# Development and Validation of Predictive & Prognostic Biomarkers with High Dimensional Data

Noah Simon & Richard Simon

# Development and Validation of Predictive & Prognostic Biomarkers that Inform Treatment Decisions with High Dimensional Data

Noah Simon & Richard Simon

# Kinds of Biomarkers

- Early detection
- Diagnostic
- Prognostic
- Predictive
- Endpoint
  - Pharmacodynamic, intermediate, surrogate

# Prognostic Biomarker

- Measured before treatment to indicate long-term outcome of patient without treatment or receiving standard treatment
- Can be used to identify which patients have such good prognosis on "minimal" treatment that they don't require more intensive regimens

# Predictive Biomarker

- Measured before treatment to indicate who is likely or unlikely to benefit from a particular treatment

# Kinds of biomarkers

- Measurement of single analyte
- Scalar function of measurements of multiple analytes
  - $p>n$ or $p<n$

Contemporary Clinical Trials 36 (2013) 636–641

Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial

# Overfitting in prediction models – Is it a problem only in high dimensions?

Jyothi Subramanian [a], Richard Simon [b,*]

[a] Emmes Corporation, USA
[b] Biometric Research Branch, National Cancer Institute, USA

ARTICLE INFO

*Article history:*
Received 25 February 2013
Revised 29 May 2013
Accepted 17 June 2013
Available online 27 June 2013

*Keywords:*
Classifiers
Prediction accuracy
Overfitting
Clinical trials
Patient selection

ABSTRACT

The growing recognition that human diseases are molecularly heterogeneous has stimulated interest in the development of prognostic and predictive classifiers for patient selection and stratification. In the process of classifier development, it has been repeatedly emphasized that in situations where the number of candidate predictor variables is much larger than the number of observations, the apparent (training set, resubstitution) accuracy of the classifiers can be highly optimistically biased and hence, classification accuracy should be reported based on evaluation of the classifier on a separate test set or using complete cross-validation. Such evaluation methods have however not been the norm in the case of low-dimensional, $p < n$ data that arise, for example, in clinical trials when a classifier is developed on a combination of clinico-pathological variables and a small number of genetic biomarkers selected from an understanding of the biology of the disease. We undertook simulation studies to investigate the existence and extent of the problem of overfitting with low-dimensional data. The results indicate that overfitting can be a serious problem even for low-dimensional data, especially if the relationship of outcome to the set of predictor variables is not strong. We hence encourage the adoption of either a separate test set or complete cross-validation to evaluate classifier accuracy, even when the number of candidate predictor variables is substantially smaller than the number of cases.

Published by Elsevier Inc.

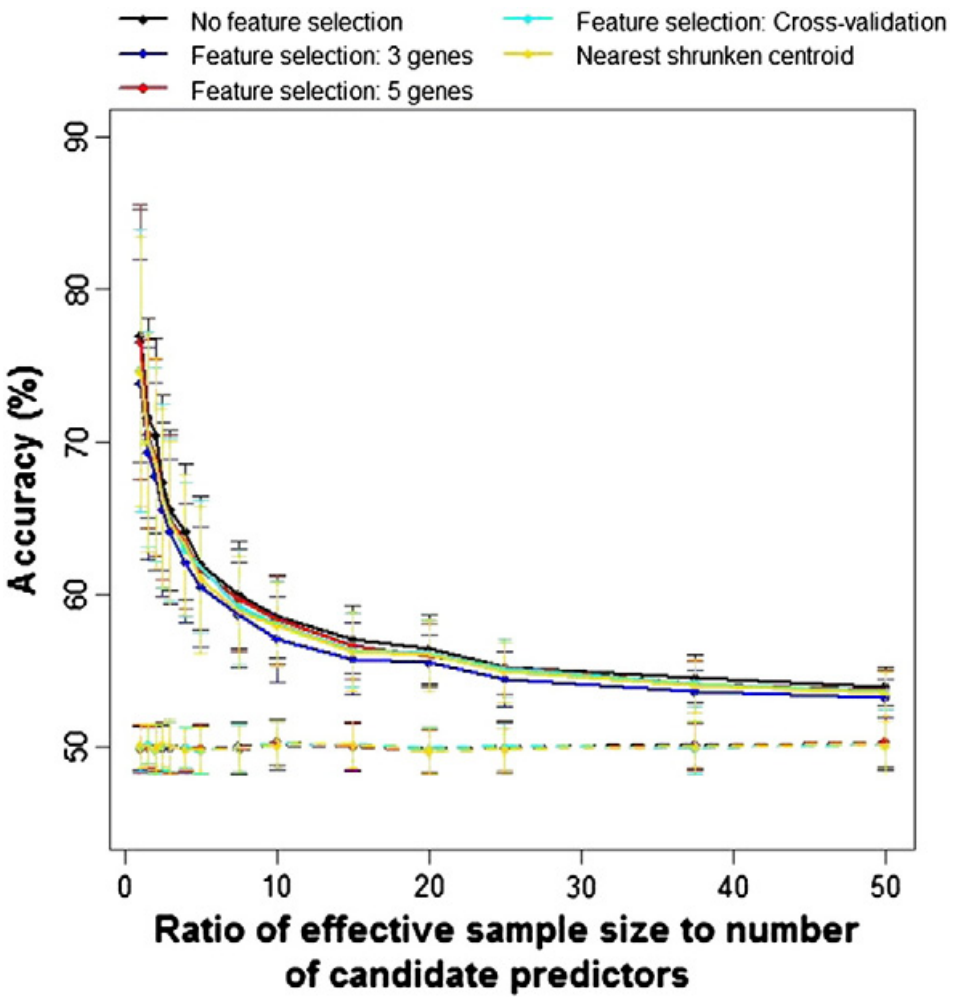



**Fig. 1.** Illustration of overfit in prediction models under the null. The dashed line represents the accuracy of classification as evaluated in the test set and the solid lines represent the accuracy of classification in the training set. The error bars represent $\pm 1$ standard deviation.

# Prognostic Biomarkers

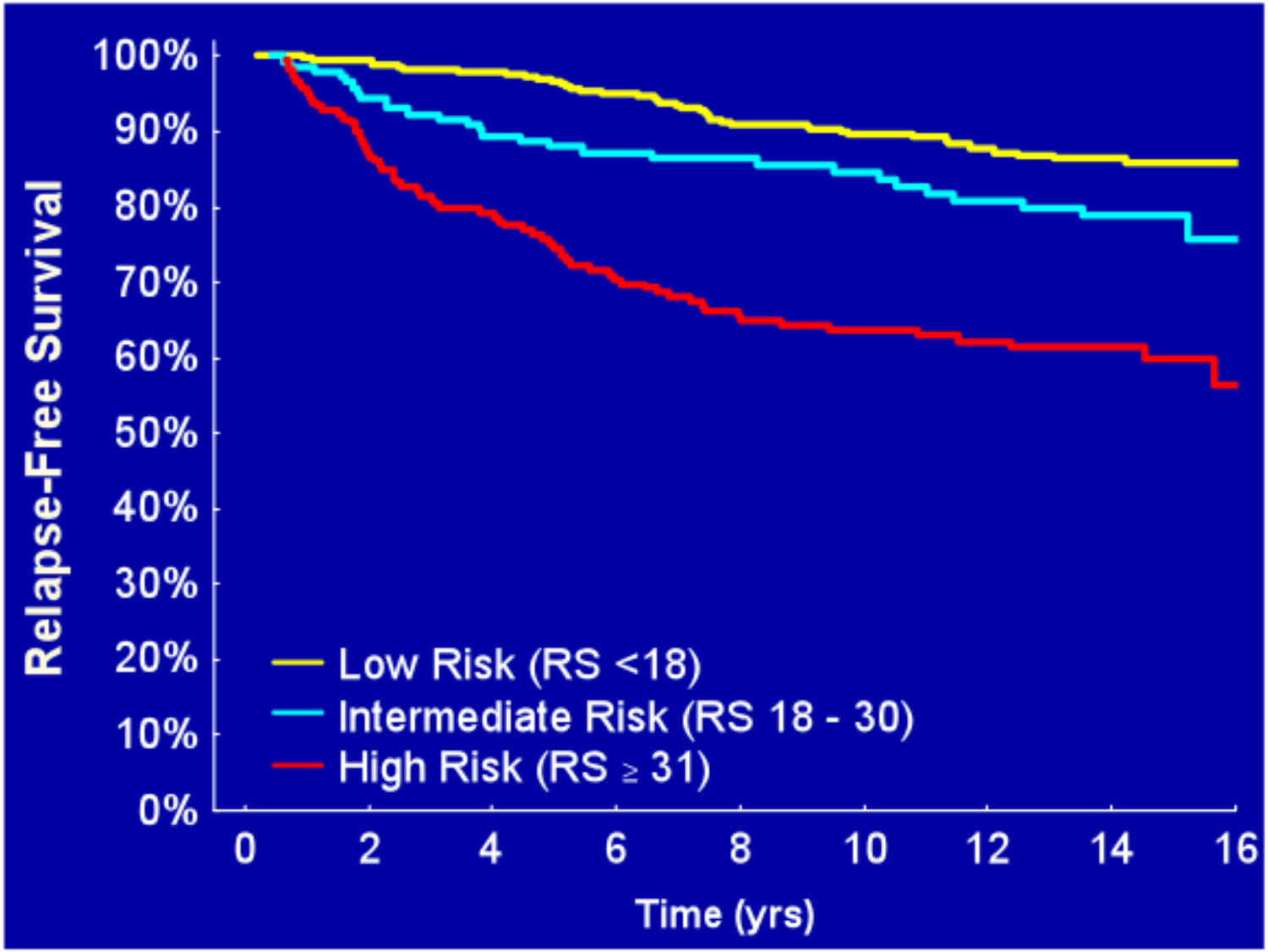
- Many prognostic studies do not develop models or biomarkers that inform treatment decisions
- They use heterogeneous convenience samples of cases
- The selection of cases and analysis are not driven by an intended use
- Over-emphasis on statistical significance and hazard ratios
- Biased estimates of prediction accuracy

# Oncotype DX

- Identify a subset of stage I, ER+ breast cancer patients who have such good outcome with only anti-estrogen therapy that they do not need chemotherapy

# Key Features of Oncotype DX Development

- Select cases with stage I, ER+ patients who received anti-estrogen therapy alone
- Analysis driven by objective of trying to identify a subset with such good outcome that they don't need other therapy
  - Not by FDR or what genes are significant or which model has the greatest separation of survival curves
- Separation of data used for model development from data used for validation


Relapse-Free Survival
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
0
2
4
6
8
10
12
14
16
Time (yrs)
Low Risk (RS <18)
Intermediate Risk (RS 18 - 30)
High Risk (RS ≥ 31)

# Key Features of Oncotype DX Development

- Avoided problems of multiplicity by not doing any model development or tweaking on the data used for model validation
- Strong analytical validation

# Kinds of Validation

- Analytical validation
  - Does it accurately measure the analyte and is it reproducible
- Clinical validation
  - Does it correlate with some clinical feature like outcome or stage
- Medical utility
  - Is it actionable in a way that benefits the patient
  - Requires clarity on intended use

# Single Arm Study with Binary Response

- Pathologic complete remission following pre-operative chemotherapy for patients with locally advanced breast cancer
- PCr indicates treatment effect on tumor even without a control group
- Whether that study informs treatment decisions depends on context of what other effective regimens are available

- So a prognostic biomarker can in some cases be established as useful for informing treatment decisions based on a single arm study with a survival (time-to-event) endpoint
  - Prospective clinical trial
  - "Prospective-retrospective" trial

- Development of a predictive biomarker model that informs selection of a new treatment or control often requires data from an RCT

# Development and Validation of Predictive Biomarker Usually Requires RCT

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.

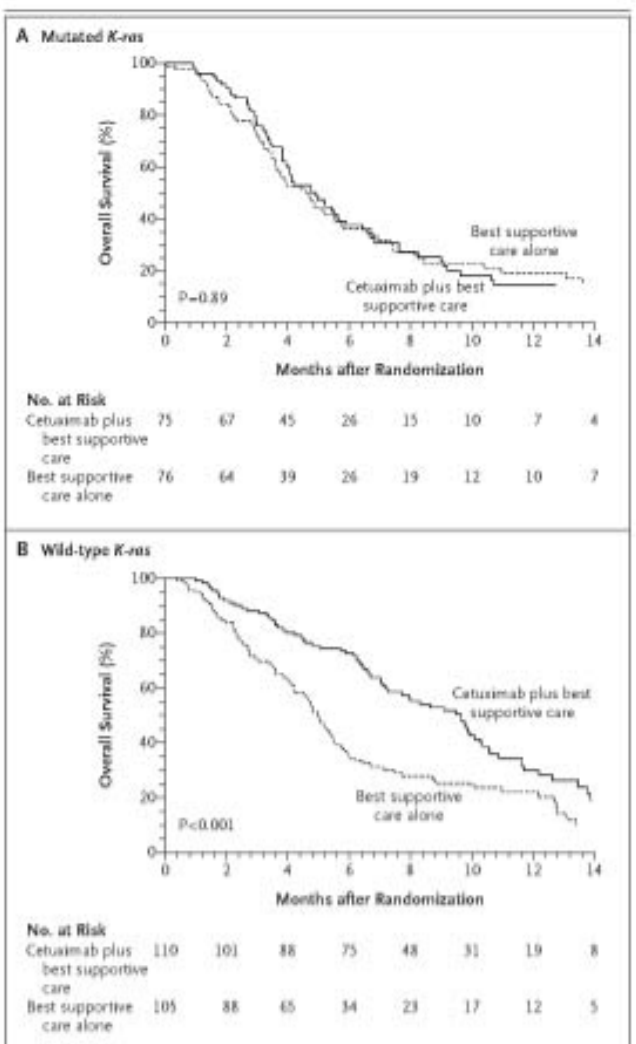



Figure 1. **Kaplan–Meier Curves for Overall Survival According to Treatment.**
Panel A shows results for patients with mutated *K-ras* tumors, and Panel B for patients with wild-type *K-ras* tumors. Cetuximab as compared with best supportive care alone was associated with improved overall survival among patients with wild-type *K-ras* tumors but not among those with mutated *K-ras* tumors. The difference in treatment effect according to mutation status was significant (test for interaction, P=0.01).

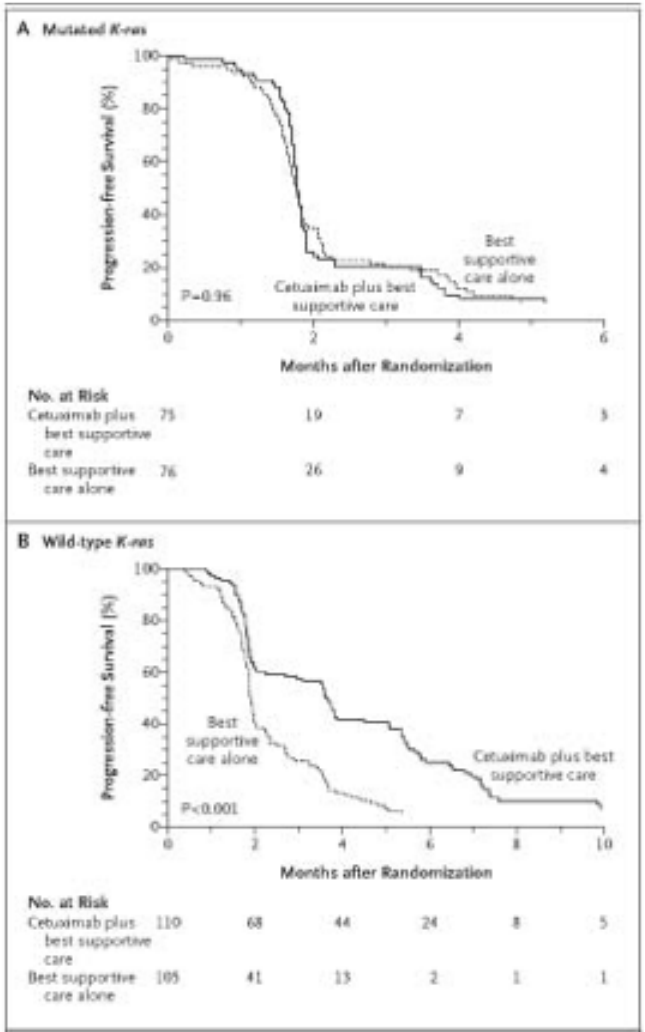



Figure 2. **Kaplan–Meier Curves for Progression-free Survival According to Treatment.**
Panel A shows results for patients with mutated *K-ras* tumors, and Panel B for patients with wild-type *K-ras* tumors. Cetuximab as compared with best supportive care alone was associated with improved progression-free survival among patients with wild-type *K-ras* tumors but not among those with mutated *K-ras* tumors. The difference in treatment effect according to mutation status was significant (test for interaction, P<0.001).