EDRN Approaches to Biomarker Validation

DMCC Statisticians
Fred Hutchinson Cancer Research Center

Margaret Pepe

Ziding Feng, Mark Thornquist, Yingye Zheng, and Ross Prentice

Biomarker for What Purpose?

- ✓ Early Detection Screening
- ✓ Diagnosis
- ✓ Prognosis
- ✓ Risk Prediction
- Treatment Selection
- Surrogate Outcome
- Exposure

Biomarker with What Performance?

- Measure of performance is context-dependent
- Acceptable levels of performance also context-dependent

Breast Cancer Collaborative Group Study

Context = diagnostic biopsy for suspicious lesion

Biomarker purpose = diagnostic

Decision rule = “biopsy only if biomarker positive”

Performance measure = reduction in unnecessary biopsies
reduction in cancers detected

Acceptable levels of performance: True Positive Rate $\geq$ 98%
False Positive Rate $\leq$ 75%

No formal decision analysis to set these criteria

Another Purpose

Reduce unnecessary mammograms by applying biomarker ***before*** mammogram

FPF $\leq$ 75% would have enormous impact

Ovarian Cancer Screening of Healthy Population

- Performance = disease detected early
  false referral for work-up
- FPF $\leq$ 2%
- TPF for late stage cancer at 1 year before clinical diagnosis > 20%

Risk Prediction

Model individual's risk of bad outcome given his marker value(s)

- Well calibrated model?
- Performance = Useful delineation of risk distribution across the population?

Risk Model of Biopsy Proven High Grade Prostate Cancer in the PCPT Study

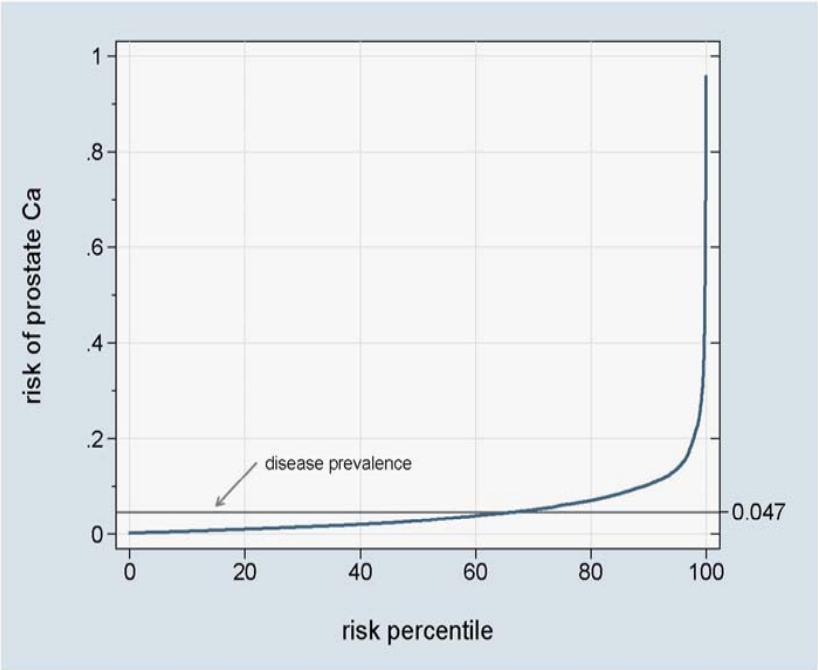
Placebo Arm (n=5519)

Factor	Log Odds Ratio	P-value
Constant	-5.94	--
Log (PSA)	1.30	<0.001
Age (years)	0.03	0.020
DRE	0.99	<0.001
Prior biopsy	-1.37	0.040

Risk Model of Biopsy Proven High Grade Prostate Cancer in the PCPT Study

Placebo Arm (n=5519)

Figure 1. Predictiveness curve for the combination of PSA and risk factors.



$R^{-1}(.02)$	=	Risk below 0.02	=	39.0%
$1 - R^{-1}(.20)$	=	Risk above 0.20	=	2.2%

Phases of Biomarker Development for Early Detection

Phase 1 Discovery
Phase 2 Diagnostic Validation
Phase 3 Early Detection Validation*
Phase 4 Prospective Application
Phase 5 Randomized Trial with Treatment

EDRN focus on phase 2 and 3 studies

Phases 4 and 5 involve actions based on biomarker result. Consequences to patients should be evaluated.

Key Design Issues in Definitive Validation

- Study population is that for intended clinical application. Sufficiently general? Multiple institutions?
- Marker well defined in advance
  - Validation separate from discovery
  - Combination pre-defined
  - Threshold need not be predefined?
- Assay as intended for clinical use?
- Minimally acceptable performance criteria to be met. Justification?
  Anticipated/desirable performance drives sample size calculations

Key Design Issues in Definitive Validation

- Cases-controls from the same population typically require prospective collection of samples for storage. Existing repositories (e.g., PLCO, WHI) or create our own
- Blinding: collection, storage and assay
- Random selection of eligible cases from repository. Stratify on disease characteristics and other factors?
- Random selection of controls
  - Several control groups possible
  - Matched to cases if appropriate (potential pitfalls here)
- Early termination rules. Adjustments in analysis?

Standards for Design of the Definitive* Study

Therapeutics: the randomized placebo controlled clinical trial

Biomarkers: the prospective collection blinded evaluation study

Questions Addressed in Analysis

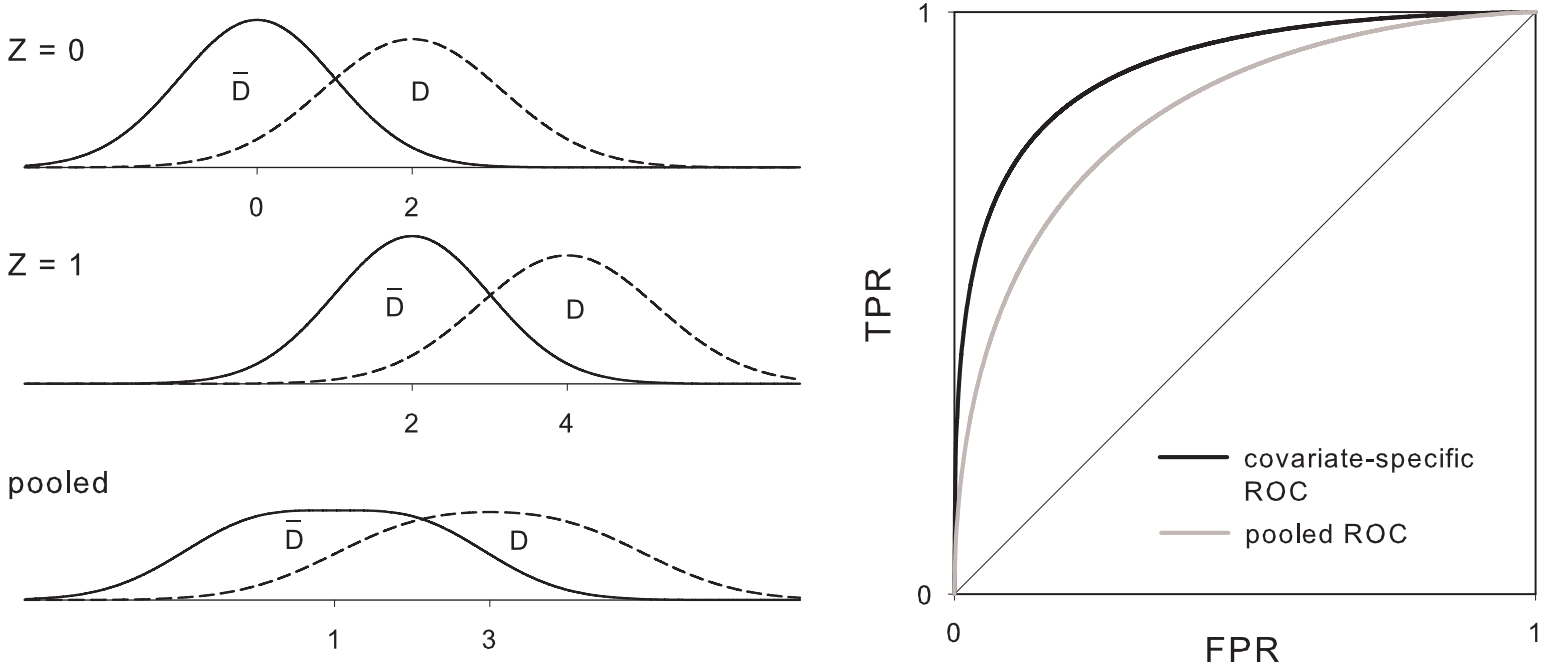
1. Is performance good enough for the clinical application?
   e.g., in the breast cancer diagnostic study: using the threshold corresponding to TPR=98% is the upper confidence limit for FPR $\leq$ 75%?

2. What factors substantially affect biomarker values in controls? Is covariate adjustment necessary?

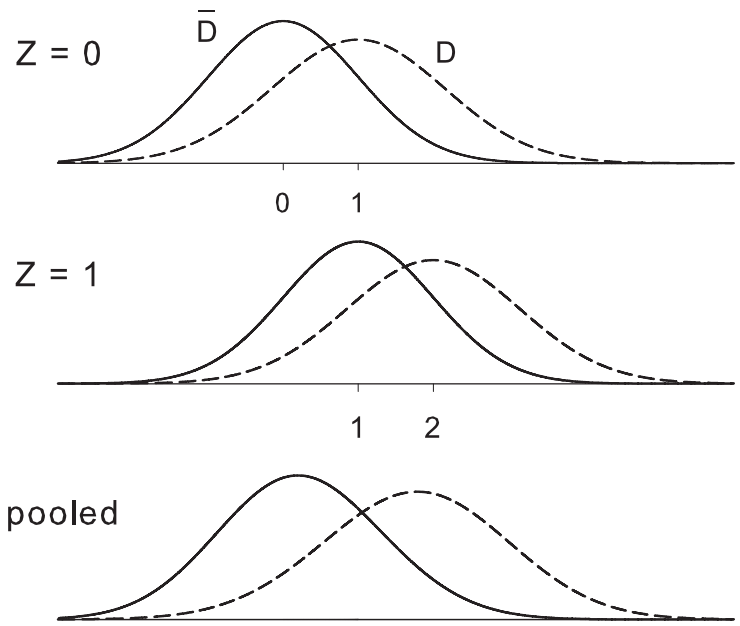
   e.g., stratify for study site?

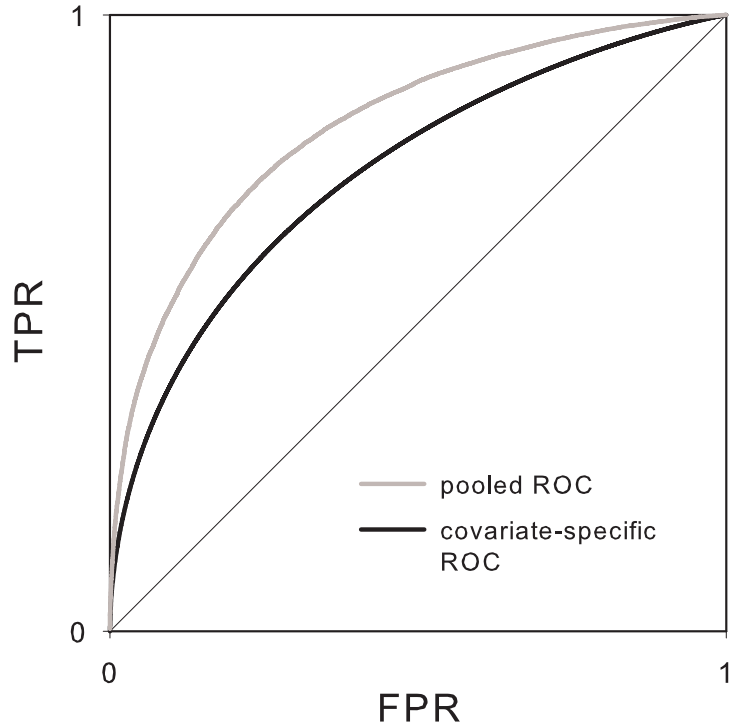
   e.g., adjust for age?

Study site (*Z*) affects biomarker distributions but not discrimination. Equal prevalence across study sites.

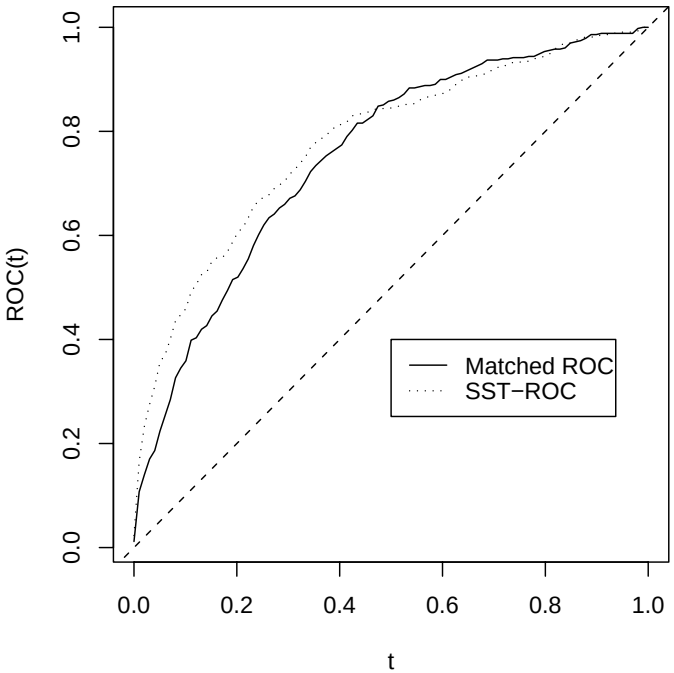


Study site (*Z*) affects biomarker distributions but not discrimination. Confounding caused by differing prevalence across study sites





Physician's Health Study, PSA as a marker for Prostate Cancer Matched on Age



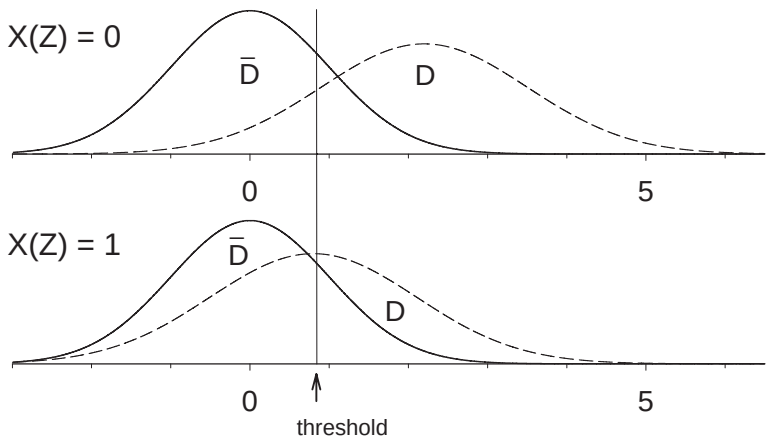
- ROC(0.05) = 0.23*A*ROC (0.05) = 0.36
- log(*PSA*) = $\alpha_0 + \alpha_1$ age + ε

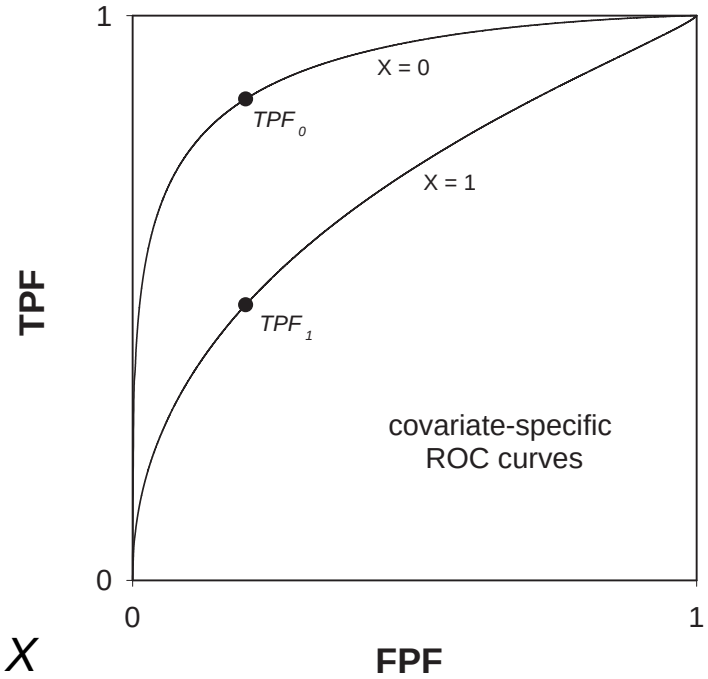
Questions Addressed in Analysis

3. What factors affect biomarker performance?

Disease specific factors : histology, stage …

Non-disease specific factors: age, study site …





Performance varies with *X*

Questions Addressed in Analysis

4. Incremental value of a marker over existing predictors

- Comparative study

- ROC curves for (i) baseline predictors
  (ii) marker and baseline

- Statistically significant effect in logistic regression is not enough

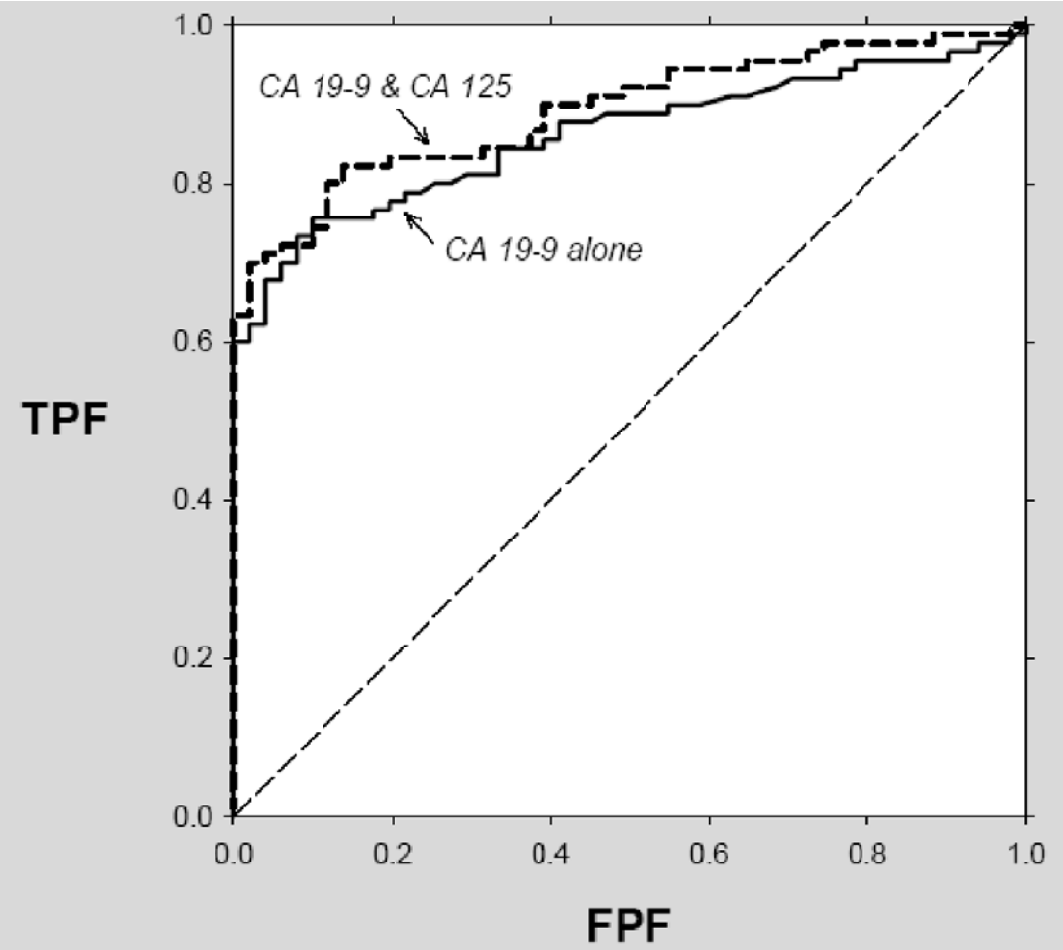
- Matching on baseline variables (e.g., age) can render incremental value non-identifiable.

Example

$X1$ = log CA19-9 X2 = log CA-125

Logit$P(D=1|X1,X2) = \alpha + \beta 1X1 + \beta 2X2$

$exp(\beta 2) = 1.54$ (p=0.002)



$$R\hat{O}C(0.05) = 68\% \ \ for \ X_1$$

$$R\hat{O}C(0.05) = 71\% \ \ for \ (X_1, X_2) \ combination$$

Summary

Biomarker evaluation must be done in the context of clinical application and population of interest.

Setting performance criteria is crucial and difficult

Standards of practice are needed for definitive validation studies

- Design standards
- Reporting standards