PRE-IMPAKT Training Course

How to Report Translational Research Results

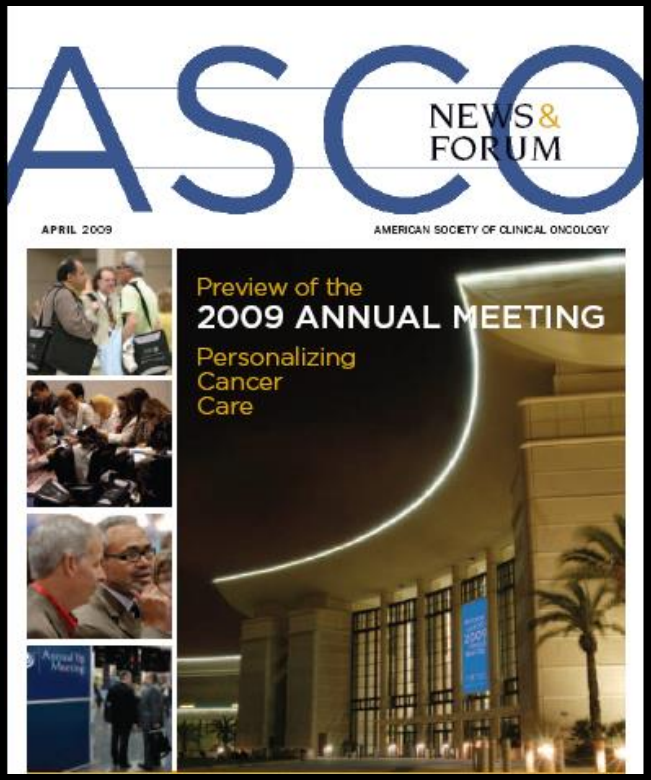
Daniel F. Hayes, M.D.

University of Michigan Comprehensive Cancer Center

Special Acknowledgement to my Colleague:

Lisa M. McShane, PhD

National Cancer Institute



The theme of the 2009 Annual Meeting, *chosen by 2008-2009 ASCO President Richard L. Schilsky, MD*, is:

"Personalizing Cancer Care."

"Each patient with cancer is different—biologically, clinically, economically, and socially—and a one-size-fits-all approach to treating cancer is not optimal," Dr. Schilsky said. "As oncologists, our focus has always been, and must remain, treating the patient, not the disease. We must each acquire the skills and make the commitment to do so in the optimal way."

Tumor Markers

- Tumor marker-based tests are integral to the practice of personalized cancer care
- Need to apply same rigor in development of marker tests as we do for treatments

A Typical Therapeutic Trial Methods Section

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Combination Anastrozole and Fulvestrant in Metastatic Breast Cancer

Rita S. Mehta, M.D., William E. Barlow, Ph.D., Kathy S. Albain, M.D., Ted A. Vandenberg, M.D., Shaker R. Dakhil, M.D., Nagendra R. Tirumali, M.D., Danika L. Lew, M.A., Daniel F. Hayes, M.D., Julie R. Gralow, M.D., Robert B. Livingston, M.D., and Gabriel N. Hortobagyi, M.D.

N Engl J Med 367:435-44, 2012

regulates the estrogen receptor by disrupting estrogen-receptor dimerization and accelerating degradation of the unstable fulvestrant–estrogen-receptor complex.[2] This effect leads to reduced cross-talk between the estrogen receptor and estrogen-independent growth factor signaling, thus delaying resistance to hormone therapy.[2] Clinically, fulvestrant at a dose of 250 mg monthly is as active as tamoxifen when used as first-line therapy for metastatic disease[3] and as active as anastrozole when administered in patients who have had disease progression after receiving tamoxifen therapy.[4,5]

In preclinical models, fulvestrant has been shown to have high efficacy in a low-estrogen environment.[6] The combination of fulvestrant and an aromatase inhibitor, as compared with either agent alone, delays the development of resistance by down-regulating several signaling molecules involved in the development of resistance.[7,8] We therefore conducted a phase 3, randomized trial to determine whether the combination of anastrozole and fulvestrant would be superior to anastrozole alone as first-line therapy for metastatic breast cancer.

METHODS

STUDY DESIGN AND OVERSIGHT

The study was designed and conducted, and the data were analyzed, by the Southwest Oncology Group (SWOG) Cooperative Group, which was funded by the National Cancer Institute (NCI), with review and collaboration by the other participating cooperative groups and the NCI Cancer Therapy Evaluation Program. The first two authors assume full responsibility for the quality and completeness of the data and vouch for the data analysis and for the fidelity of the study to the protocol. All drafts of the manuscript were prepared and approved by all the authors, and members of the SWOG made the decision to submit it for publication. The trial data were reviewed by a data and safety monitoring committee every 6 months.

text of this article at NEJM.org.

ELIGIBILITY

Eligible patients were postmenopausal women w HR-positive metastatic breast cancer (estroge receptor–positive, progesterone-receptor–positi or both), diagnosed according to local institutio standards. Women were eligible if they had h no prior chemotherapy, hormonal therapy, or i munotherapy for metastatic disease. Neoad vant or adjuvant chemotherapy had to have be completed more than 12 months before enr ment. In the original protocol, women who h received prior adjuvant therapy with an aromata inhibitor or fulvestrant were excluded, but th who had received prior adjuvant tamoxifen th apy were eligible. In an early amendment, wom who had received prior adjuvant therapy with aromatase inhibitor were also eligible if the th apy had been completed more than 12 mont before enrollment. Patients were not allowed receive concurrent chemotherapy or other h monal therapy during the study treatment peri (bisphosphonates were allowed). Women wi either measurable or nonmeasurable disease w eligible. Other major eligibility criteria includ no known metastases in the central nervous s tem and a Zubrod's performance score of 0 t (with a score of 0 indicating that the patient fully active, 1 that the patient is restricted strenuous activity but is ambulatory, and 2 t the patient is unable to work but is ambulat and capable of self-care).[9] Patients with ble ing diathesis or long-term anticoagulant ther (except antiplatelet therapy) were ineligible. tients with other cancers were ineligible unl the cancer had been adequately treated or h been in remission for at least 5 years. All p tients provided written informed consent bef enrollment.

RANDOMIZATION AND TREATMENT

Randomization was performed at a central lo tion, with stratification according to prior rece or no prior receipt of adjuvant tamoxifen therapy. Patients were randomly assigned, in a 1:1 ratio, to anastrozole alone (group 1) or to fulvestrant in combination with anastrozole (group 2). Patients in group 1 received 1 mg of anastrozole orally each day. Patients in group 2 received 1 mg of anastrozole orally each day, as well as an initial loading dose (500 mg) of fulvestrant administered intramuscularly on day 1, followed by 250 mg (low-dose fulvestrant) administered intramuscularly on day 14 and day 28 of the first cycle, and thereafter every 28 days. Treatment was continued until disease progression, the development of unacceptable toxic effects, a delay in treatment of 4 weeks or longer, or withdrawal of the patient from the trial. After progression, the treating physician could choose the appropriate therapy, although crossover to low-dose fulvestrant was strongly recommended for patients in group 1 after discontinuation of anastrozole, and fulvestrant was provided free of charge to encourage crossover to that agent. After a higher monthly dose of fulvestrant (500 mg) was shown to be superior to the low dose[10] and the Food and Drug Administration approved the higher monthly dose, the protocol was amended (on February 2, 2011) to allow patients in either group to receive the 500-mg dose after progression.

ASSESSMENT OF PROGRESSION AND SURVIVAL

Progression was assessed every 3 months and was defined according to the Response Evaluation Criteria in Solid Tumors (RECIST) in the case of measurable disease[11] and according to an assessment of the worsening of symptoms or increasing disease (as determined by the patient's oncologist) in the case of nonmeasurable disease. After progression, overall survival was assessed every 6 months for the first 2 years from the time of random assignment and then annually for the next 2 years. Follow-up beyond 4 years was not required, although 32 patients without progression at 48 months continued to be followed.

ASSESSMENT OF TOXIC EFFECTS

Toxic effects were measured according to the Common Terminology Criteria for Adverse Events, version 3.0 (http://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/ctcaev3.pdf). Patients with grade 3 or grade 4 toxic effects could have treatment interrupted for up to 4 weeks to allow resolution of the toxic effects to grade 2 or less. The study treatment was withdrawn in the case of grade 3 or 4 toxic effects that did not resolve by 4 weeks.

STATISTICAL ANALYSIS

The primary outcome was progression-free survival, which was defined as the time from random assignment to disease progression or death from any cause. Data from patients who were alive and progression-free at the time of cutoff of the data (September 29, 2011) were censored at the last follow-up visit at which progression had not yet been observed. Overall survival, which was a secondary outcome, was defined as the time from random assignment to death from any cause. We calculated the rates of clinical benefit using the number of patients with a complete or partial response or stable disease as the numerator and the number of all patients (even those in whom a response could not be assessed or for whom response data were missing) as the denominator. The rate of objective response was calculated only for patients with measurable disease, whereas the rate of clinical benefit applied to all patients. Both the primary analysis of progression-free survival and the analysis of overall survival were specified as log-rank tests stratified according to prior receipt or no prior receipt of adjuvant tamoxifen therapy. Kaplan–Meier methods were used to construct survival plots and to estimate the survival percentages and the median times to progression-free and overall survival. Cox regression was used to estimate hazard ratios and 95% confidence intervals.

Post hoc subgroup analyses were performed on the basis of the stratification variable (status with respect to prior adjuvant tamoxifen therapy); therefore, the results should be interpreted cautiously. A forest plot was used to compare the overall hazard ratio with the hazard ratios obtained in subgroups defined on the basis of several potentially prognostic or predictive factors. P values for interaction were obtained from a Cox regression analysis. Two interim analyses of progression-free survival were performed when 50% and 75% of the expected events had occurred. The final analysis was set at an alpha level of 0.02 (one-sided) so that the one-sided cumulative alpha level was 0.025 or the two-sided alpha level was 0.05. All tests were two-sided, so the P value for the final analysis of the primary outcome (progression-free survival) had to be 0.04 or less to indicate statistical significance. We estimated that with a sample of 690 patients and an expected median progression-free survival of 10 months in group 1 and 13 months in group 2, the trial would have 90% overall power to show a between-group difference in the primary outcome. The projected medians for overall survival were 36 months and 48 months, respectively.

A Typical Tumor Biomarker Methods Section

This article has been retracted: N Engl J Med 2011. DOI: 10.1056/NEJMc1101915.

ORIGINAL ARTICLE

How were these patients treated? Does treatment affect results?

What was definition of these endpoints? Who determined them?

Why n= 89, 25, and 84 from three groups?

Could you reproduce these data from this section?

LUNG CANCER IS THE LEADING CAUSE OF death from cancer among both men and women in the United States, and non–small-cell lung cancer (NSCLC) accounts for almost 80 percent of such deaths.[1,2] The clinical staging system has been the standard for determining lung-cancer prognosis.[3-5] Although other clinical and biochemical markers have prognostic significance,[6,7] none are more accurate than the clinicopathological stage.[8]

The current standard of treatment for patients with stage I NSCLC is surgical resection, despite the observation that nearly 30 to 35 percent will relapse after the initial surgery and thus have a poor prognosis,[2,4] indicating that a subgroup of these patients might benefit from adjuvant chemotherapy. Similarly, as a population, patients

refine the clinical prognosis and the context in which improved prognostic capability could be used to alter a clinical treatment decision were not clear. Thus, we evaluated the use of gene-expression patterns as a means of stratifying risk and treatment in NSCLC.

METHODS

PATIENTS AND TUMOR SAMPLES

We analyzed 439 tumor samples from three cohorts of patients with NSCLC. The training cohort consisted of 89 patients enrolled through the Duke Lung Cancer Prognostic Laboratory. The independent validation cohorts included patients in two multicenter cooperative group trials: 25 patients from the American College of Surgeons Oncology Group (ACOSOG) Z0030 study and 84 from the prospective Cancer and Leukemia Group B (CALGB) 9761 trial. Table 1 lists the clinical and demographic characteristics of the patients in each cohort and their tumors, and complete details are listed in Table 1 of the Supplementary Appendix, available with the full text of this article at www.nejm.org. All patients were enrolled according to protocols approved by the institutional review board of Duke University, after written informed consent had been obtained.

HISTOPATHOLOGICAL EVALUATION

For each cohort, a single pathologist reviewed all slides to determine whether they met the histopathological criteria for NSCLC of the World Health Organization, including the subtype of adenocarcinoma and the degrees of differentiation, lymphatic invasion, and vascular invasion. Only samples with a tumor-cell content of [illegible] percent were used in the analysis.

EXPRESSION ARRAYS

RNA was extracted from the [illegible] Neasy Kits (Qiagen). The RNA was [illegible] with the use of a bioanalyzer (Agilent). Hybridization targets were prepared from [illegible] RNA according to standard protocols (described in detail in the Supplementary Appendix, along with the [illegible] scanning of the arrays and [illegible] the resulting data). The microarrays were carried out with Affymetrix GeneChip Human Genome U133 Plus2). All raw data and data transformed with the use of the robust multiarray average expression measure for the Duke, ACOSOG, and CALGB data sets are available elsewhere (accession number GSE3593 in the Gene Expression Omnibus database at www.ncbi.nlm.nih.gov/geo).

STATISTICAL ANALYSIS

We performed statistical analyses using the metagene construction and binary prediction tree analysis, as described previously[25-29] and in detail in the Supplementary Appendix. The metagene for a cluster of genes is the dominant singular factor (principal component), as computed with the use of a singular value decomposition of gene-expression levels in the gene cluster in all samples. The metagene represents the dominant average pattern of expression of the gene cluster across the tumor samples.[25]

We then used the set of metagenes and the clinical variables previously shown to be of prognostic value (age, sex, tumor diameter, stage of

initial diagnosis of NSCLC were made in terms of the estimated risk probabilities.[26,30,31] In the analysis, many classification trees were computed, weighed, and integrated to provide overall risk predictions for each patient. The dominant metagenes that constituted the final model are described in the Supplementary Appendix.

To compare the prognostic efficacy of the metagenes and clinical strategies, the clinical variables [illegible] factors or principal components [illegible] treatment of metagenes in the lung [illegible]) in a classification-tree analysis [illegible] clinical model. The end result was [illegible] of recurrence, which represents the [illegible] prognostic value of the individual [illegible]s. Using GraphPad software, we [illegible] statistic (comparable to the area [illegible] in a receiver-operating-characteristic [illegible] prediction of binary outcomes) [illegible] that included just the clinical variables [illegible] for a model that included just [illegible] and a C statistic for a model that [illegible] clinical and genomic variables.

The accuracy of each model was defined with the use of a probability of 0.5 as a cutoff. An estimated probability of recurrence of more than 0.5 was classified as a high risk of recurrence; an estimated probability of recurrence of 0.5 or less was classified as a low risk of recurrence.

Simple univariate and multivariate logistic regressions for recurrence (with and without the metagene-based assessment of the risk) were also computed to assess the baseline prognostic value of each clinical variable (age, sex, tumor diameter, stage of disease, histologic subtype, and smoking history) in the cohorts. We also calculated the sensitivity, specificity, and positive and negative predictive values using a probability of recurrence of 0.5 as the cutoff value. Standard Kaplan–Meier survival curves were generated for the high-risk and low-risk groups of patients with the use of GraphPad software; the survival curves were compared with the use of the log-rank test. This test generates a two-tailed P value that tests the null hypothesis, which was that the survival curves were identical among the cohorts.

clinical characteristics of the patients in each cohort and their tumors) used to develop and test the prognostic model (Fig. 1).

USE OF GENE-EXPRESSION PROFILES TO IMPROVE PROGNOSIS

Lung cancer is a heterogeneous disease resulting from the acquisition of multiple somatic mutations; given this complexity, it would be surprising if a single gene-expression pattern could effectively describe and ultimately predict the clinical course of the disease for all patients. Recognizing the importance of addressing this complexity, we have previously described methods to integrate various forms of data, including clinical variables and multiple gene-expression profiles, to build robust predictive models for the individual patient.[25,26] There are two critical components of this methodologic approach. First, we generated a collection of gene-expression profiles, termed "metagenes" (an example is given in Fig. 2A), that provide the basis for the predictive models. Second, we used clas-

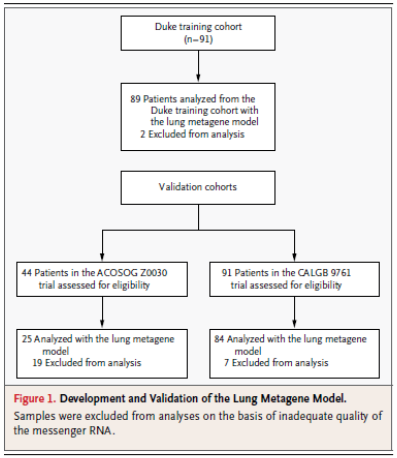


Figure 1. Development and Validation of the Lung Metagene Model.
Samples were excluded from analyses on the basis of inadequate quality of the messenger RNA.

Tumor Markers

- **A bad tumor marker is as harmful as a bad drug!**
- **Would you use a drug if:**
  - **You aren't sure how it is mixed?**
  - **You aren't sure what the concentration is?**
  - **You don't have clinical data about how the drug might be useful?**
  - **You don't have reliable clinical research data to determine how much efficacy it might have?**

Efforts to Facilitate Better Interpretation of Tumor Marker Literature

COMMENTARY

Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK)

Lisa M. McShane, Douglas G. Altman, Willi Sauerbrei, Sheila E. Taube, Massimo Gion, Gary M. Clark for the Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics

McShane, et al., J Natl Cancer Inst 97:1180-4, 2005
McShane, et al., J Clin Oncol 23:9067-72, 2005
McShane, et al., Br J Cancer 93:387-91, 2005
McShane, et al., Breast Cancer Res Treat 100:229-35, 2006

OPEN ACCESS Freely available online PLoS MEDICINE

Guidelines and Guidance

Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): Explanation and Elaboration

Douglas G. Altman[1]*, Lisa M. McShane[2], Willi Sauerbrei[3], Sheila E. Taube[4]

Altman, et al., BMC Medicine 10:2012
Altman, et al., PLoS Med 9:e1001216, 2012

Biospecimen Reporting for Improved Study Quality (BRISQ)

Helen M. Moore, PhD[1]; Andrea B. Kelly, PhD[2]; Scott D. Jewell, PhD[3]; Lisa M. McShane, PhD[4]; Douglas P. Clark, MD[5]; Renata Greenspan, MD[6]; Daniel F. Hayes, MD[7]; Pierre Hainaut, PhD, MS[8]; Paula Kim[9]; Elizabeth A. Mansfield, PhD[10]; Olga Potapova, PhD[11]; Peter Riegman, PhD[12]; Yaffa Rubinstein, PhD[13]; Edward Seijo, MS[14]; Stella Somiari, PhD[15]; Peter Watson, MB, BChir[16]; Heinz-Ulrich Weier, PhD[17]; Claire Zhu, PhD[18]; and Jim Vaught, PhD[1]

Moore, et al., Biopreserv Biobank 9:57-70, 2011
Moore, et al., Clin Chim Acta 413:1305, 2012
Moore, et al., J Proteome Res 10:3429-38, 2011
Moore, et al., Cancer Cytopathol 119:92-101, 2011

Efforts to Facilitate Better Interpretation of Tumor Marker Literature

VOLUME 30 · NUMBER 34 · DECEMBER 1 2012

JOURNAL OF CLINICAL ONCOLOGY

REVIEW ARTICLE

Publication of Tumor Marker Research Results: The Necessity for Complete and Transparent Reporting

Lisa M. McShane and Daniel F. Hayes

Journal of Clinical Oncology 30:4223-32, 2012

Tumor Biomarker Publications and Use: Definitions

- **Analytical Validity**
  - **Does the assay accurately and reproducibly measure what you say?**
- **Clinical (or "Biologic") Validity**
  - **Does the assay actually identify a biologic difference ("pos" vs. "neg") that may or may not be clinically useful?**
- **Clinical Utility**
  - **Do results of the assay lead to a clinical decision that has been shown with high level of evidence to improve outcomes?**

Teutsch S.M., et al. Genet Med. 11:3-14, 2009

Definitions

- **Analytical Validity**
  - **Does the assay accurately and reproducibly measure what you say?**
- **Clinical (or "Biologic") Validity**
  - **Does the assay actually identify a biologic difference ("pos" vs. "neg") that may or may not be clinically useful?**
- **Clinical Utility**
  - **Do results of the assay lead to a clinical decision that has been shown with high level of evidence to improve outcomes?**

Teutsch S.M., et al. Genet Med. 11:3-14, 2009

Analytical Validity

- **Pre-analytical Validity: BRISQ Criteria**

- **Where is Specimen From?**
- **How was Specimen:**
  - **Collected**
  - **Processed**
  - **Stored**
  - **Treated**

	Data Elements	Examples
❐	**Biospecimen type** *Solid tissue, whole blood, or another product derived from a human being*	Serum, Urine
❐	**Anatomical site** *Organ of origin or site of blood draw*	Liver, Antecubital area of the arm
❐	**Disease status of patients** *Controls or individuals with the disease of interest*	Diabetic, Healthy control
❐	**Clinical characteristics of patients** *Available medical information known or believed to be pertinent to the condition of the biospecimens*	Pre-menopausal breast cancer patients
❐	**Vital State of patients** *Alive or deceased patient when biospecimens were obtained*	Postmortem
❐	**Clinical diagnosis of patients** *Patient clinical diagnoses (determined by medical history, physical examination, and analyses of the biospecimen) pertinent to the study*	Breast cancer
❐	**Pathology diagnosis** *Patient pathology diagnoses (determined by macro and/or microscopic evaluation of the biospecimen at the time of diagnosis and/or prior to research use) pertinent to the study*	Her2-negative intraductal carcinoma
❐	**Collection mechanism** *How the biospecimens were obtained*	Fine needle aspiration, Pre-operative blood draw
❐	**Type of stabilization** *The initial process by which biospecimens were stabilized during collection*	Heparin, On ice
❐	**Type of long-term preservation** *The process by which the biospecimens were sustained after collection*	Formalin fixation, freezing
❐	**Constitution of preservative** *The make-up of any formulation used to maintain the biospecimens in a non-reactive state*	10% neutral-buffered formalin, 10 USP Heparin Units/mL
❐	**Storage temperature** *The temperature or range thereof at which the biospecimens were kept until distribution/analysis.*	-80 °C, 20 to 25 °C
❐	**Storage duration** *The time or range thereof between biospecimen acquisition and distribution or analysis.*	8 days, 5 to 7 years
❐	**Shipping temperature** *The temperature or range thereof at which biospecimens were kept during shipment or relocation.*	-170 °C to -190 °C
❐	**Composition assessment & selection** *Parameters used to choose biospecimens for the study*	Minimum 80% tumour nuclei & maximum 50% necrosis

Analytical Validity

- **Technical and Biological Issues**
  - **How is the assay performed?**
  - **What type of specimen is required?**
  - **How accurately is the analyte measured?**
  - **Are measurements reproducible (within lab, between labs, between operators, between different portions of the specimen)?**
  - **Do different assay methods yield similar biomarker values?**

REMARK

Assay Methods (#5 on Checklist):
Specify the assay method used and provide (or reference) a detailed protocol, including specific reagents or kits used, quality control procedures, reproducibility assessments, quantitation methods, and scoring and reporting protocols. Specify whether and how assays were performed blinded to the study endpoint

Definitions

- Analytical Validity
  - Does the assay accurately and reproducibly measure what you say?
- **Clinical (or "Biologic") Validity**
  - **Does the assay actually identify a biologic difference ("pos" vs. "neg") that may or may not be clinically useful?**
- Clinical Utility
  - Do results of the assay lead to a clinical decision that has been shown with high level of evidence to improve outcomes?

Teutsch S.M., et al. Genet Med. 11:3-14, 2009

Clinical Validity

- **Is there an association between the biomarker and a clinical endpoint?**
  - **In what patient population?**
  - **In what clinical setting?**
  - **What clinical endpoint?**

REMARK

MATERIALS AND METHODS	
Patients	
2	Describe the characteristics (for example, disease stage or co-morbidities) of the study patients, including their source and inclusion and exclusion criteria.
3	Describe treatments received and how chosen (for example, randomized or rule-based).
Study design	
6	State the method of case selection, including whether prospective or retrospective and whether stratification or matching (for example, by stage of disease or age) was used. Specify the time period from which cases were taken, the end of the follow-up period, and the median follow-up time.
7	Precisely define all clinical endpoints examined.

Clinical Validity

- **Is there an association between the biomarker and a clinical endpoint?**
  - **Nature of the association/ magnitude of effect?**
    - **Form of marker**
      - **With continuous biomarker?**
      - **With dichotomized biomarker?**
    - **Prognostic vs. predictive**
  - **Does marker add information beyond standard variables?**

REMARK

8	List all candidate variables initially examined or considered for inclusion in models.
9	Give rationale for sample size; if the study was designed to detect a specified effect size, give the target power and effect size.
Statistical analysis methods	
10	Specify all statistical methods, including details of any variable selection procedures and other model-building issues, how model assumptions were verified, and how missing data were handled.

Clinical Validity

- **Was association established in statistically appropriate way?**
  - **Positive vs. Negative Cutoff determination**
    - **Arbitrary**
      - **0 vs. any, or >10% pos, or Mean, Median**
      - **Mean + 2SD of normal (often done with circulating markers)**
      - **Mean of normal + sufficient to be above coefficient of variation of assay**
    - **Data Driven**
      - **Cutpoint "optimization" to produce lowest p-value may create spurious associations**

REMARK

11	Clarify how marker values were handled in the analyses; if relevant, describe methods used for cutpoint determination.

Definitions

- Analytical Validity
  - Does the assay accurately and reproducibly measure what you say?
- Clinical (or "Biologic") Validity
  - Does the assay actually identify a biologic difference ("pos" vs. "neg") that may or may not be clinically useful?
- **Clinical Utility**
  - **Do results of the assay lead to a clinical decision that has been shown with high level of evidence to improve outcomes?**

Teutsch S.M., et al. Genet Med. 11:3-14, 2009

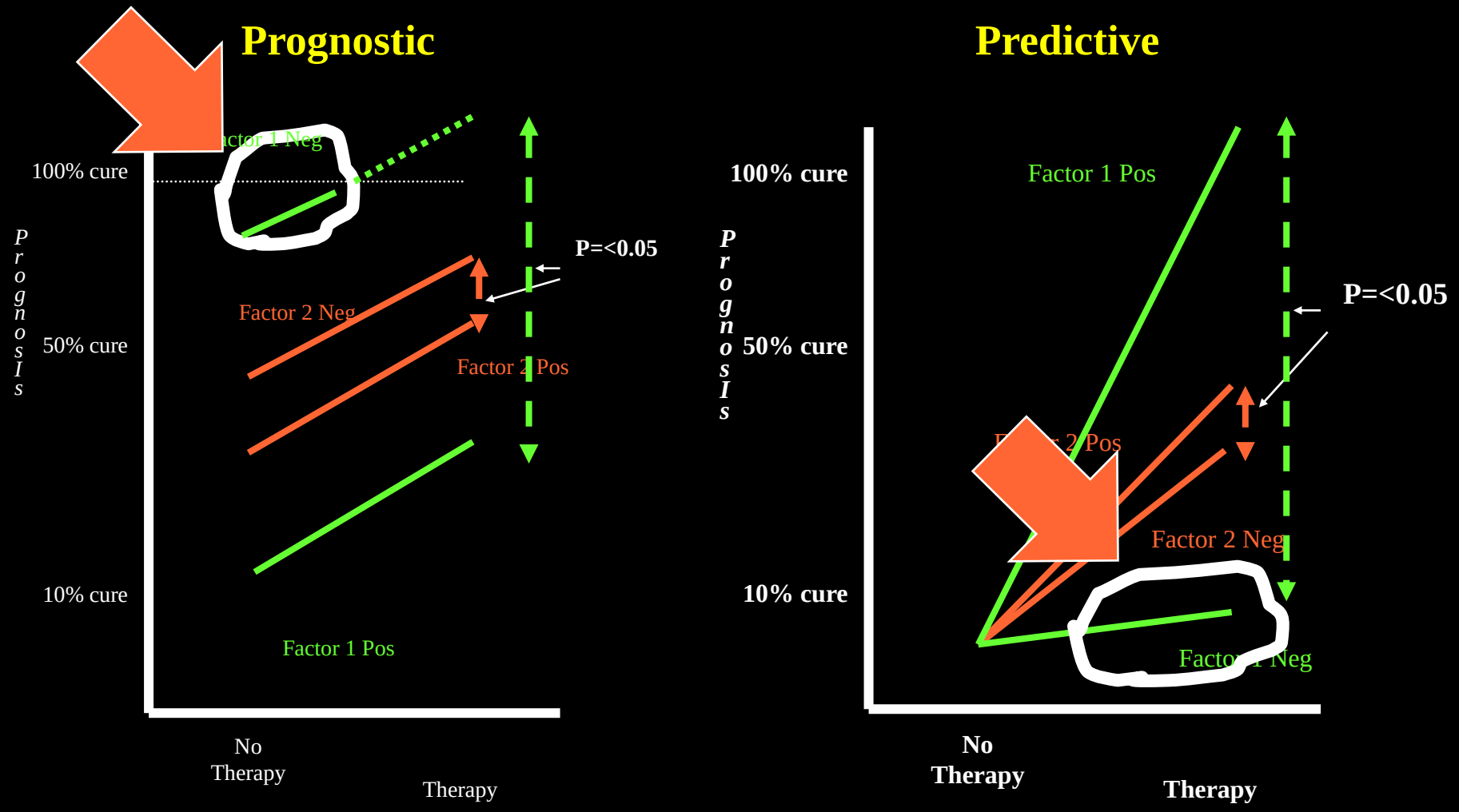
When is a Marker Clinically Useful?

- **It is either prognostic or predictive**
- **The magnitude of effect is sufficiently large that clinical decisions based on the data result in outcomes that are acceptable**
  - ***Greater chance for benefit***
  - ***Smaller toxicity risk***
- **The estimate of magnitude of effect is reliable**
  - ***Assay is reproducible***
  - ***Clinical trial/marker study design is appropriate***
  - ***Results are validated in subsequent well-designed studies***

Henry N.L., Hayes D.F. Oncologist. 11:541-52, 2006

Prognostic and Predictive Factors:

What Are We Trying to Find?



When is a Marker Clinically Useful?

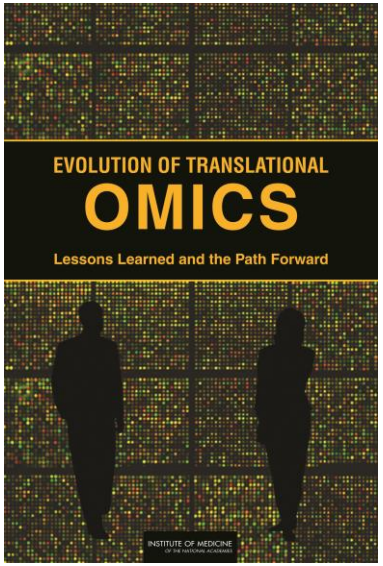
- **It is either prognostic or predictive**
- **The magnitude of effect is sufficiently large that clinical decisions based on the data result in outcomes that are acceptable**
  - ***Greater chance for benefit***
  - ***Smaller toxicity risk***
- **The estimate of magnitude of effect is reliable**
  - ***Assay is reproducible***
  - ***Clinical trial/marker study design is appropriate***
  - ***Results are validated in subsequent well-designed studies***

Henry N.L., Hayes D.F. Oncologist. 11:541-52, 2006
Simon R., Paik S., & Hayes DF., J Natl Cancer Inst 101:1446-52, 2009

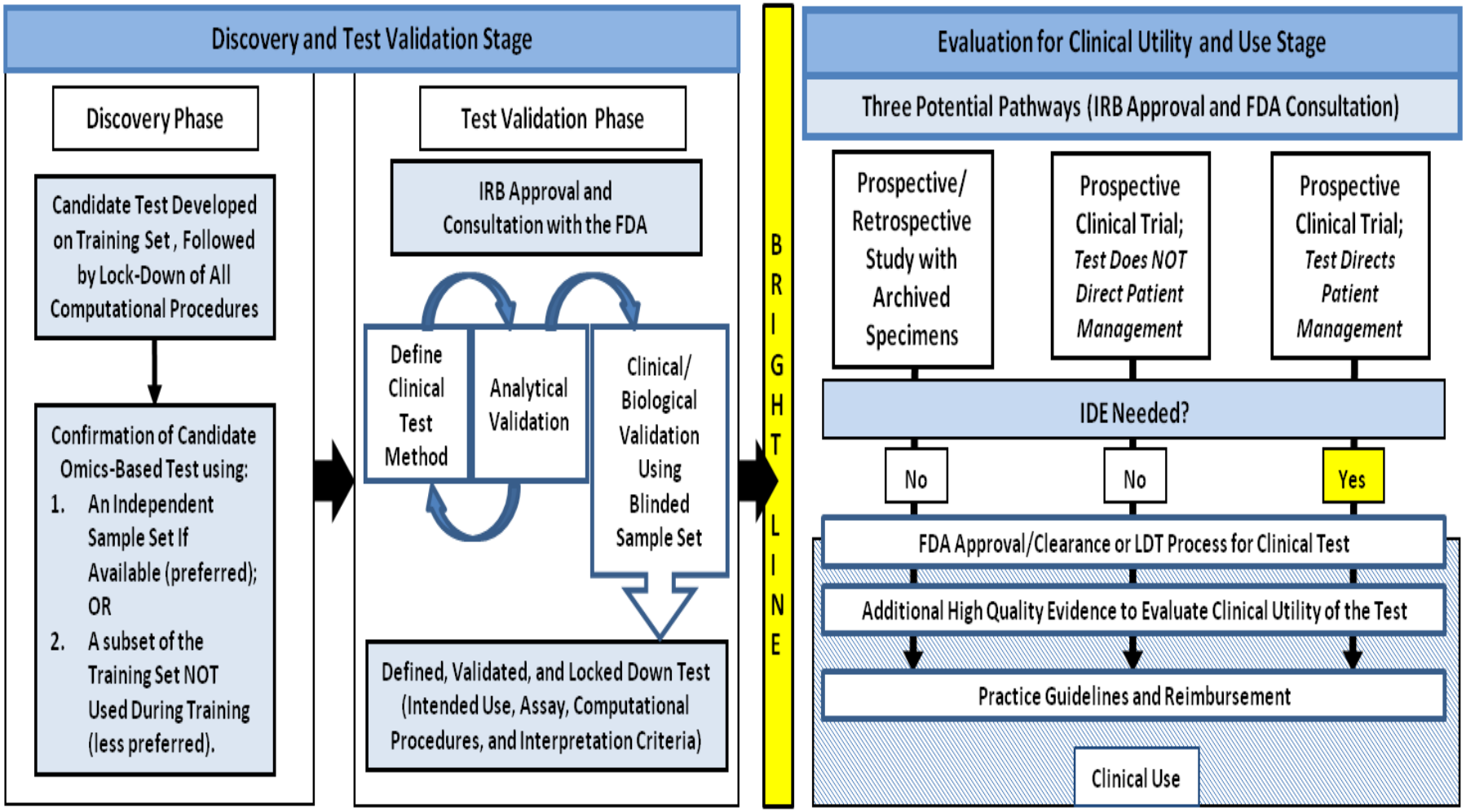
Overview

Gilbert S. Omenn, MD, PhD

University of Michigan



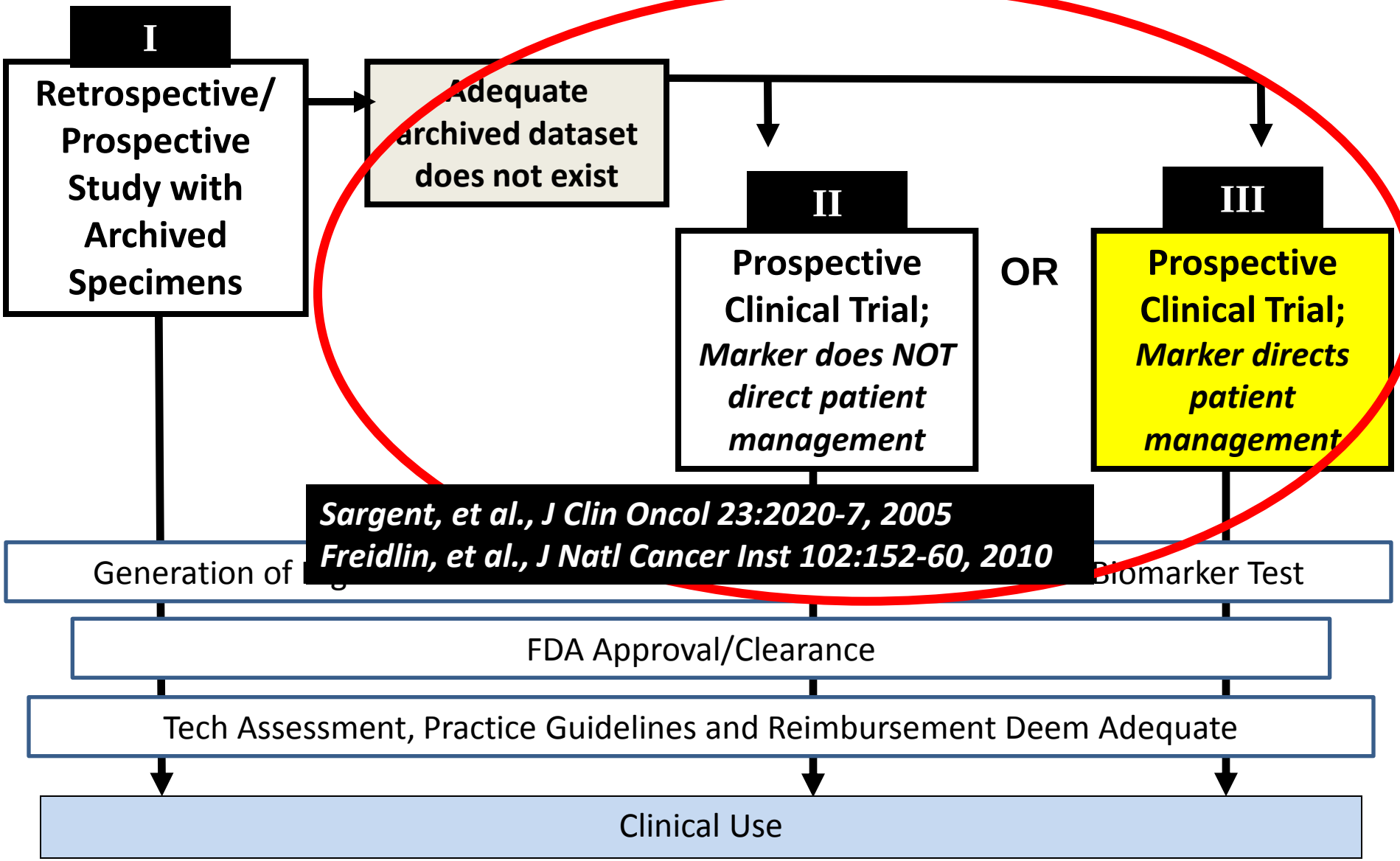
Evaluation for Clinical Utility and Use



Tumor Markers: Carrots and Sticks

- **Clinical Research: Various Strategies to *"Test the Test"***
  - ***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
  - ***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*

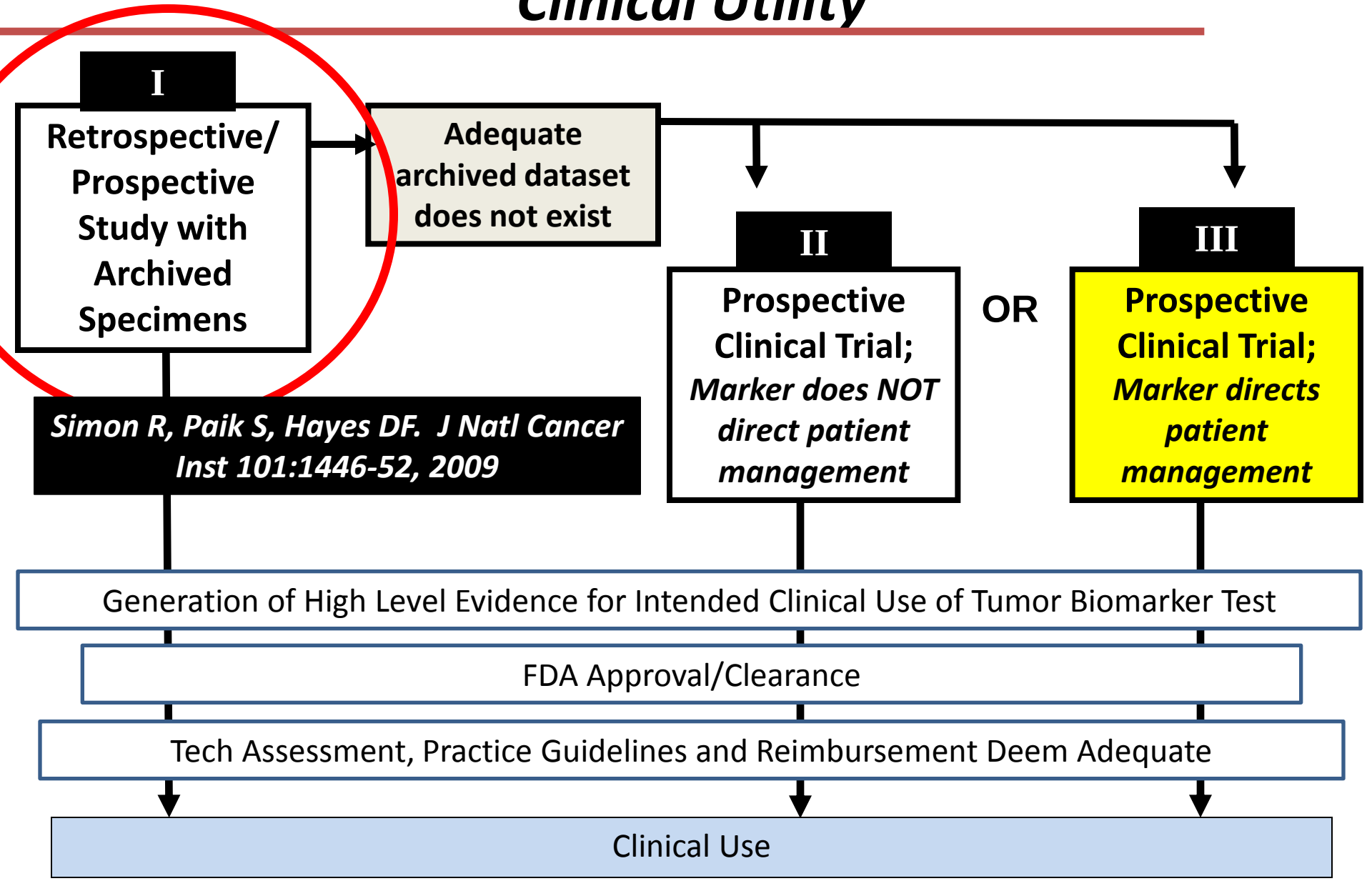
Recommended Pathways to Generate LOE I Data for Clinical Utility



Tumor Markers: Carrots and Sticks

- **Clinical Research: Various Strategies to *"Test the Test"***
  - ***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
  - ***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***

Recommended Pathways to Generate LOE I Data for Clinical Utility



Clinical Design Issues

REMARK

Table 1. The REMARK checklist [1–7].

INTRODUCTION	
1	State the marker examined, the study objectives, and any pre-specified hypotheses.
MATERIALS AND METHODS	
Patients	
2	Describe the characteristics (for example, disease stage or co-morbidities) of the study patients, including their source and inclusion and exclusion criteria.
3	Describe treatments received and how chosen (for example, randomized or rule-based).
6	State the method of case selection, including whether prospective or retrospective and whether stratification or matching (for example, by stage of disease or age) was used. Specify the time period from which cases were taken, the end of the follow-up period, and the median follow-up time.
7	Precisely define all clinical endpoints examined.
8	List all candidate variables initially examined or considered for inclusion in models.
9	Give rationale for sample size; if the study was designed to detect a specified effect size, give the target power and effect size.
Statistical analysis methods	
10	Specify all statistical methods, including details of any variable selection procedures and other model-building issues, how model assumptions were verified, and how missing data were handled.
11	Clarify how marker values were handled in the analyses; if relevant, describe methods used for cutpoint determination.

Clinical Analytical Issues

REMARK

RESULTS	
Data	
12	Describe the flow of patients through the study, including the number of patients included in each stage of the analysis (a diagram may be helpful) and reasons for dropout. Specifically, both overall and for each subgroup extensively examined report the number of patients and the number of events.
13	Report distributions of basic demographic characteristics (at least age and sex), standard (disease-specific) prognostic variables, and tumor marker, including numbers of missing values.
Analysis and presentation	
14	Show the relation of the marker to standard prognostic variables.
15	Present univariable analyses showing the relation between the marker and outcome, with the estimated effect (for example, hazard ratio and survival probability). Preferably provide similar analyses for all other variables being analyzed. For the effect of a tumor marker on a time-to-event outcome, a Kaplan-Meier plot is recommended.
16	For key multivariable analyses, report estimated effects (for example, hazard ratio) with confidence intervals for the marker and, at least for the final model, all other variables in the model.
17	Among reported results, provide estimated effects with confidence intervals from an analysis in which the marker and standard prognostic variables are included, regardless of their statistical significance.
18	If done, report results of further investigations, such as checking assumptions, sensitivity analyses, and internal validation.

Conclusions/Discussion

REMARK

DISCUSSION	
19	Interpret the results in the context of the pre-specified hypotheses and other relevant studies; include a discussion of limitations of the study.
20	Discuss implications for future research and clinical value.

Prospective Registry of Tumor Biomarker Studies

FOCUS ON PERSONALIZED MEDICINE

For Prospective Retrospective Studies:

Biomarker studies: a call for a comprehensive biomarker study registry

Fabrice Andre, Lisa M. McShane, Stefan Michiels, David F. Ransohoff, Douglas G. Altman, Jorge S. Reis-Filho, Daniel F. Hayes and Lajos Pusztai

Andre, et al., Nat Rev Clin Oncol 8:171-6, 2011

http://win.biomarkerregistry.org

Prospective: Preferably registered in ClinicalTrials.gov

Circulating Tumor Biomarker Studies

Clinical Chemistry 59:1
000–000 (2013)

Special Report

Design of Tumor Biomarker–Monitoring Trials: A Proposal by the European Group on Tumor Markers

György Sölétormos,[1] Michael J. Duffy,[2] Daniel F. Hayes,[3] Catharine M. Sturgeon,[4*] Vivian Barak,[5] Patrick M. Bossuyt,[6] Eleftherios P. Diamandis,[7,8] Massimo Gion,[9] Per Hyltoft-Petersen,[10] Rolf M. Lamerz,[11] Dorte L. Nielsen,[12] Paul Sibley,[13] Bengt Tholander,[14] Malgorzata K. Tuxen,[12] and Johannes M.G. Bonfrer[15]

"MONITOR Guidelines"
Soletormos, et al., Clin Chem 59:52-9, 2013

- **Suggested Trial Designs**
- **BRISQ and REMARK still pertain**

When is a Marker Clinically Useful?

- ***REGARDLESS OF THE PATHWAY YOU CHOOSE TO GET TO THE ANSWER, YOU NEED TO USE THE SCIENTIFIC METHOD TO WALK DOWN IT!***

- ***LUCK IS NOT A GOOD STRATEGY IN GOLF OR SCIENCE……..***

Tumor Markers: Carrots and Sticks

- **Research**
  - **Funding:** NCI Cancer Biomarkers Study Section

    www.cms.csr.nih.gov
  - **Tumor Marker Study Registry** ***(=clinicaltrials.gov)*****:**

(Andre, F., et al.; Nat Rev Clin Oncol; 2011)

- **Publication:** Recommended Guidelines
  - **BRISQ:** Moore HM, Kelly AB, Jewell SD, et al. Biospecimen Reporting for Improved Study Quality (BRISQ). J Proteome Res 2011.
  - **REMARK:** Mcshane et al, REporting Recommendations for Tumor MARker Prognostic Studies (REMARK) J Clin Oncol, 2005
  - **MONITOR:** Soletormis et al. Design of Tumor Biomarker–Monitoring Trials: A Proposal by the European Group on Tumor Markers. Clin Chem 2013

Thanks to Many Colleagues

- Co-authors of Vicious Cycle
- ASCO TM Guidelines Committee
- Steve Gutman; Formerly FDA,
- Richard Simon; NCI
- Lisa McShane; NCI