Metrics for Quantifying and Comparing Markers Used for Treatment Selection

November 22, 2010

Holly Janes and Margaret Sullivan Pepe
Fred Hutchinson Cancer Research Center

with

Patrick Bossuyt and Bill Barlow

FRED HUTCHINSON
CANCER RESEARCH CENTER

A LIFE OF SCIENCE

Many Roles for Biomarkers

- surrogate outcomes
- early diagnosis of disease (EDRN)
- prognosis
- treatment selection (predictive) (prescriptive)

A Motivating Example

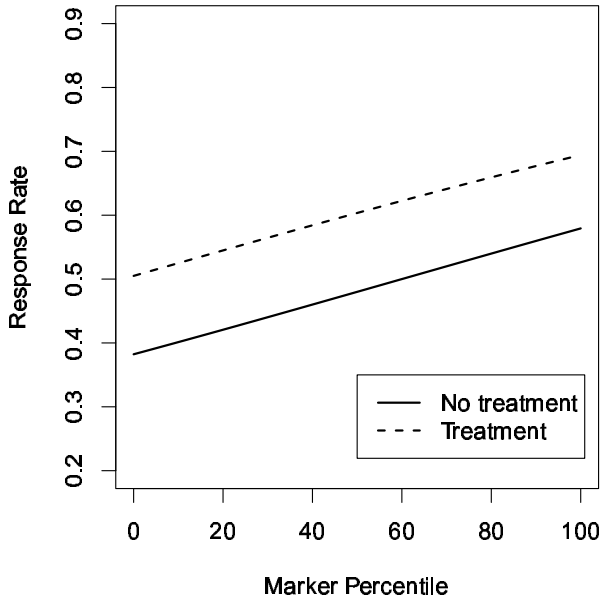
- most women with ER+ breast cancer are treated with chemotherapy but likely only a subset benefit
- Onco*type* DX is an RT-PCR assay on 21 genes: 16 cancer related, 5 reference
- Onco*type* DX useful for guiding the decision to have chemotherapy
- MammaPrint is another candidate marker

Not Targeted Therapy

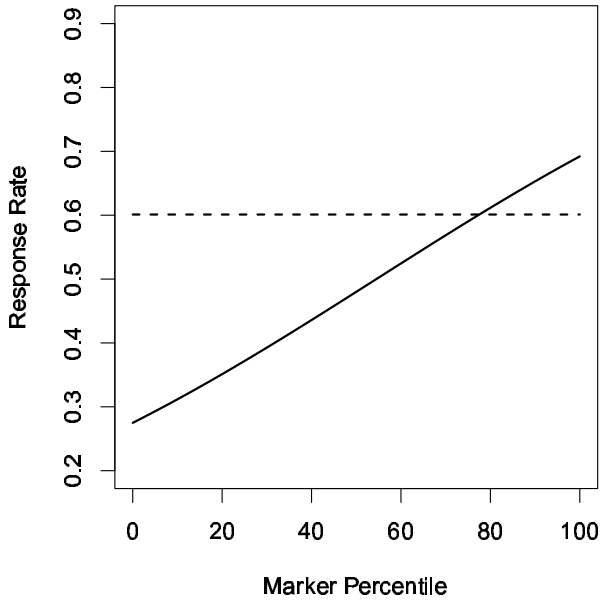
A Common Approach to Evaluating a Treatment Selection Marker

- is the marker predictive of response among subjects treated?
- not adequate

Marker equally prognostic on both treatments

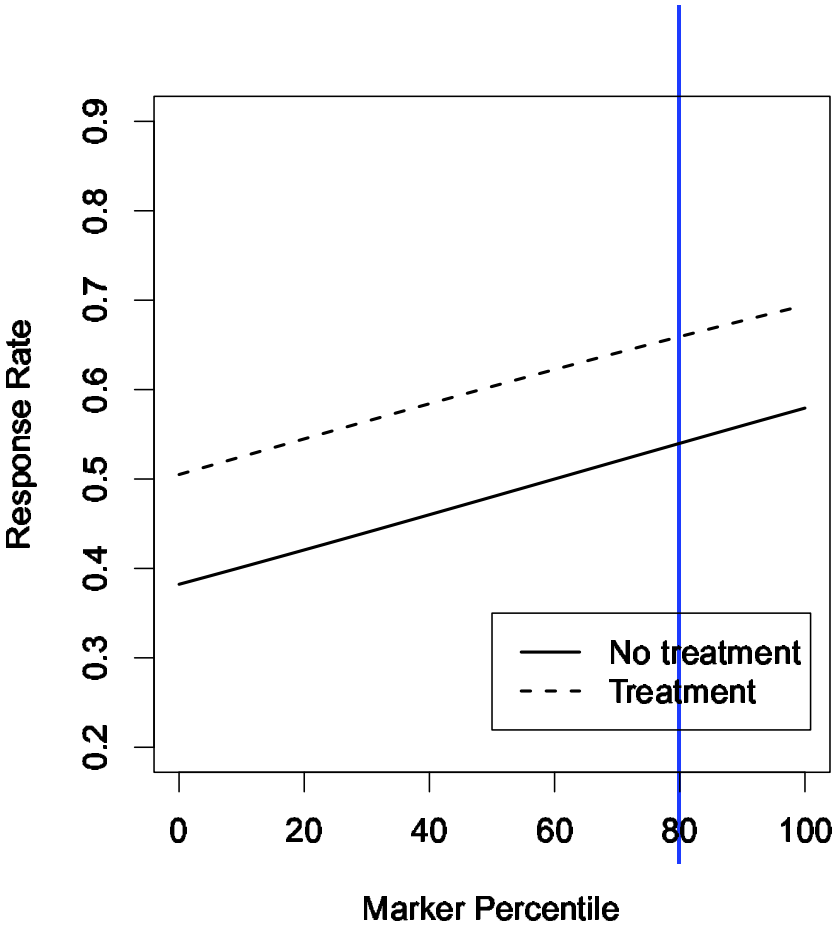


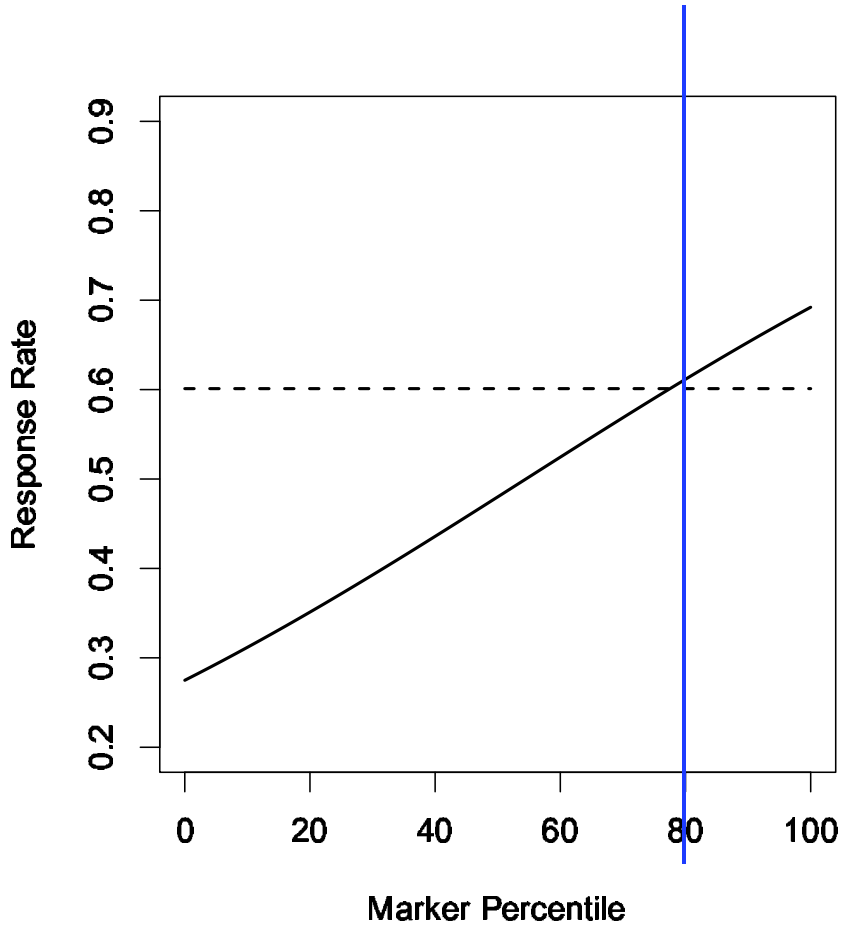
Marker not prognostic on treatment but useful for treatment selection



A Common approach to Evaluating a Treatment Selection Marker

- is treatment effective among marker positive patients?
- not adequate for evaluating *the marker*

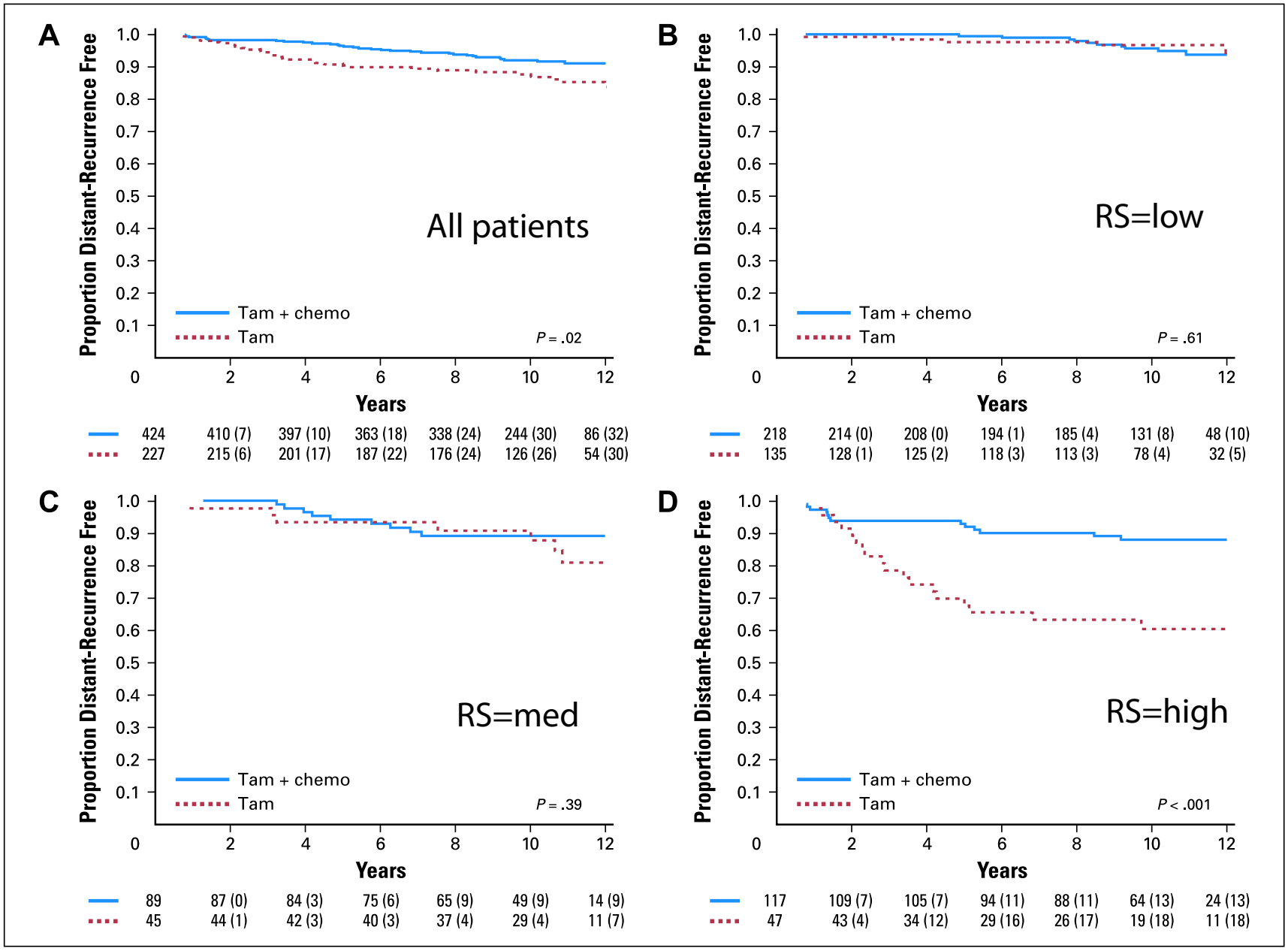




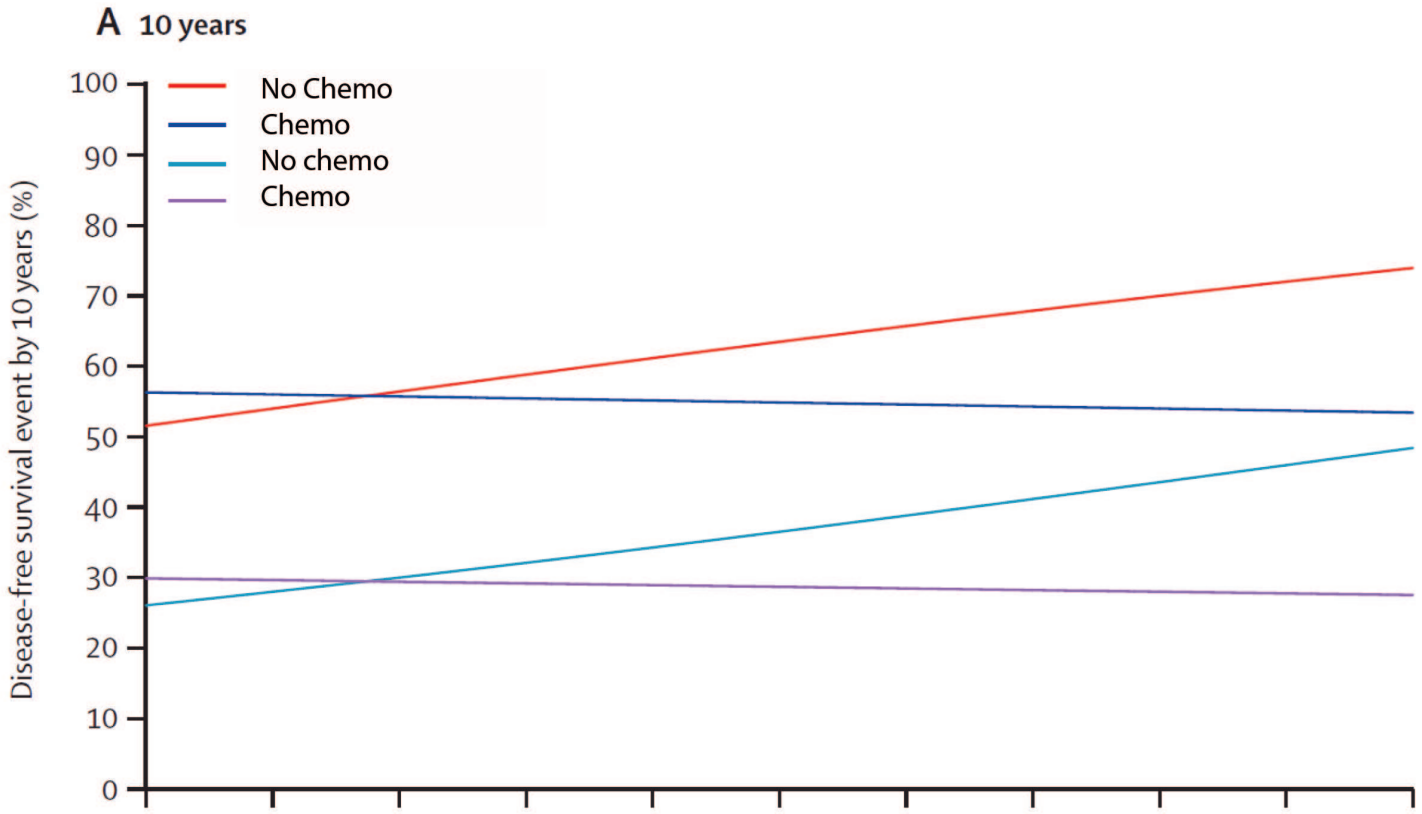
History of Onco*type* DX

- developed as a *prognostic* marker for women not receiving chemotherapy (Paik 2004)
- evaluated as a treatment selection marker
  - for node negative breast cancer (Paik et al 2006)
  - for node positive breast cancer (Albain et al 2010)

NSABP Trial (Paik et al 2006) Node Negative Breast Cancer



Breast Cancer Intergroup Study (Albain et al 2010)
Node Positive Breast cancer



The Paradigm

- biospecimens collected **p**rospectively before treatment in a RCT
- **r**etrospectively **b**linded **e**valuation of stored specimens
- PRoBE design for prognostic biomarkers extended to address treatment (Pepe et al *JNCI* 2008)

Simulated Data for Illustration

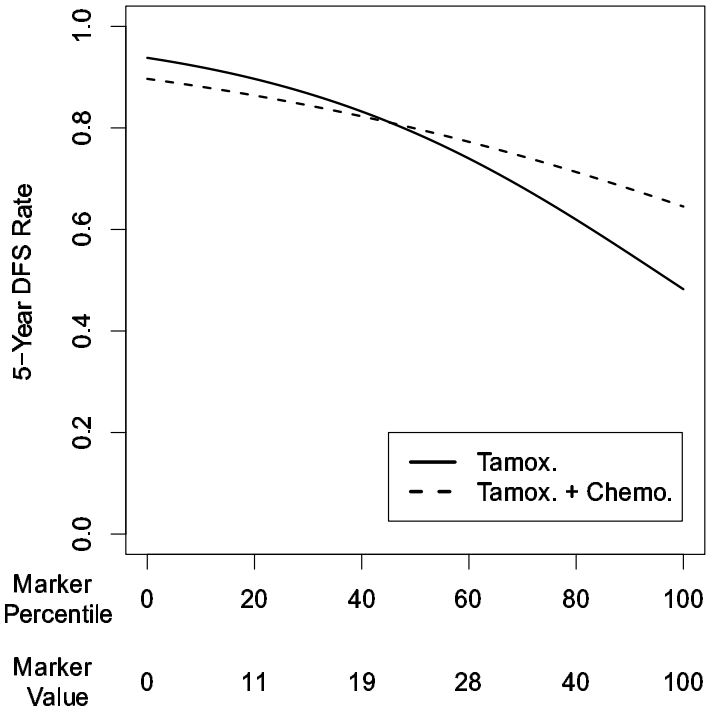
- Randomized trial with 3,000 participants
- Comparing no treatment ($T = 0$) vs. treatment ($T = 1$)
- 1,500 subjects each arm
- Response is 5-year disease-free survival (R)
- Chemotherapy is marginally effective:

$$P(R = 1|T = 1) = 79\% \text{ vs } P(R = 0|T = 0) = 76\%$$

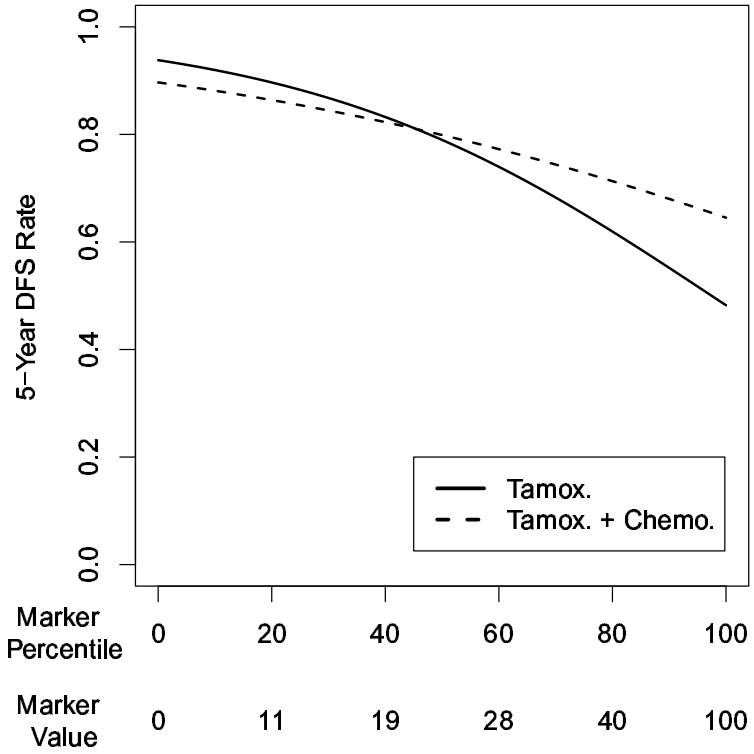
- Two markers measured at baseline on all trial participants:
  - Y_1 has performance similar to Onco*type* DX
  - Y_2 has better performance

Proposed Descriptive: Marker by Treatment Predictiveness Curves

- response rate as a function of marker percentile and treatment

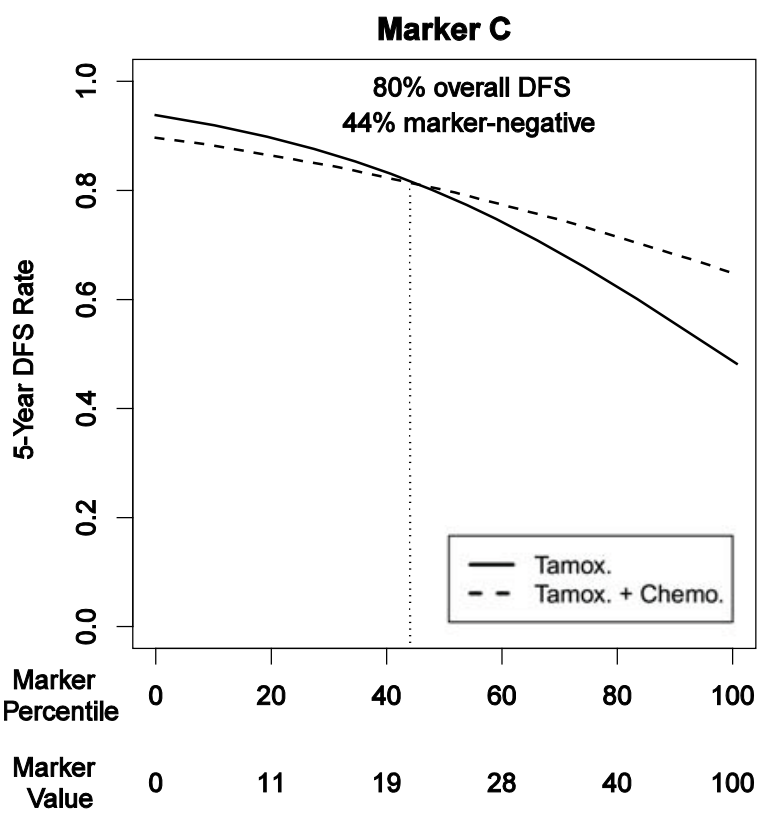


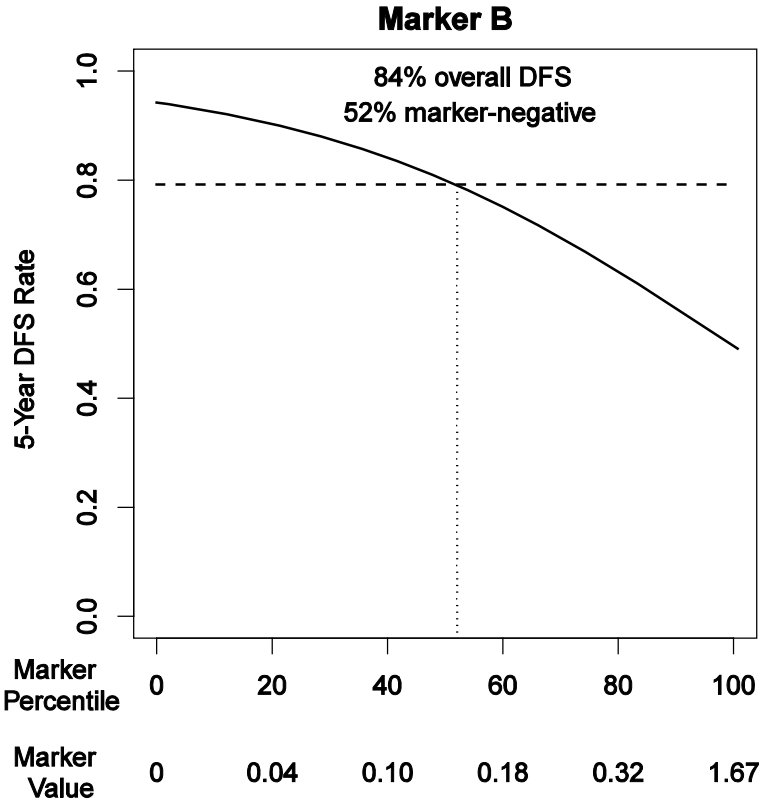
- similar to plots in Paik et al (2006) and Albain et al (2010)



- useful for guiding individual treatment decisions
- percentile scale shows % patients that opt to forego chemotherapy

- useful for comparing markers





Estimation

- $P(R = 1|T, Y)$ versus $F(Y)$ $T = 0, 1$
- extension of predictiveness curves for prognostic markers $P(R = 1|Y)$ versus $F(Y)$
- Huang, Pepe, Feng (*Biometrics* 2007) combine fitted risk model $P_{\widehat{\theta}}(R = 1|Y)$ with empirical $\widehat{F}(Y)$ using cohort data

More Estimation (Huang and Pepe)

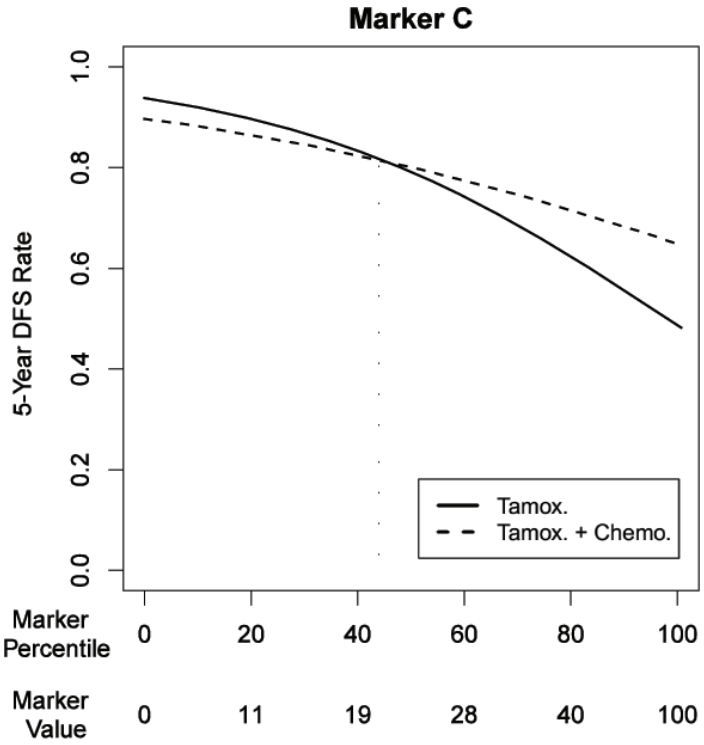
- nested case-control data
  - select on the basis of $R = 1$ and $R = 0$ for measuring Y
  - relevant also for Treatment Selection Markers
- semiparametric efficient estimation (*Biometrika* 2009)
- nonparametric estimation (*Statistics in Medicine* 2010)

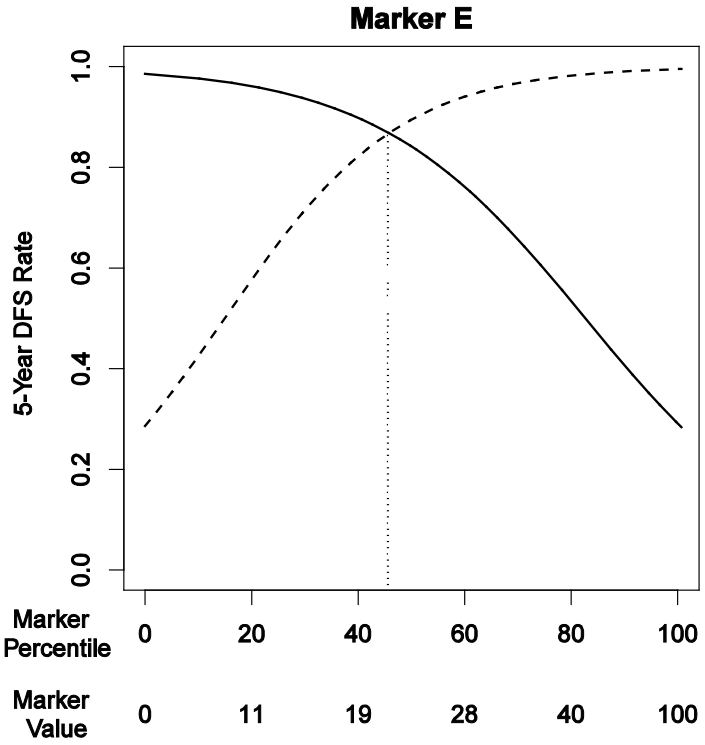
Estimation Accommodating Covariates

- how does the marker perform in subpopulations?
  e.g. node$+$ versus node$-$ patients
  e.g. younger versus older patients
- covariate specific predictiveness curves (*JRSS C* 2010)
- accommodates matching controls ($R = 0$) to cases ($R = 1$) in regards to covariates (and treatment)

Summary Index

- $\widehat{P}$[response|use of marker] $= \int max(\widehat{PC}_{T=1}(v), \widehat{PC}_{T=0}(v))dv$
- for comparing markers
- for discovery work



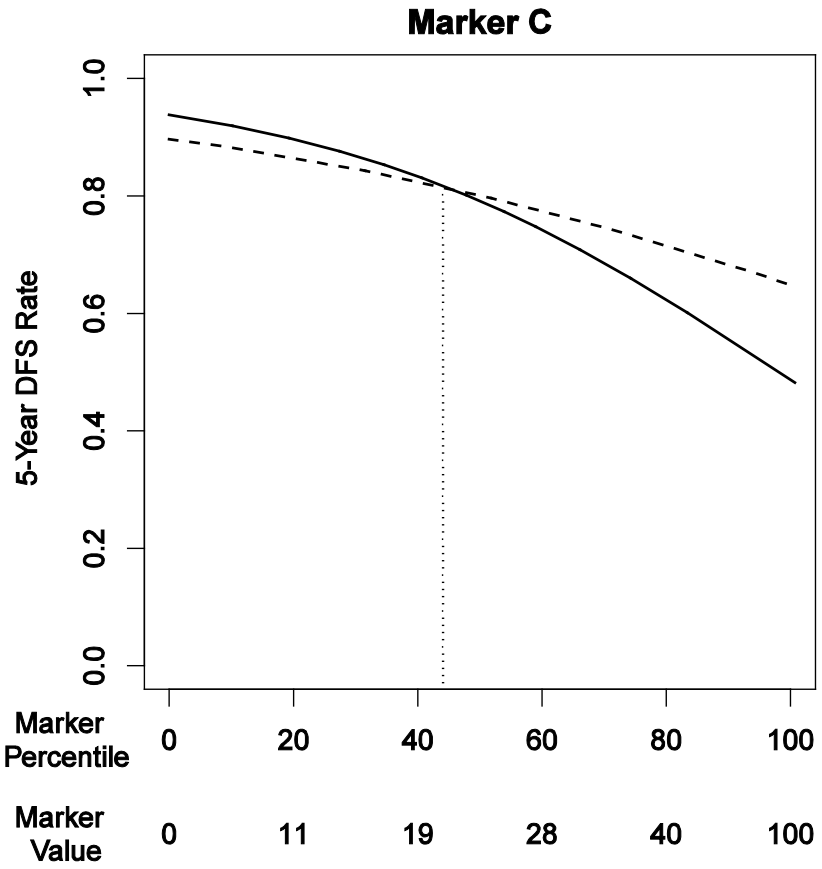


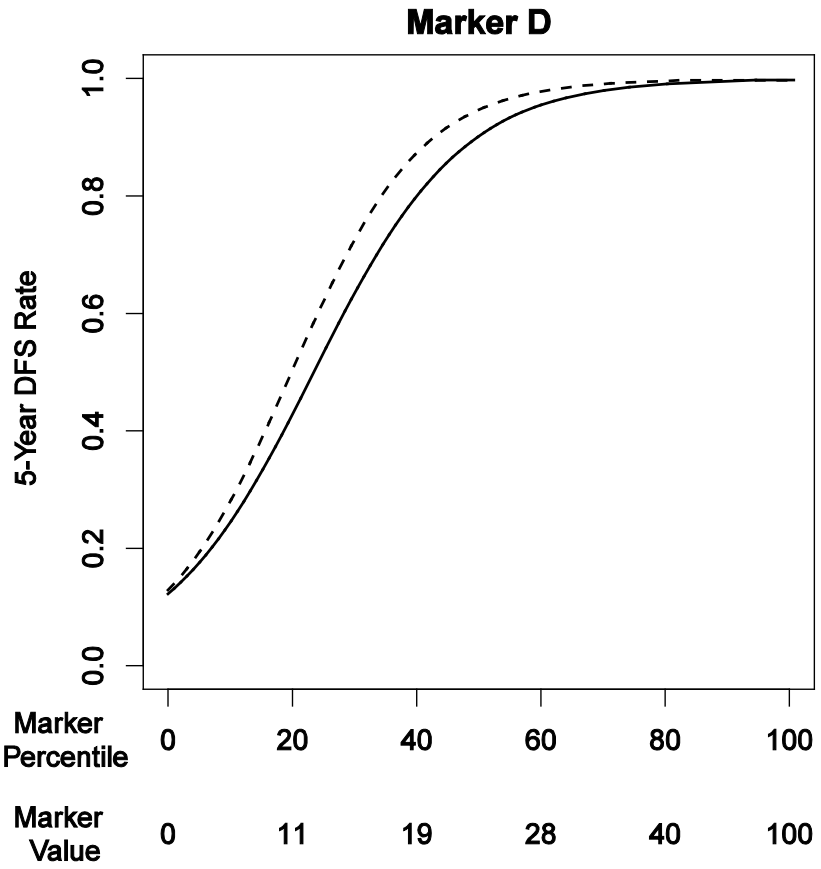
% 5-year DFS using marker	80 (78,82)	95 (98,96)

- CI and inference using the bootstrap (for now)
- p-value < 0.001

Interactions

- are not sufficient
- interactions do not *quantify* performance





- same interactions but different performance possible

Combining Markers

- multiple markers may be available
  - Onco*type* DX is a combination of 21 markers
  - combine MammaPrint and Onco*type* DX together
- add marker to existing clinical information
  - age, nodal status, ...

How to Combine Markers?

- $P(R = 1|\text{use of marker})$ is maximized by the rule

$$\text{Treat}(Y) : P(R = 1|T = 1, Y) > P(R = 1|T = 0, Y) \Rightarrow T = 1$$

$$\text{otherwise} \Rightarrow T = 0$$

Example

$$\text{logit}P(R = 1|T, Y) = \alpha_0 + \alpha_1 T + \alpha_2 Y + \alpha_3 TY$$

$$P(R = 1|T = 1, Y) > P(R = 1|T = 0, Y)$$

$$\Longleftrightarrow \alpha_1 + \alpha_3 Y > 0$$

- optimal marker combination $= \underset{\sim}{\alpha_3} \underset{\sim}{Y}$
- interaction of Y and T

Cost-Benefit Analysis Framework

- treatment selection marker needed if treatment has positive and negative effects
- chemotherapy $\begin{cases} \text{positive effects: tumor response} \\ \text{negative effects: toxicity} \end{cases}$
- how to put positive and negative effects on the same scale?
- composite outcome “any bad event”
  e.g., Onco*type* DX “disease free survival”
- more generally: assign costs and benefit values
- Expected benefit replaces $Prob(R = 1)$

Expected Benefit with Use of Marker (B)

C= cost of treatment

C_1= additional cost for those who respond in the absence of treatment

B_0= benefit for subjects that do not respond in the absence of treatment

$$B(Y) = P(R = 0|T = 0, Y)B_0 - P(R = 1|T = 0, Y)C_1 - C$$

When Costs and Benefits of Treatment do not Vary with Y

$$B(Y) = P(R = 0|T = 0, Y)B_0 - P(R = 1|T = 0, Y)C_1 - C$$

- prognostic marker can serve as treatment selection marker if it is adequately prognostic
- prognostic score is the right function for combining markers
- Gail (*JNCI* 2009) used the Gail model for breast cancer risk as a treatment selection marker assuming uniform costs and benefits
- Onco*type* DX is a prognostic score but costs and benefit of treatment may vary with components of it

Further Work

1. Visual displays for multiple markers
2. Evaluation of Incremental value
3. Extensions of the PRoBE design: selection of cases and controls; matching; sample size
4. Cost-Benefit Analysis Methods
5. Meta-analysis
6. Observational studies

Bibliography

Albain KS, Barlow WE, Shak S et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised clinical trial. Lancet Oncology 2010;11(1):55–65.

Gail MH. Value of Adding Single-Nucleotide Polymorphism Genotypes to a Breast Cancer Risk Model. J Natl Cancer Inst 2009;10(13):959–963.

Huang Y, Pepe MS. Assessing risk prediction models in case-control studies using semiparametric and nonparametric methods. Stat in Med 2010;29:1391–1410.

Huang Y, Pepe MS. Semiparametric methods for evaluating risk prediction markers in case-control studies. Biometrika 2009;96(4):991–997.

Huang Y, Pepe MS Semiparametric methods for evaluating the covariate-specific predictiveness of continuous markers in matched case-control studies. JRSS C Applied Statistics 2010;59:437–456.

Huang Y, Pepe MS, and Feng Z. Evaluating the predictiveness of a continuous marker. Biometrics 2007;63:1181–1188.

Paik S, Shak S, Tang G, Kim C, Baker J, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med. 2004;351:2817-2826.

Paik S, Tang G, Shak S, et al: Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol 2006;24:3726–3734.

Pepe MS, Feng Z, Janes H, Bossuyt P and Potter J. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: Standards for study design. J Natl Cancer Inst 2008;100(20):1432–1438.