Personalized Predictive Medicine and Genomic Clinical Trials

Richard Simon, D.Sc.
Chief, Biometric Research Branch
National Cancer Institute
http://brb.nci.nih.gov

brb.nci.nih.gov

- Powerpoint presentations
- Reprints
- BRB-ArrayTools software
- Web based Sample Size Planning

Personalized Oncology is Here Today and Rapidly Advancing

- Key information is generally in the tumor genome, not in inherited genetics
- Personalization is based on limited stratification of traditional diagnostic categories, not on individual genomes (so far)

Personalized Oncology is Here Today

- Estrogen receptor over-expression in breast cancer
  - tamoxifen, aromatase inhibitors
- HER2 amplification in breast cancer
  - Trastuzumab, Lapatinib
- OncotypeDx in breast cancer
  - Low score for ER+ node - = hormonal rx
- KRAS in colorectal cancer
  - WT KRAS = cetuximab or panitumumab
- EGFR mutation or amplification in NSCLC
  - EGFR inhibitor

These Diagnostics Have Medical Utility

- They are actionable; they inform therapeutic decision-making leading to improved patient outcome
- Tests with medical utility help patients and can reduce medical costs

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 OCTOBER 23, 2008 VOL. 359 NO. 17

K-ras Mutations and Benefit from Cetuximab in Advanced Colorectal Cancer

Christos S. Karapetis, M.D., Shirin Khambata-Ford, Ph.D., Derek J. Jonker, M.D., Chris J. O'Callaghan, Ph.D., Dongsheng Tu, Ph.D., Niall C. Tebbutt, Ph.D., R. John Simes, M.D., Haji Chalchal, M.D., Jeremy D. Shapiro, M.D., Sonia Robitaille, M.Sc., Timothy J. Price, M.D., Lois Shepherd, M.D.C.M., Heather-Jane Au, M.D., Christiane Langer, M.D., Malcolm J. Moore, M.D., and John R. Zalcberg, M.D., Ph.D.*

ABSTRACT

BACKGROUND

Treatment with cetuximab, a monoclonal antibody directed against the epidermal growth factor receptor, improves overall and progression-free survival and preserves the quality of life in patients with colorectal cancer that has not responded to chemotherapy. The mutation status of the *K-ras* gene in the tumor may affect the response to cetuximab and have treatment-independent prognostic value.

METHODS

We analyzed tumor samples, obtained from 394 of 572 patients (68.9%) with colorectal cancer who were randomly assigned to receive cetuximab plus best supportive care or best supportive care alone, to look for activating mutations in exon 2 of the *K-ras* gene. We assessed whether the mutation status of the *K-ras* gene was associated with survival in the cetuximab and supportive-care groups.

RESULTS

Of the tumors evaluated for *K-ras* mutations, 42.3% had at least one mutation in exon 2 of the gene. The effectiveness of cetuximab was significantly associated with *K-ras* mutation status ($P=0.01$ and $P<0.001$ for the interaction of *K-ras* mutation status with overall survival and progression-free survival, respectively). In patients with wild-type *K-ras* tumors, treatment with cetuximab as compared with supportive care alone significantly improved overall survival (median, 9.5 vs. 4.8 months; hazard ratio for death, 0.55; 95% confidence interval [CI], 0.41 to 0.74; $P<0.001$) and progression-free survival (median, 3.7 months vs. 1.9 months; hazard ratio for progression or death, 0.40; 95% CI, 0.30 to 0.54; $P<0.001$). Among patients with mutated *K-ras* tumors, there was no significant difference between those who were treated with cetuximab and those who received supportive care alone with respect to overall survival (hazard ratio, 0.98; $P=0.89$) or progression-free survival (hazard ratio, 0.99; $P=0.96$). In the group of patients receiving best supportive care alone, the mutation status of the *K-ras* gene was not significantly associated with overall survival (hazard ratio for death, 1.01; $P=0.97$).

CONCLUSIONS

Patients with a colorectal tumor bearing mutated *K-ras* did not benefit from cetuximab, whereas patients with a tumor bearing wild-type *K-ras* did benefit from cetuximab. The mutation status of the *K-ras* gene had no influence on survival among patients treated with best supportive care alone. (ClinicalTrials.gov number, NCT00079066.)

From Flinders Medical Centre and Flinders University, Adelaide, Australia (C.S.K.); Bristol-Myers Squibb Research and Development, Princeton, NJ (S.K.-F.); Ottawa Hospital Research Institute, University of Ottawa, Ottawa (D.J.J.); National Cancer Institute of Canada Clinical Trials Group, Kingston, ON (C.J.O., D.T., S.R., L.S.); Austin Health, Melbourne, Australia (N.C.T.); National Health and Medical Research Council Clinical Trials Centre, University of Sydney, Sydney (R.J.S.); Allan Blair Cancer Centre, Regina, SK, Canada (H.C.); Cabrini Hospital and Alfred Hospital, Melbourne, Australia (J.D.S.); Queen Elizabeth Hospital and University of Adelaide, Adelaide, Australia (T.J.P.); Cross Cancer Institute, Edmonton, AB, Canada (H.-J.A.); Bristol-Myers Squibb, Wallingford, CT (C.L.); Princess Margaret Hospital, Toronto (M.J.M.); and Peter MacCallum Cancer Centre and University of Melbourne, Melbourne, Australia (J.R.Z.). Address reprint requests to Dr. Karapetis at the Department of Medical Oncology, Flinders Medical Centre, Flinders Dr., Bedford Park, SA 5042, Australia, or at c.karapetis@flinders.edu.au.

*Other participants in the CO.17 trial from the National Cancer Institute of Canada Clinical Trials Group and the Australasian Gastro-Intestinal Trials Group are listed in the Supplementary Appendix, available with the full text of this article at www.nejm.org.

N Engl J Med 2008;359:1757-65.

Home
Meetings
Abstracts & Virtual Meeting
Practice Resources
Education & Training
News
Press Center
ASCO News & Forum
• Feature Articles
Podcasts
Legislative & Regulatory
Quality Care & Guidelines

About ASCO
ASCO Bookstore
Careers in Oncology
Downloads & Technology
Foundation Grants & Awards
Membership
Research Policy
State Affiliates

Information for Patients CancerNet
The Journal of Clinical Oncology
The Journal of Oncology Practice
ASCO News & Forum
ASCO Cancer Foundation
The ASCO Press Center

ASCO Releases its First Provisional Clinical Opinion (PCO)

Patients with metastatic colorectal cancer who are candidates for anti-EFGR therapy should have their tumors tested for *KRAS* gene mutations, according to ASCO's first Provisional Clinical Opinion (PCO).

If a patient has a mutated form of the *KRAS* gene, the Society recommends *against* the use of anti-EFGR antibody therapy, based on recent studies indicating this treatment is only effective in patients with the normal (wild-type) form of the *KRAS* gene. It is estimated that 40% of patients with colon cancer have the *KRAS* mutation.

"Personalized medicine is the next frontier in cancer care," said Richard L. Schilsky, MD, ASCO President. "Using *KRAS* testing to guide colorectal cancer treatment is a prime example of where cancer care is heading."

"Basing cancer treatment on the unique genetic characteristics of the tumor or the individual with cancer will improve patient outcomes and help avoid unnecessary costs and side effects for patients who are unlikely to benefit," Dr. Schilsky added.

PCOs are intended to offer timely preliminary clinical direction to oncologists following the publication or presentation of potentially practice-changing data from major studies. ASCO's PCO on *KRAS* gene testing was given prior to the January 15-17, 2009 Gastrointestinal Cancers Symposium in San Francisco, California. The Symposium was co-sponsored by ASCO, the American Gastroenterological Association (AGA), the American Society for Radiation Oncology (ASTRO), and the Society of Surgical Oncology (SSO).

Among the 500 presentations was an important economic and scientific study that discussed the possibility of more than half a billion dollars in savings for the United States healthcare system. The study showed that routine testing for *KRAS* gene mutations in patients with metastatic colorectal cancer could save the U.S. health system up to $604 million per year by identifying who would benefit from the drug cetuximab.

Information on the PCO is currently available on ASCO.org, and the entire report will be published in the February, 1 2009 issue of the *Journal of Clinical Oncology* (JCO).

- Although the randomized clinical trial remains of fundamental importance for predictive genomic medicine, some of the conventional wisdom of how to design and analyze rct's requires re-examination
- The concept of doing an rct of thousands of patients to answer a single question about average treatment effect for a target population presumed homogeneous with regard to the direction of treatment efficacy in many cases no longer has an adequate scientific basis

- Cancers of a primary site often represent a heterogeneous group of diverse molecular diseases which vary fundamentally with regard to
  - the oncogenic mutations that cause them
  - their responsiveness to specific drugs

- How can we develop new drugs in a manner more consistent with modern tumor biology and obtain reliable information about what regimens work for what kinds of patients?

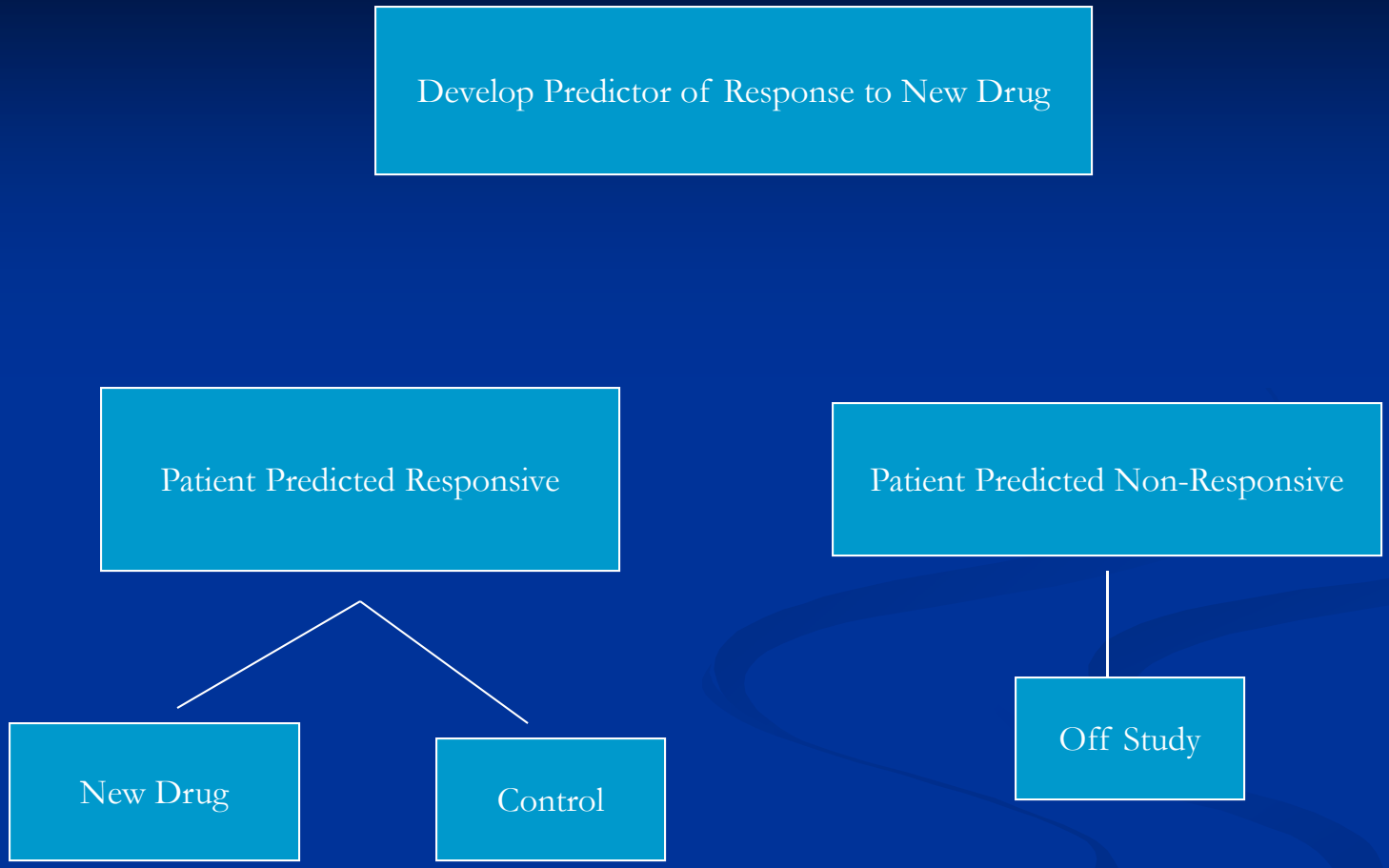
Developing a drug with a companion test increases complexity and cost of development but should improve chance of success and has substantial benefits for patients and for the economics of health care

Phase III Trial Development When the Biology is Clear

1. Develop a completely specified genomic classifier of the patients likely (or unlikely) to benefit from a new drug
2. Develop an analytically validated assay for the classifier
3. Design a focused clinical trial to evaluate effectiveness of the new treatment and how it relates to the test

Targeted (Enrichment) Design

- Restrict entry to the phase III trial based on the binary classifier

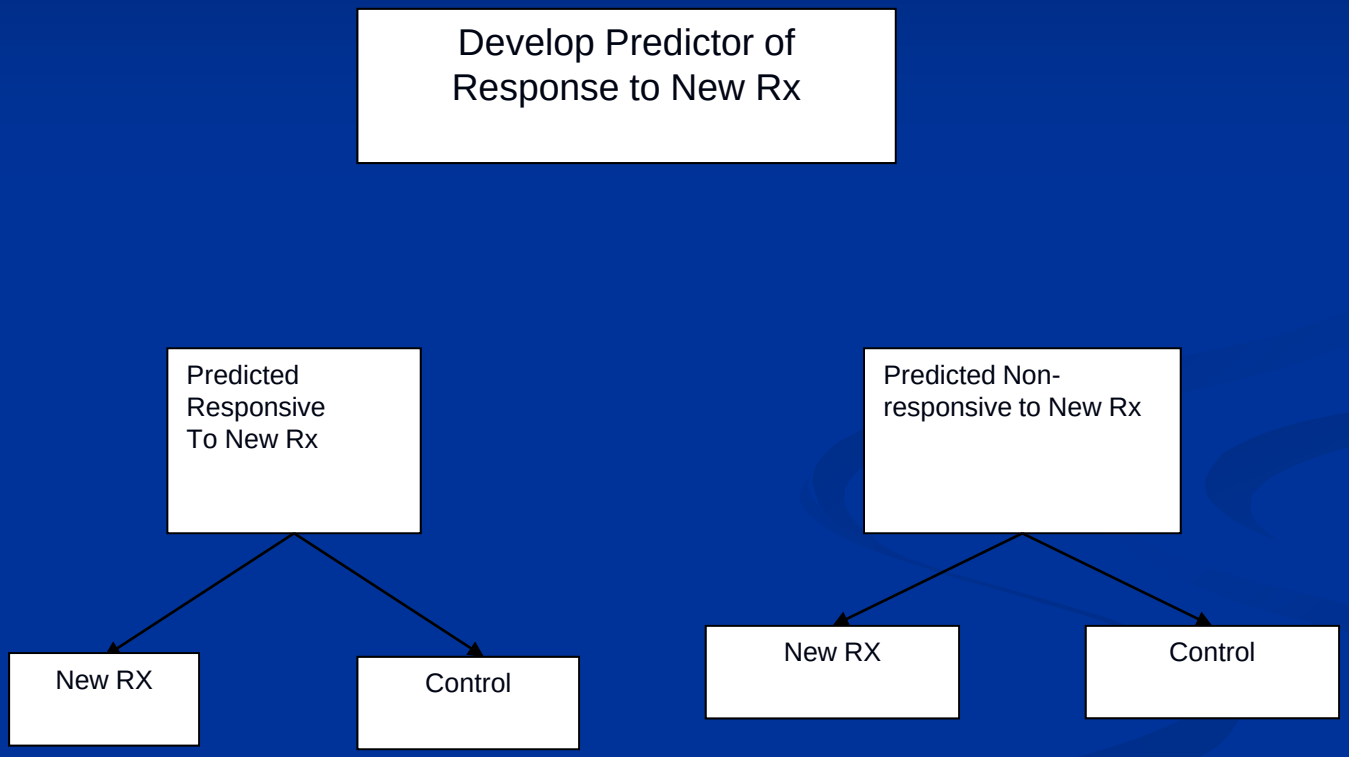

Develop Predictor of Response to New Drug
Patient Predicted Responsive
Patient Predicted Non-Responsive
New Drug
Control
Off Study

Evaluating the Efficiency of Targeted Design

- Simon R and Maitnourim A. Evaluating the efficiency of targeted designs for randomized clinical trials. Clinical Cancer Research 10:6759-63, 2004; Correction and supplement 12:3229, 2006
- Maitnourim A and Simon R. On the efficiency of targeted clinical trials. Statistics in Medicine 24:329-339, 2005.
- reprints and interactive sample size calculations at http://linus.nci.nih.gov

- Relative efficiency of targeted design depends on
  - proportion of patients test positive
  - effectiveness of new drug (compared to control) for test negative patients
- When less than half of patients are test positive and the drug has little or no benefit for test negative patients, the targeted design requires dramatically fewer randomized patients

Stratification Design



- Develop prospective analysis plan for evaluation of treatment effect and how it relates to biomarker
  - type I error should be protected
  - Trial sized for evaluating treatment effect overall and in subsets defined by test
- Stratifying" (balancing) the randomization is useful to ensure that all randomized patients have the test performed but is not necessary for the validity of comparing treatments within marker defined subsets

- R Simon. Using genomics in clinical trial design, Clinical Cancer Research 14:5984-93, 2008
- R Simon. Designs and adaptive analysis plans for pivotal clinical trials of therapeutics and companion diagnostics, Expert Opinion in Medical Diagnostics 2:721-29, 2008

Fallback Analysis Plan

- Compare the new drug to the control overall for all patients ignoring the classifier.
  - If $p_{overall} \leq 0.03$ claim effectiveness for the eligible population as a whole
- Otherwise perform a single subset analysis evaluating the new drug in the classifier + patients
  - If $p_{subset} \leq 0.02$ claim effectiveness for the classifier + patients.

Does the RCT Need to Be Significant Overall for the T vs C Treatment Comparison?

- No
  - That requirement has been traditionally used to protect against data dredging. It is inappropriate for focused trials with a prospective plan for a subset analysis with protected type I error

Web Based Software for Planning Clinical Trials of Treatments with a Candidate Predictive Biomarker

- http://brb.nci.nih.gov

Biomarker Stratified Randomized Design

Stratified design randomizes both marker positive and negative patients.

See references 73-75 in Technical Reports Section

- **Stratified Design with Prospective Analysis Plan and Binary Endpoint**
- **Stratified Design with Prospective Analysis Plan and Time-to-Event Endpoint**

The Biology is Often Not So Clear

- Cancer biology is complex and it is not always possible to have the right single completely defined predictive classifier identified and analytically validated by the time the pivotal trial of a new drug is ready to start accrual

Biomarker Adaptive Threshold Design

Wenyu Jiang, Boris Freidlin & Richard Simon

JNCI 99:1036-43, 2007

Biomarker Adaptive Threshold Design

- Have identified a candidate predictive biomarker score B but threshold of "positivity" has not been established
- Randomized trial of T vs C
- Eligibility not restricted by biomarker
- Time-to-event data

Procedure A
Fallback Procedure

- Compare T vs C for all patients
  - If results are significant at level .03 claim broad effectiveness of T
  - Otherwise proceed as follows

Procedure A

- Test T vs C restricted to patients with biomarker B > b
  - Let S(b) be log likelihood ratio statistic for rx effect
- Repeat for all values of b
- Let S* = max{S(b)}
- Compute null distribution of S* by permuting treatment labels
- If the data value of S* is significant at 0.02 level, then claim effectiveness of T for a patient subset
- Compute point and bootstrap confidence interval estimates of the threshold b

Multiple Biomarker Design

- Have identified K candidate binary classifiers B_1, …, B_K thought to be predictive of patients likely to benefit from T relative to C
- Eligibility not restricted by candidate classifiers
- For notation let B_0 denote the classifier with all patients positive

- Test T vs C restricted to patients positive for B_k for k=0,1,…,K
  - Let $S(B_k)$ be log partial likelihood ratio statistic for treatment effect in patients positive for B_k (k=1,…,K)
- Let $S^* = \max\{S(B_k)\}$, $k^* = \text{argmax}\{S(B_k)\}$
- For a global test of significance
  - Compute null distribution of S* by permuting treatment labels
  - If the data value of S* is significant at 0.05 level, then claim effectiveness of T for patients positive for B_{k^*}

Adaptive Signature Design

Boris Freidlin and Richard Simon

Clinical Cancer Research 11:7872-8, 2005

Biomarker Adaptive Signature Design

- Randomized trial of T vs C
- Large number of candidate predictive biomarkers available
- Eligibility not restricted by any biomarker

Cross-Validated Adaptive Signature Design

Freidlin B, Jiang W, Simon R

Clinical Cancer Research 16(2) 2010

Prediction Based Analysis of Clinical Trials

- This approach can be used with any set of candidate predictor variables

- Define an algorithm A for developing a classifier of whether patients benefit preferentially from a new treatment T relative to C
- For patients with covariate vector x, the classifier predicts preferred treatment
- Using algorithm A on the full dataset $\mathcal{D}$ provides a classifier model M(x;A, $\mathcal{D}$)
  - M(x;A, $\mathcal{D}$)) = T or M(x;A,$\mathcal{D}$)=C

- At the conclusion of the trial randomly partition the patients into K approximately equally sized sets $P_1, \ldots, P_K$
- Let D_{-i} denote the full dataset minus data for patients in P_i
- Using K-fold complete cross-validation, omit patients in P_i
- Apply the defined algorithm to analyze the data in D_{-i} to obtain a classifier M_{-i}
- For each patient j in P_i record the treatment recommendation i.e. $M_{-i}(x_j)$=T or C

- Repeat the above for all K loops of the cross-validation
- All patients have been classified as what their optimal treatment is predicted to be

- Let S_T denote the set of patients for whom treatment T is predicted optimal i.e. $S_T = \{j:M(x_j;A,D_{-i})=T$ where $x_j \varepsilon D_{-i}$
- Compare outcomes for patients in S_T who actually received T to those in S_T who actually received C
  - Compute Kaplan Meier curves of those receiving T and those receiving C
  - Let z_T = standardized log-rank statistic

Test of Significance for Effectiveness of T vs C

- Compute statistical significance of z_T by randomly permuting treatment labels and repeating the entire cross-validation procedure
  - Do this 1000 or more times to generate the permutation null distribution of treatment effect for the patients in each subset
- The significance test based on comparing T vs C for the adaptively defined subset S_T is the basis for demonstrating that T is more effective than C for some patients.

- By applying the analysis algorithm to the full RCT dataset $\mathcal{D}$, recommendations are developed for how future patients should be treated
  - M(x;A, $\mathcal{D}$) for all x vectors.
- The stability of the indication can be evaluated by examining the consistency of classifications $M(x_i;A, B)$ for bootstrap samples B from $\mathcal{D}$.

- The size of the T vs C treatment effect for the indicated population is (conservatively) estimated by the Kaplan Meier survival curves of T and of C in S_T

- Although there may be less certainty about exactly which types of patient benefit from T relative to C, classification may be better than for many standard clinical trial in which all patients are classified based on results of testing the single overall null hypothesis

70% Response to T in Sensitive Patients
25% Response to T Otherwise
25% Response to C
30% Patients Sensitive

	ASD	CV-ASD
Overall 0.05 Test	0.830	0.838
Overall 0.04 Test	0.794	0.808
Sensitive Subset 0.01 Test	0.306	0.723
Overall Power	0.825	0.918

35% Response to T
25% Response to C
No Subset Effect

	ASD	CV-ASD
Overall 0.05 Test	0.586	0.594
Overall 0.04 Test	0.546	0.554
Sensitive Subset 0.01 Test	0.009	0
Overall Power	0.546	0.554

25% Response to T
25% Response to C
No Subset Effect

	ASD	CV-ASD
Overall 0.05 Test	0.047	0.056
Overall 0.04 Test	0.04	0.048
Sensitive Subset 0.01 Test	0.001	0
Overall Power	0.041	0.048

Table 6. Results of applying CVASD to Bonnefoi et al. (2007) EORTC 10994 Neoadjuvant breast cancer data

Overall comparison	
$P = 0.79$	
Arm	**Observed pCR rate (%) (no of patients)**
FEC	42% (66)
TET	45% (58)

Sensitive subset comparison		
$P = 0.006$*		
Arm	**Estimates of pCR rates in the sensitive subpopulation**	
	Resubstitution	**CV**
FEC	20% (15)	29% (14)
TET	100% (8)	83% (12)

*P value based on permutation distribution of the cross-validated treatment effect in sensitive subset.

cross-validation procedure compared with the classifier developed using the full data set.

To illustrate our approach, we searched the National Center for Biotechnology Information Gene Expression Omnibus depository for publicly available gene expression data from a RCT. The only randomized cancer trial data with both expression and clinical outcome available that we were able to identify were data on a subset of 124 hormone receptor–negative breast cancer patients treated on EORTC 10994 (reported in ref. 9 and available at National Center for Biotechnology Information Web site).[3] EORTC 10994 was a phase III neoadjuvant breast cancer RCT that compared nontaxane regimen of 5-fluorouracil, cyclophosphamide, and epirubicin (FEC) with a taxane regimen of epirubicin and docetaxel (TET). In 66 patients treated with FEC, 28 had pathologic complete response (pCR), and in 58 patients treated with TET, 26 had pCR. CVASD was applied to these data (see Appendix B for details), and results are presented in Table 6. There was no overall difference in pCR rates between TET and FEC arms (pCR rates 45% and 42%, respectively; $P = 0.79$). The CVASD algorithm indicated the existence of a significant ($P = 0.006$) sensitive subset where TET is substantially more effective than FEC: the conservative (CV) estimate of the treatment effect was 83% pCR (TET) versus 29% pCR (FEC). Although providing a biological rationale for the sensitive patient signature is beyond the scope of this article, we note that two of the probes in the signature (Hs.310359.0.A1_3p_at and g4507484_3p_a_at) are related to the mitogen-activated protein kinase pathway that has been reported to be associated with anthracycline resistance in hormone receptor–negative breast cancer (10).

[3] http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE6861

As this is a retrospective application of CVASD to a subset of a reported RCT, these results will need an independent confirmation.

Discussion

Our results show that the cross-validation approach can considerably enhance the ASD performance. Cross-validation permits the maximization of the portion of study patients contributing to the development of the diagnostic signature [as shown by Molinaro et al. (11), this is critical in the high-dimensional data setting where the sample size or signal to noise ratio is limited]. Cross-validation also maximizes the size of the sensitive patient subset used to test (validate) the signature (this is important in settings where the fraction of the sensitive patients is small).

In this presentation, we used 80% to 20% allocation of the error rates between the overall and subset tests. This allocation represents a conservative approach that is aimed at preserving the ability of detecting the overall treatment effect without increasing the overall sample size. Depending on the amount of preliminary evidence that the treatment effect is limited to a subpopulation, one might allocate a higher proportion (up to 50%) of the overall error to the subset effect (i.e., using $\alpha_1 = 0.025$ and $\alpha_2 = 0.025$ for an overall 0.05 level design). The study sample size could be increased to achieve a desired power for the overall analysis or for the subset analysis. Sample size depends not only on the proportion of patients in the sensitive subset and the treatment effect in that subset, but also on aspects of the data used for classifier development. We plan to study sample size planning for the CVASD.

An important step in interpreting a trial that indicates that the effect of the new therapy is limited to a subset of patients is to provide an explicitly defined diagnostic test to identify the subpopulation of the future patients

- This approach can also be used to identify the subset of patients who don't benefit from a new regimen C in cases where T is superior to C overall at the first stage of analysis. The patients in $S_C = D - S_T$ are not predicted to benefit from T. Survivals of T vs C can be examined for patients in that subset and a permutation based confidence interval for the hazard ratio calculated.

Example of Classifier with Time to Event Data

- Fit proportional hazards model to dataset $\mathcal{D}$ or D_{-i}
  - With many candidate covariates, use L1 penalized proportional hazards regression
- f(x) = for patient with covariate vector x , log hazard if patient receives T minus log hazard if patient receives C
- M(x)=T if f(x)>k, M(x)=C otherwise
  - k optimized with inner cross-validation or a-priori based on toxicity of T

506 prostate cancer patients were randomly allocated to one of four arms: Placebo and 0.2 mg of diethylstilbestrol (DES) were combined as control arm C

1.0 mg DES, or 5.0 mg DES were combined as E.

The end-point was overall survival (death from any cause).

Covariates:
Age: In years

Performance status (pf): Not bed-ridden at all vs other

Tumor size (sz): Size of the primary tumor (cm2)

Index of a combination of tumor stage and histologic grade (sg)

After removing records with missing observations in any of the covariates, 485 observations remained.

A proportional hazards regression model was developed using patients in both E and C groups. Main effect of treatment, main effect of covariates and treatment by covariate interactions were considered.

$\log[HR(z,x)] = a\,z + b'x + z\,c'x$

$z = 0,1$ treatment indicator ($z=0$ for control)

x = vector of covariates

$\log[HR(1,x)] - \log[HR(0,x)] = a + c'x$

Define classifier $C(X) = 1$ if $a + c'x < c$

$= 0$ otherwise

c was fixed to be the median of the $a + c'x$ values in the training

Figure 1: Overall analysis. The value of the log-rank statistic is 2.9 and the corresponding p-value is 0.09. The new treatment thus shows no benefit overall at the 0.05 level.

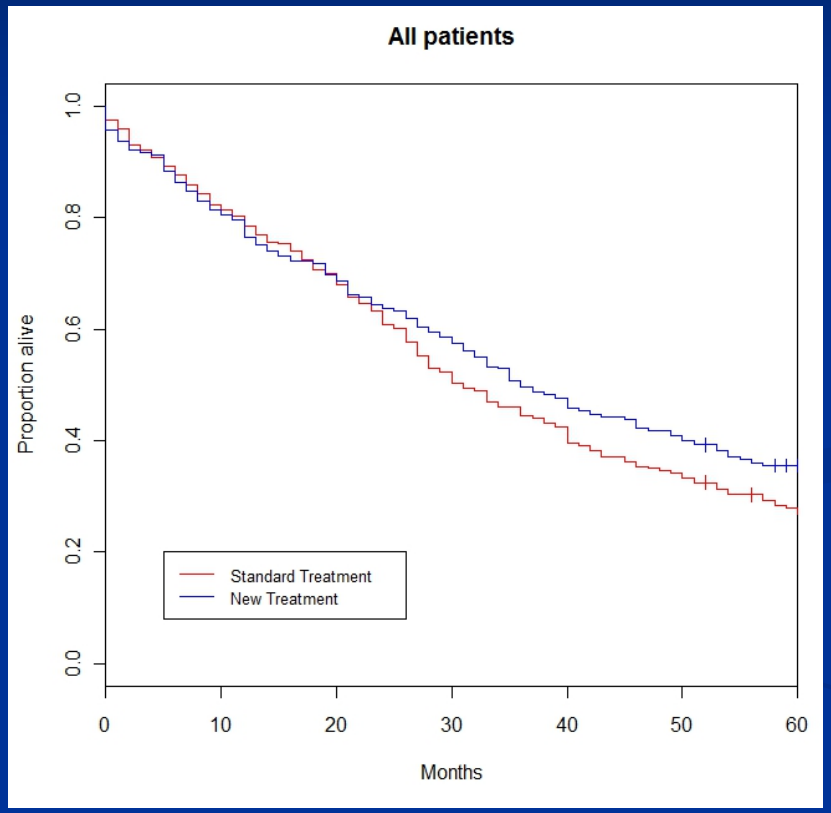


Figure 2: Cross-validated survival curves for patients predicted to benefit from the new treatment. log-rank statistic = 10.0, permutation p-value is .002

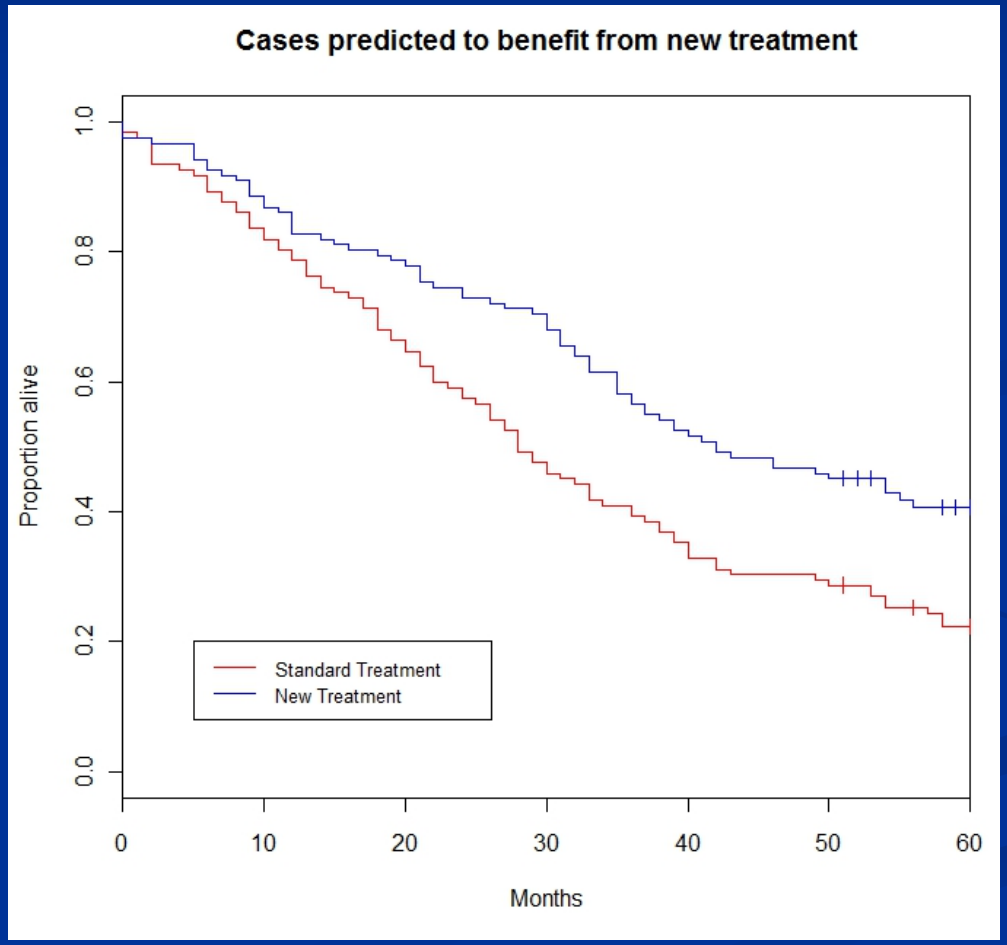
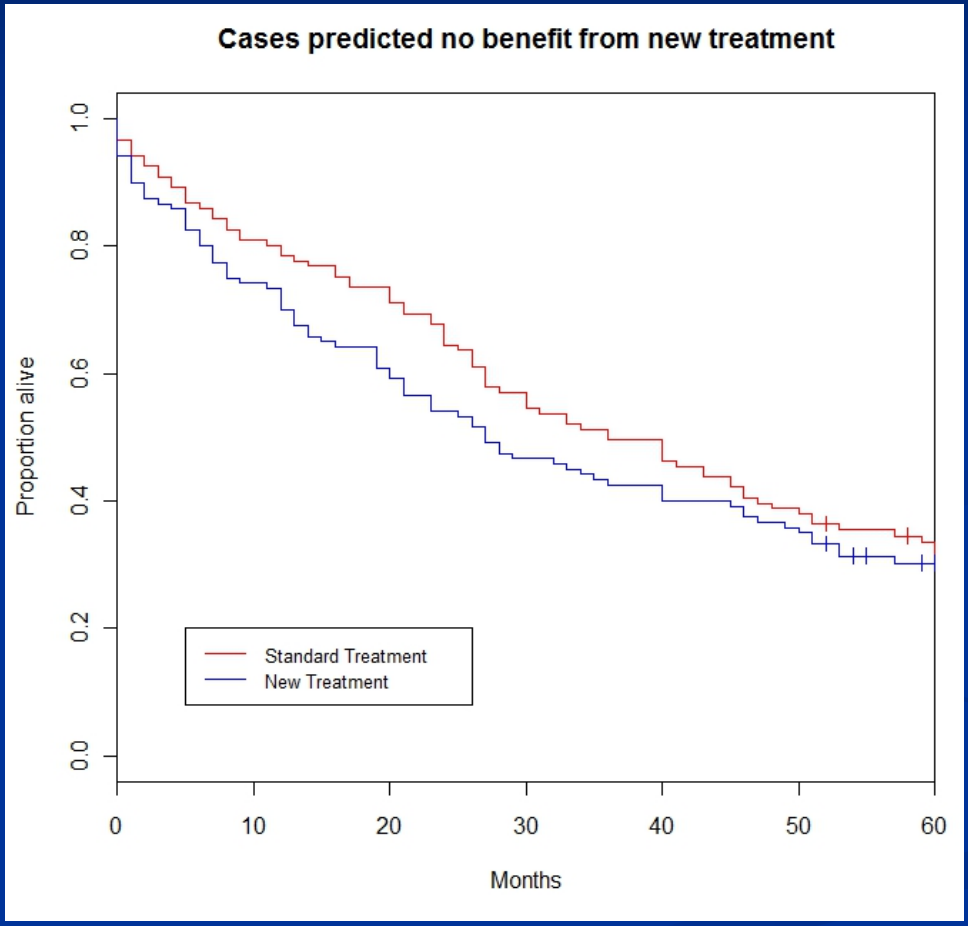


Figure 3: Survival curves for cases predicted not to benefit from the new treatment. The value of the log-rank statistic is 0.54.

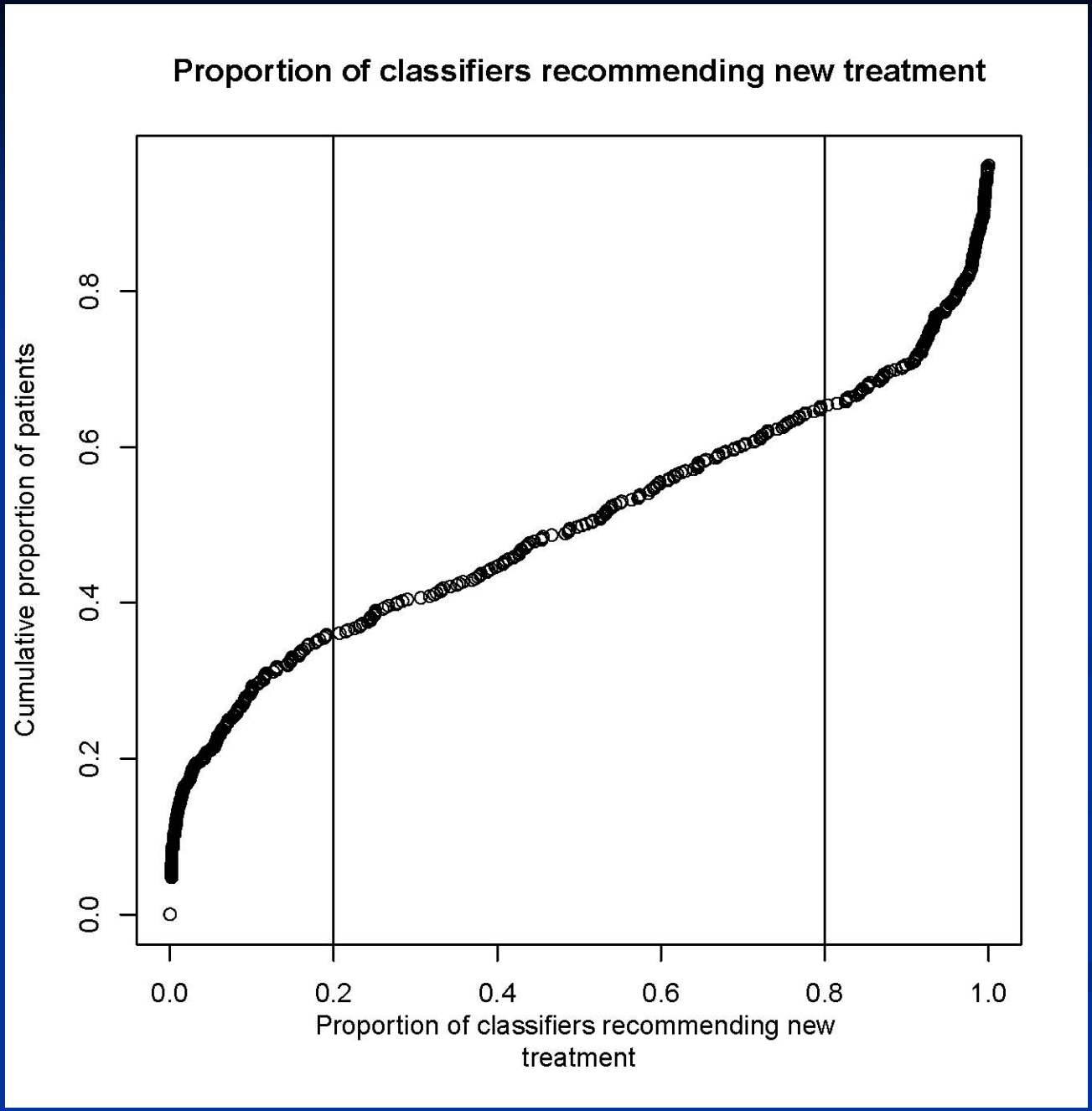


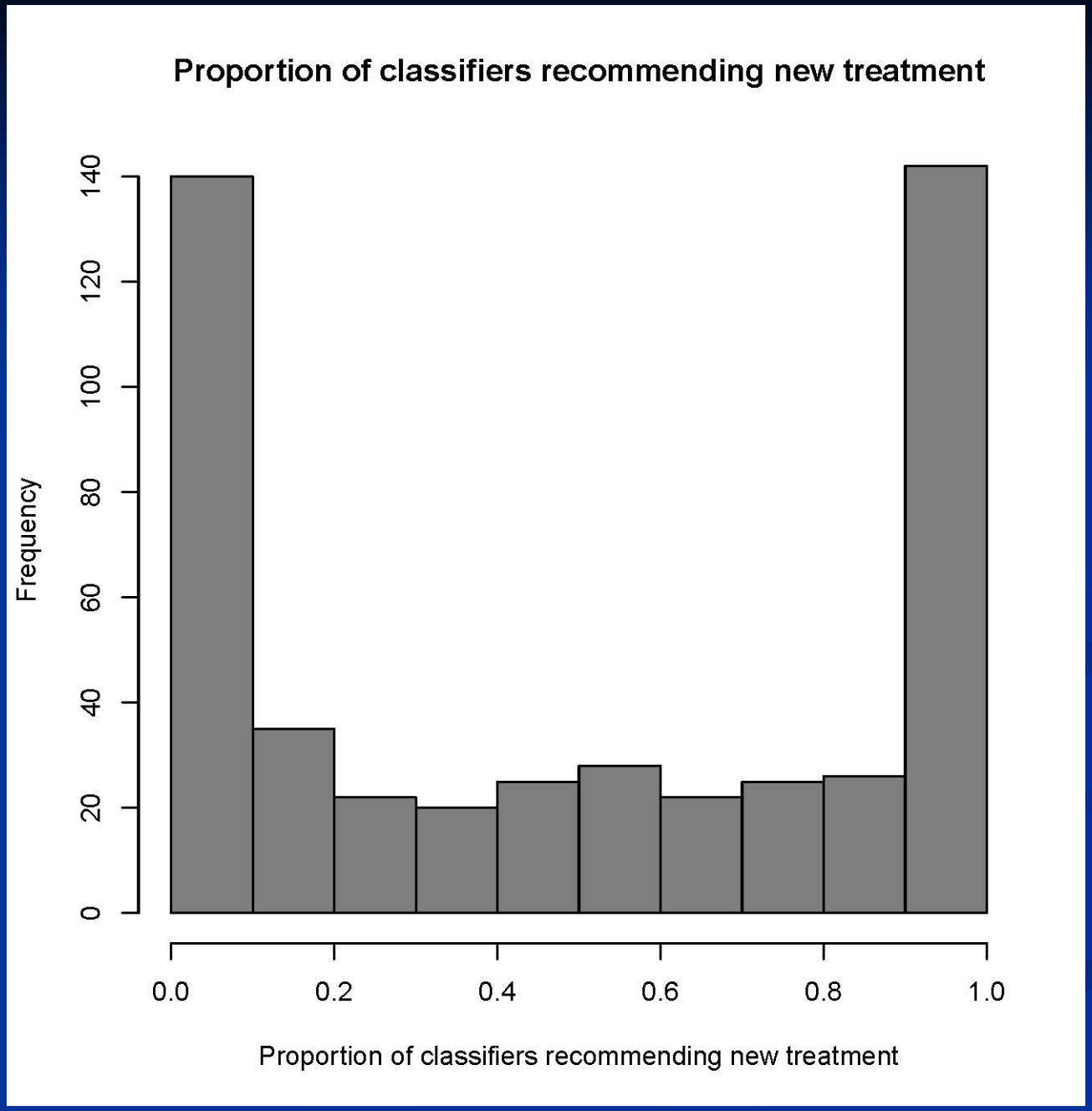
Proportional Hazards Model Fitted to Full Dataset

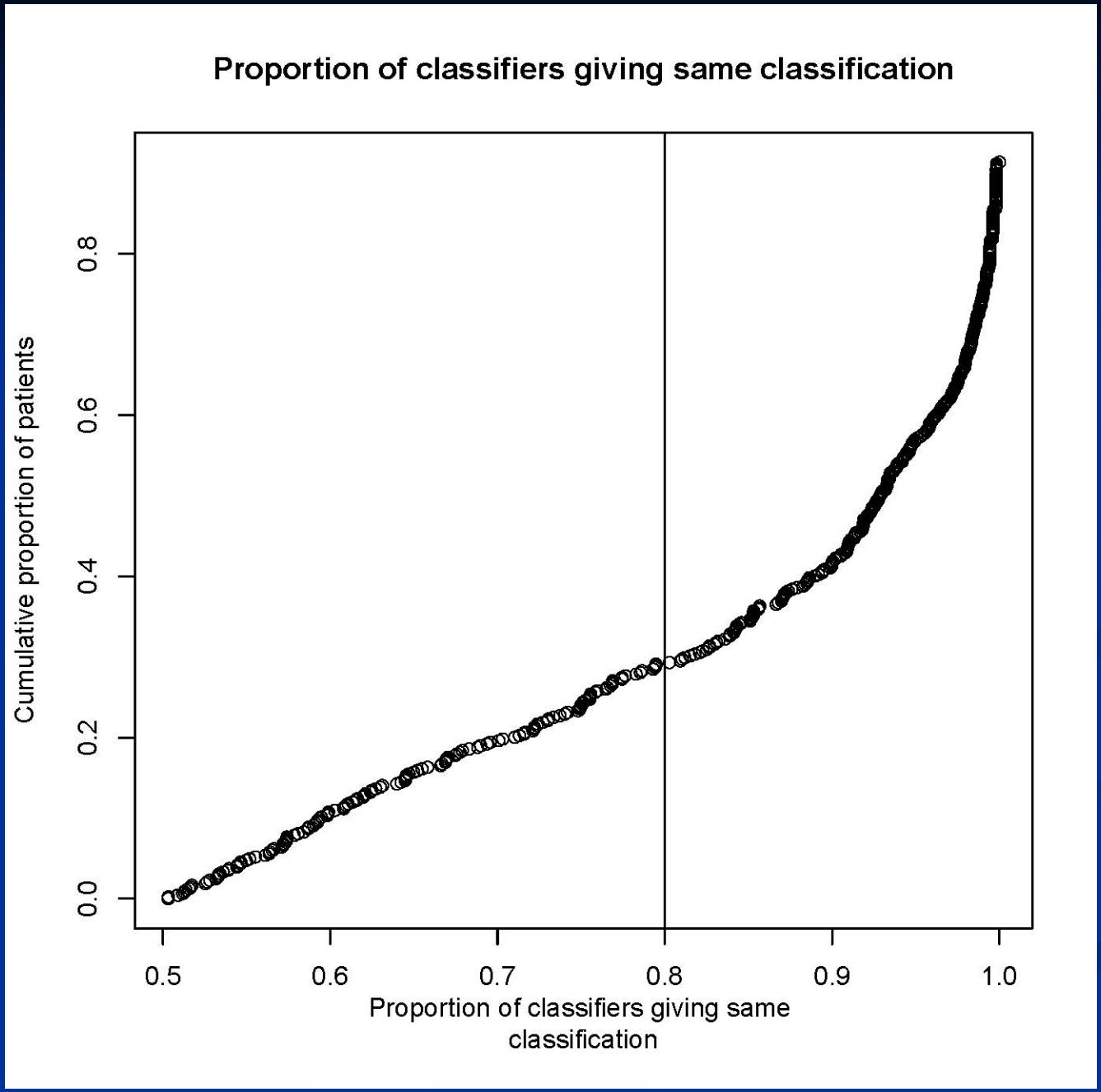
	coef	p-value
Treatment	-2.195	0.12
age	0.002	0.85
pf(Normal.Activity)	-0.260	0.25
sz	0.020	0.001
sg	0.113	0.004
ap	0.002	0.21
Treatment*age	0.050	0.003
Treatment*pf(Normal.Activity)	-0.743	0.026
Treatment*sz	-0.010	0.26
Treatment*sg	-0.074	0.19
Treatment*ap	-0.003	0.11

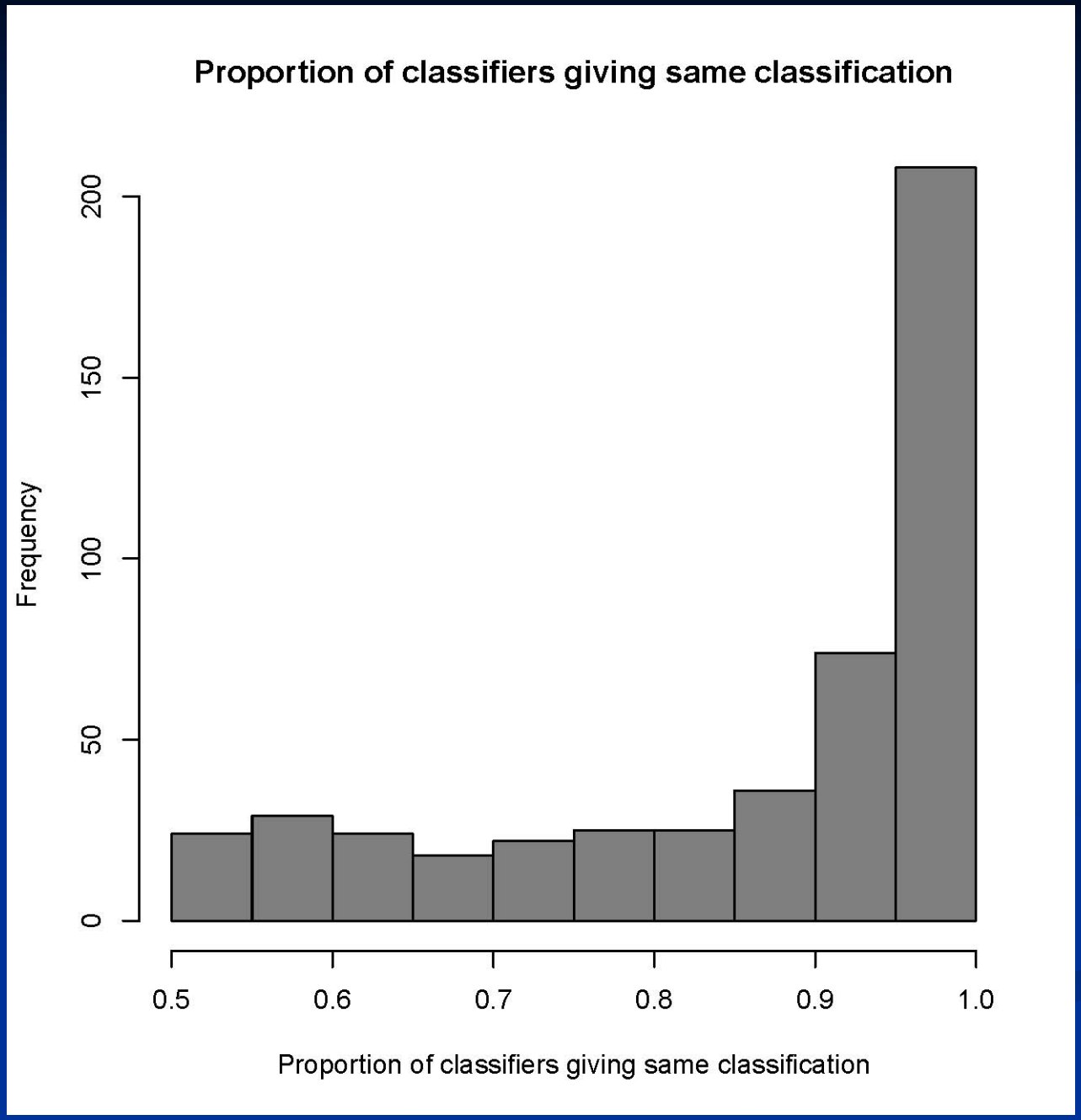
- By applying the analysis algorithm to the full RCT dataset D, recommendations are developed for how future patients should be treated; i.e. M(x; A,D) for all x vectors.

- The stability of the recommendations can be evaluated based on the distribution of M(x;A,D(b)) for non-parametric bootstrap samples D(b) from the full dataset D.


Proportion of classifiers recommending new treatment
Cumulative proportion of patients
0.0
0.2
0.4
0.6
0.8
0.0
0.2
0.4
0.6
0.8
1.0
Proportion of classifiers recommending new treatment


Proportion of classifiers recommending new treatment
Frequency
0
20
40
60
80
100
120
140
0.0
0.2
0.4
0.6
0.8
1.0
Proportion of classifiers recommending new treatment


Proportion of classifiers giving same classification
Cumulative proportion of patients
0.0
0.2
0.4
0.6
0.8
0.5
0.6
0.7
0.8
0.9
1.0
Proportion of classifiers giving same classification


Proportion of classifiers giving same classification
Frequency
0
50
100
150
200
0.5
0.6
0.7
0.8
0.9
1.0
Proportion of classifiers giving same classification

Biotechnology Has Forced Biostatistics to Focus on Prediction

- This has led to many exciting methodological developments
  - p>n problems in which number of covariates is much greater than the number of cases
- Statistics has over-emphasized inference and sometimes failed to adequately distinguish between inference and prediction problems
  - using prediction methods for inference and inferential methods for prediction
  - Failing to recognize the importance of prediction as a component of the analysis of clinical trials

Prediction Based Clinical Trials

- New methods for determining from RCTs which patients, if any, benefit from new treatments can be evaluated directly using the actual RCT data in a manner that separates model development from model evaluation, rather than basing treatment recommendations on the results of a single hypothesis test.

Prediction Based Clinical Trials

- Using cross-validation and careful prospective planning, we can more adequately evaluate new methods for analysis of clinical trials in terms of improving patient outcome by informing therapeutic decision making

Acknowledgements

- Boris Freidlin
- Yingdong Zhao
- Wenyu Jiang
- Aboubakar Maitournam