation	Result unlikely to be play of chance Although preferred, validation not required	Result more likely to be play of chance that A, but less likely than C. Requires one or more validation studies	Result very likely to be play of chance. Requires subsequent validation studies	Result very likely to be play of chance. Requires subsequent validation

Simon R.M., Paik S, Hayes DF. J Natl Cancer Inst. 101:1446-52, 2009

Use of Archived Tissues To Determine Clinical Utility of Tumor Markers

Category	A	B	C	D
Trial Design	Prospective	Prospective using archived samples	Prospective /observational	Retrospective /observational
Clinical trial	PRCT designed to address tumor marker	Prospective trial not designed to address tumor marker, but design accommodates tumor marker utility. Accommodation of predictive marker requires PRCT	Prospective observational registry, treatment and follow up not dictated	No prospective aspect to study
Patients and patient data	Prospectively enrolled, treated, and followed in PRCT	Prospectively enrolled, treated, and followed in clinical trial and, especially if a predictive utility is considered, a PRCT addressing the treatment of interest	Prospectively enrolled in registry, but treatment and follow up standard of care	No prospective stipulation of treatment or follow up; patient data collected by retrospective chart review
Specimen collection, processing, and archival	Specimens collected, processed and assayed for specific marker in real time	Specimens collected, processed, and archived prospectively using generic SOPs. Assayed after trial completion.	Specimens collected, processed, and archived prospectively using generic SOPs. Assayed after trial completion.	Specimens collected, processed and archived with no prospective SOPs
Statistical Design and analysis	Study powered to address tumor marker question.	Study powered to address therapeutic question; underpowered to address tumor marker question. Focused analysis plan for marker question developed prior to doing assays	Study not prospectively powered at all. Retrospective study design confounded by selection of specimens for study. Focused analysis plan for marker question developed prior to doing assays	Study not prospectively powered at all. Retrospective study design confounded by selection of specimens for study. No focused analysis plan for marker question developed prior to doing assays
Validation	Result unlikely to be play of chance Although preferred, validation not required	Result more likely to be play of chance that A, but less likely than C. Requires one or more validation studies	Result very likely to be play of chance. Requires subsequent validation studies	Result very likely to be play of chance. Requires subsequent validation

Simon R.M., Paik S, Hayes DF. J Natl Cancer Inst. 101:1446-52, 2009

Use of Archived Tissues To Determine Clinical Utility of Tumor Markers

Category	A	B	C	D
Trial Design	Prospective	Prospective using archived samples	Prospective /observational	Retrospective/observational
Clinical trial	PRCT designed to address tumor marker	Prospective trial not designed to address tumor marker, but design accommodates tumor marker utility. Accommodation of predictive marker requires PRCT	Prospective observational registry, treatment and follow up not dictated	No prospective aspect to study
Patients and patient data	Prospectively enrolled, treated, and followed in PRCT	Prospectively enrolled, treated, and followed in clinical trial and, especially if a predictive utility is considered, a PRCT addressing the treatment of interest	Prospectively enrolled in registry, but treatment and follow up standard of care	No prospective stipulation of treatment or follow up; patient data collected by retrospective chart review
Specimen collection, processing, and archival	Specimens collected, processed and assayed for specific marker in real time	Specimens collected, processed, and archived prospectively using generic SOPs. Assayed after trial completion.	Specimens collected, processed, and archived prospectively using generic SOPs. Assayed after trial completion.	Specimens collected, processed and archived with no prospective SOPs
Statistical Design and analysis	Study powered to address tumor marker question.	Study powered to address therapeutic question; underpowered to address tumor marker question. Focused analysis plan for marker question developed prior to doing assays	Study not prospectively powered at all. Retrospective study design confounded by selection of specimens for study. Focused analysis plan for marker question developed prior to doing assays	Study not prospectively powered at all. Retrospective study design confounded by selection of specimens for study. No focused analysis plan for marker question developed prior to doing assays
Validation	Result unlikely to be play of chance Although preferred, validation not required	Result more likely to be play of chance that A, but less likely than C. Requires one or more validation studies	Result very likely to be play of chance. Requires subsequent validation studies	Result very likely to be play of chance. Requires subsequent validation

Simon R.M., Paik S, Hayes DF. J Natl Cancer Inst. 101:1446-52, 2009

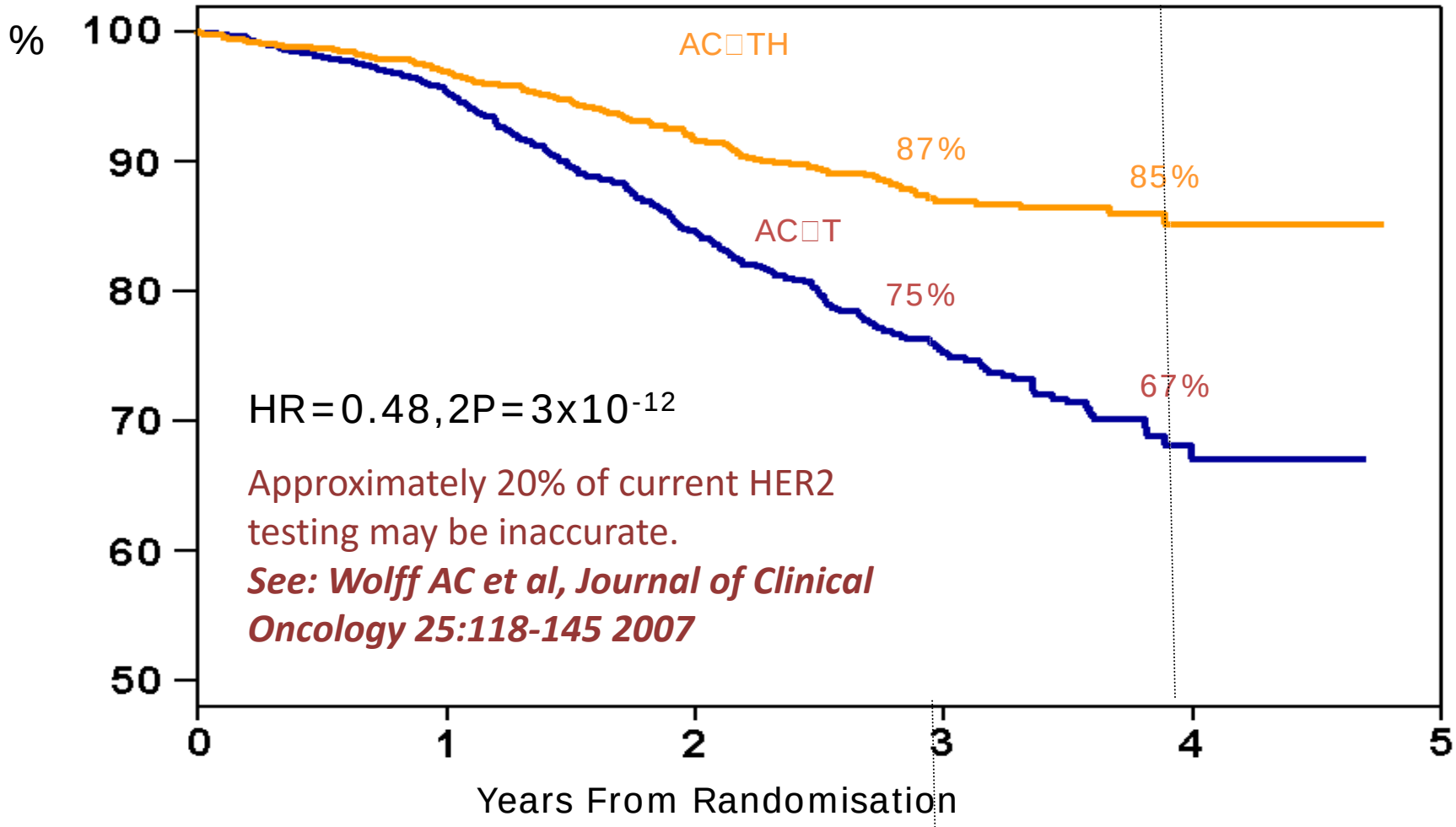
Retrospective validation: Tissue is the issue; BRISQ

- Tissue banks are critical to multiple stages of biomarker development.
  - https://www.ctrnet.ca/
- The "quality" of tissue banks impacts levels of evidence reached.
  - BRISQ: Biospecimen reporting for improved study quality J Proteome Res 2011, 10:3429-38.
- Level I evidence requires meta-analysis of 2 (or more) retrospective clinical trial tissue banks.
  - *Simon R.M., Paik S, Hayes DF. J Natl Cancer Inst. 101:1446-52, 2009*
  - REMARK: Reporting recommendations for tumor MARKer prognostic studies. McShane L et al, Nature Clinical Practice Oncology (2005) 2:416-422

Analytic validity: If the test result is not reproducible, what then?

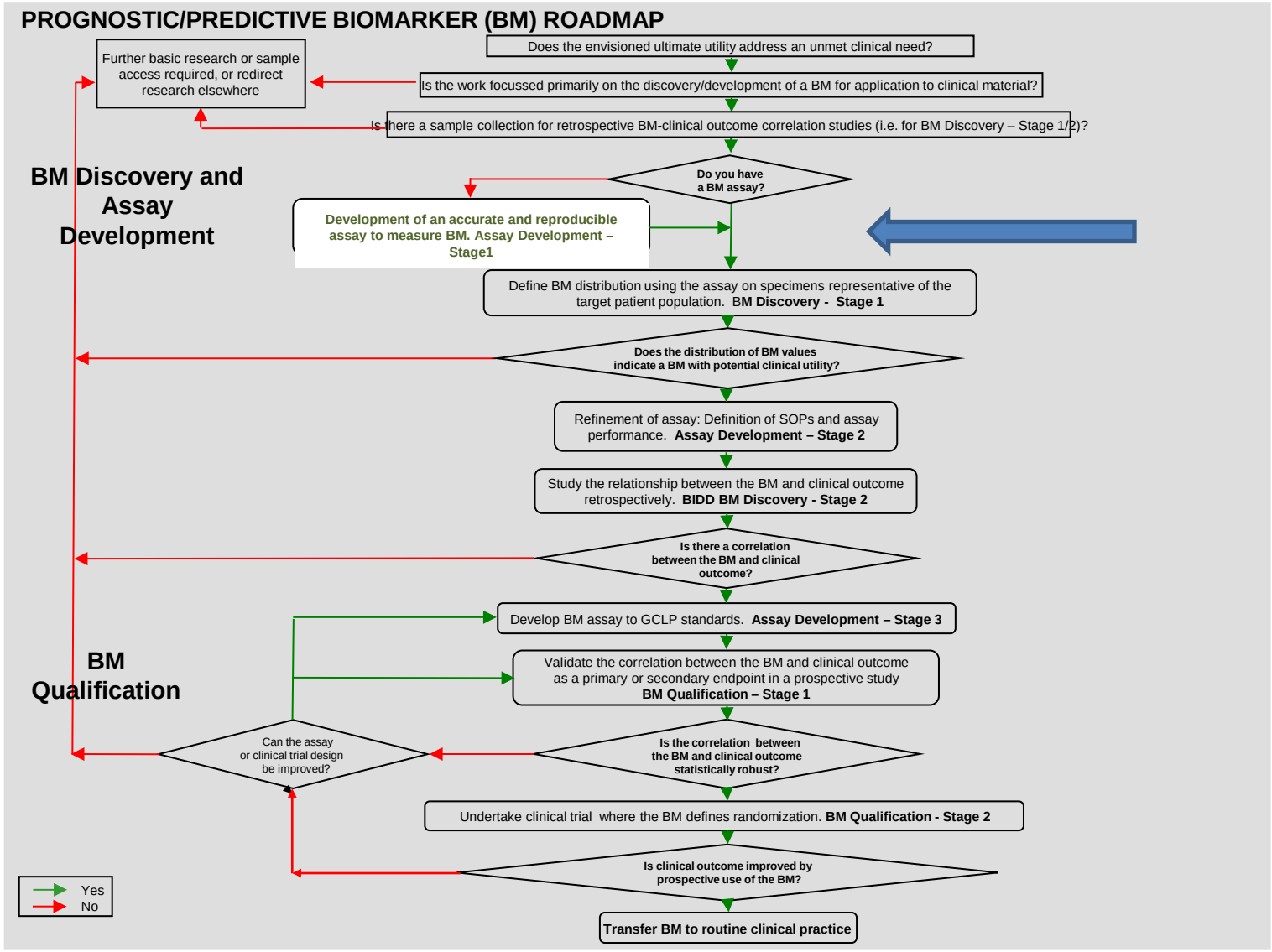
- Without analytic validity, you have worse than nothing
- Accuracy, reproducibility,
- Sensitivity, specificity
- Must be assessed for every test modality used

HER2 – selective for Herceptin Benefit



"Here we are 10 years into it, and we don't know how to test for it."

Analytical validity: without this .. nothing



ASCO Guidelines: ANALYTICAL

- Accuracy
  - How close the measured values are to a "true" value
  - Implicit that a suitable "gold standard" exists
- With binary measurements (eg positive vs. negative)
  - Sensitivity: % of positive test results when evaluating true positives
  - Specificity: % of negative test results when evaluating true negatives
  - Accuracy: % concordance between evaluated assay and gold standard

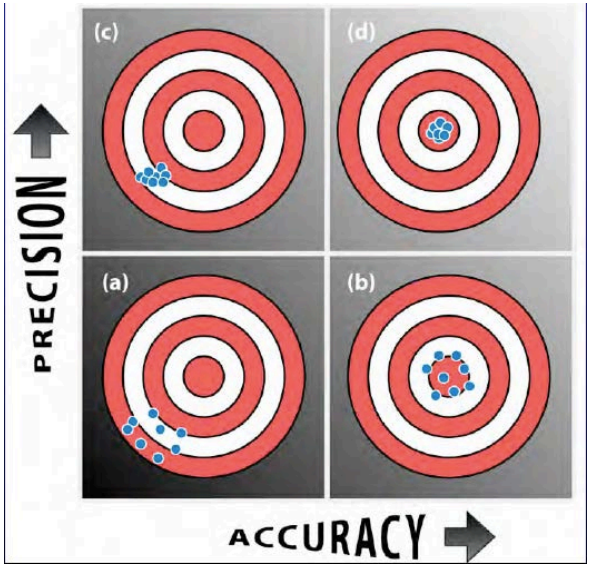
accurate determination of HER2 status must not be viewed exclusively in terms of benefit from anti-HER2 therapy, like trastuzumab.

Wolff AC et al, Journal of Clinical Oncology 25:118-145 2007

NB: Predictive value does NOT equal accuracy – it reflects clinical utility/validity.

ASCO 2007: Accuracy is paramount:

- "A precise definition of accuracy is how close the measured values are to a supposed true value"
  - *Which system most accurately determines HER2 status?*
  - *Which is least error prone in routine clinical practice?*



Measuring Accuracy HER2:

Q-IHC as a “gold standard” allows assessment of Accuracy
Ring studies show reproducibility

	Herceptest™	CB11	Pathvysion™	*HER2* copy
Concordance	0.81	0.84	0.93	0.91
Kappa	0.67	0.74	0.97	0.93
PPV	92.9%	63.2%	88.1%	78.4%
NPV	89.8%	94.7%	95.3%	97.1%
Accuracy	87.4%	83.8%	93.2%	91.6%

Concordance = concordance index (see text); Kappa = kappa statistic for inter-observer variation (see text); PPV = positive predictive value = percentage of true-positive results/all positive results; NPV = negative predictive value = percentage of true-negative results/all negative results. Accuracy = percentage (true positive + true negative)/all test results.

Accuracy may be influenced by:

- -extent of amplification
- -tumour heterogeneity
- -”cherry picking”
- -pre-analytic considerations
- -probe used
- -workflow in institution

What is your *analytical* “gold standard”?
If one does not exist – can you manufacture one?

Quality Assurance

- **Internal Quality Control (IQC)**
  - monitors within- & between-analytical run variability
  - Internal standards - plot performance over time.
    - *Ideally included in <u>every</u> analytical run.*
    - *In both research and clinical diagnostic setting*

- **External Quality Assessment (EQA)**
  - participating clinical laboratories are sent samples on a regular basis which they test as if they had come from patients.
  - Results are returned to EQA centres which provide a report that compares the participant's performance with that of all laboratories and/or groups of laboratories using the same test method(s).
  - Ensures high quality of testing – may affect ability to deliver tests.

Cost effectiveness:

- Includes the cost of the test
- Also includes the overall cost or savings to the health care system if the test is used
- Will be influenced by criteria for accessing test, test modality, and clinical utility

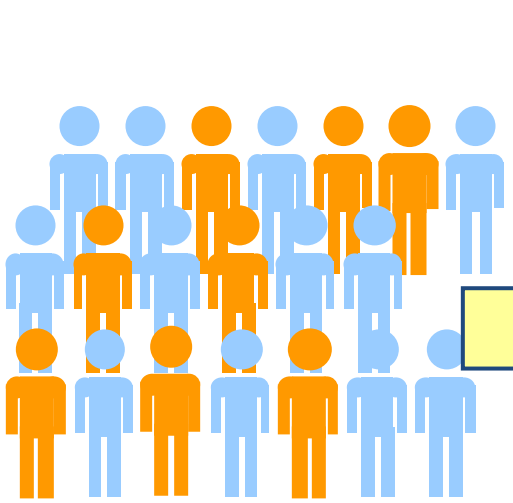
Study (Country) (Perspective) (Time horizon) (Discount rate)	Study aim	Methods	Population	Alternatives examined	Results (2007 US$)	Sensitivity analysis	Conclusions	Ref.
Dendukuri *et al* (2007) (Canada) (3rd party payer) (Cross-sectional) (N/A)	Meta-analysis of studies published between 2000 and 2005, to estimate the percentage of $HER2^+$ patients whose status was accurately determined by IHC and CEA of seven alternative strategies to test HER2 status	– Cross-sectional Bayesian meta-analysis to estimate the distribution of IHC scores and the probability of a positive FISH result in each IHC score category from a systematic review of IHC and FISH studies – CEA of screening strategies	Hypothetical cohcrt of BC patients tested with assays licensed by Health Canada	– S1: IHC + FISH confirmation of $IHC2^+$; $HER2^+$ defined as $IHC3^+$ or $FISH^+$ (base case) – S2: IHC only; $HER2^+$ defined as $IHC2^+$ or 3^+ – S3: IHC only; $HER2^+$ defined as $IHC3^+$ – S4: IHC + FISH confirmation of $IHC1^+$ or 2^+; $HER2^+$ defined as $IHC3^+$ or $FISH^+$ – S5: IHC + FISH confirmation of $IHC2^+$ and 3^+;$HER2^+$ defined as $FISH^+$ –S6: IHC + FISH confirmation of $IHC1^+$, 2^+ and 3^+; $HER2^+$ defined as	Strategies 2, 3 and 4 were ruled out by classic or extended dominance. The incremental costs per accurately determined HER2 status (ICERs) for the remaining nondominated strategies were: S5 over S1: $5,686; S6 over S5: $7,423; S7 over S6: $7,736.	– PSA performed by varying all parameters of the economic evaluation simultaneously over plausible ranges – The cost of FISH did effect CEA results	The strategy with the lowest ICER compared to current practice is to screen all BC patients with IHC plus confirmation of $IHC2^+$ and 3^+ by FISH (S5)	[4]

Final thoughts

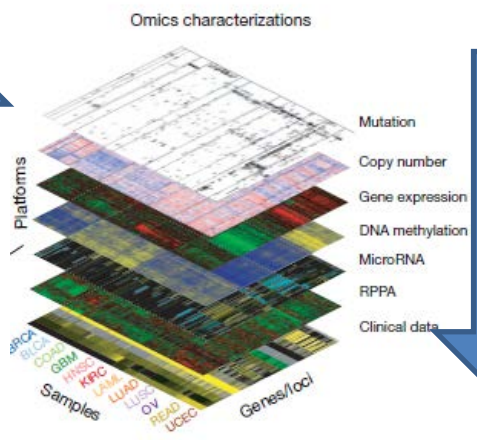
Clinical Development: Key Concepts

- Begin with the end/goal in mind: Unmet medical need – public health value
  - Begin with goal and design backwards
  - Data: Clinical, regulatory and health economic
  - Demonstrate clear, population specific benefit / risk
  - Efficient and timely as possible
- Understand the pathophysiology of the disease
  - Targets and/or molecules within a target
  - Prevalence within human populations
- Ensure that the agent/device/biomarker test is optimized for human testing and use before it is used in humans
  - Pharmacology, safety/toxicology, accuracy/reproducibility
- Plan for failure
  - Go/no go criteria to exit early if risk/benefit unacceptable
  - Intervention should be better than what is currently used

A new Taxonomy – “omics” as diagnostics?

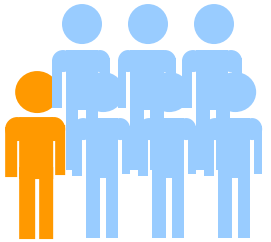


Practicality?



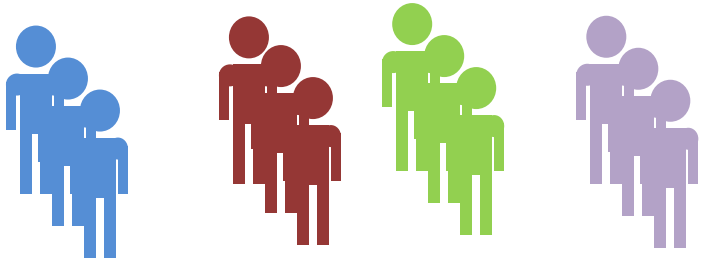
Functionality?

Less therapy



Current therapy

Alternate Therapies



*Well, **less is more**, Lucrezia: (Robert Browning 1855)*

- Highly multiplex assays are, generally, less robust and stable
  - Increased FDR and and technical errors.
- Whole genome sequencing costs are decreasing
  - the cost of bio-informatics however is not.
- Single gene predictors are rare
  - But we don't need to measure everything all the time

A final caveat

r^2 = 0.99 (p=0.00000001)