A tale of two models: PSA and its links to Progression Free Survival

Hitesh Mistry
Manchester Pharmacy School

What is PSA?

- Prostate Specific Antigen
  - Secreted by prostate epithelial cells
    - Increase in epithelial cells => increase in PSA
    - Increase in testosterone etc. => increase in PSA
    - Could be a useful marker for epithelial derived tumours
  - Measure of both pathway output and cell number
    - Makes it difficult to interpret
  - Other factors effect it's serum levels: Prostatitis, Benign Prostate Hyperplasia, Sex etc.
    - Makes it even more difficult to interpret
  - The setting is very important

These difficulties has led to PSA having a turbulent history when used as a prognostic/predictive marker within prostate cancer

Medics still see it as the gold standard though despite its caveats

ROC curves – Medical Obsession

- New plasma biomarkers in Oncology are routinely assessed by use of ROC curves
  - Medical community loves them!
- What are they?
  - Response variable has binary outcome
  - Threshold of a discriminant is varied through a range of values and true positive and true negative values are recorded
  - The curve is plot of sensitivity (y-axis) as a function of 1-specificity (x-axis)
- What comes out?
  - We can use Youdens Index = Specificity+Sensitivity-1 to select the optimal threshold point aka:

"Magic Value"

Plan

- Data:
  - 2 PhIII placebo arms within a similar patient population, PFS as main endpoint
- Build time-to-event model on a training set (one placebo arm)
  - Also use the linear predictor to develop a classifier to identify those that will progress and those that don't
- Assess models performance within the test set-data-set not used to build the model (other placebo arm)
- We assess baseline values first

Data-sets

- 2, one has a higher proportion of disease progression events in it than the other
- Both from a similar patient population

Will build a model on the red-study and test on the black

RED:

N=611

Events = 185

BLACK:

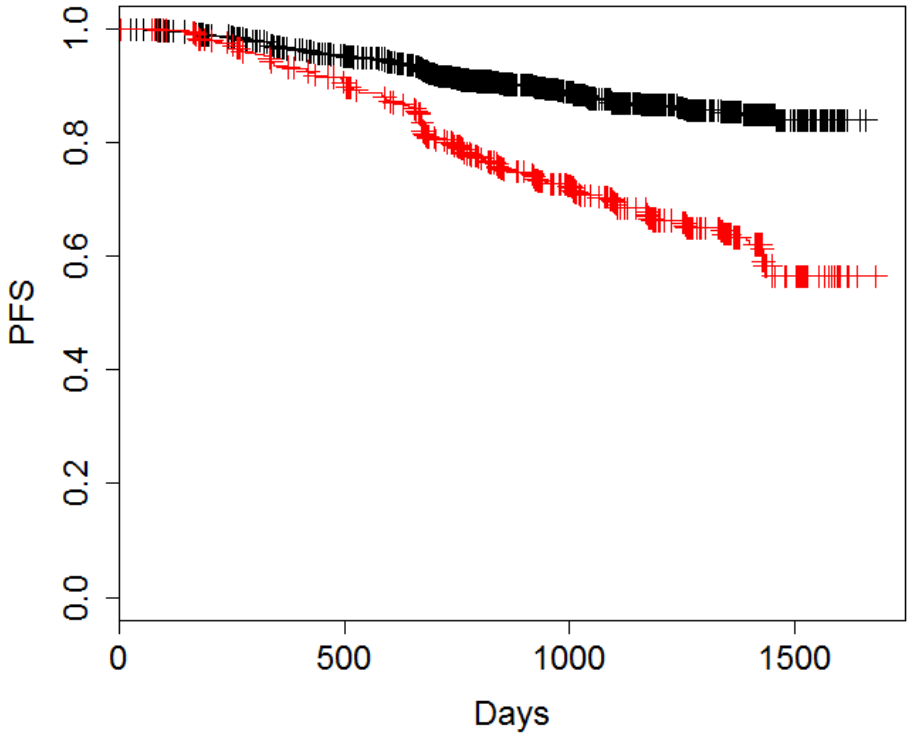
N=1803

Events = 188

Why that way round?

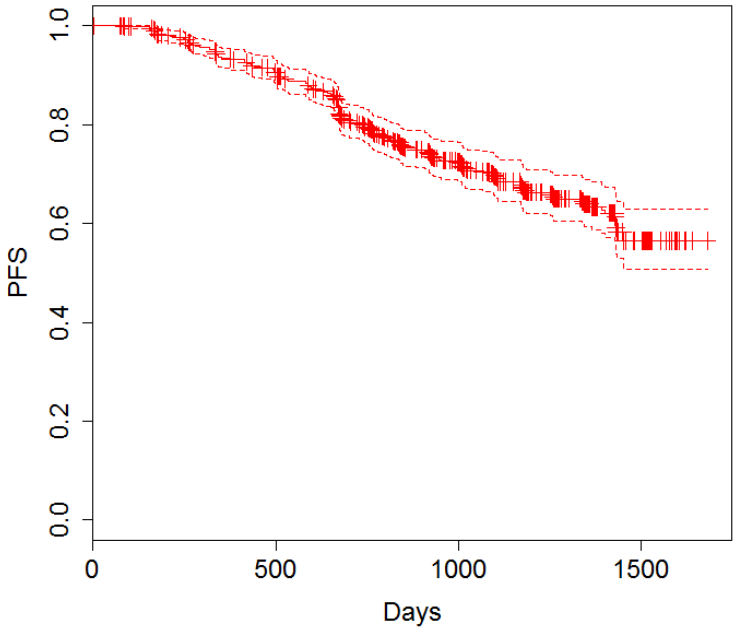
A more balanced data-set!

As a betting man I would not start with the BLACK arm!



Training Study

- Placebo arm of a PhIII study in locally invasive non-metastatic prostate cancer
  - n=611
  - 185 disease progression events
    - Increase in local tumour size
    - Presence of metastases
  - A long study
  - Covariates
    - Tumour Classification (TUMCLS)
      - How much it's spreading, scale 1:4
    - Histological Grading (HISTGD)
      - Gleason Score, describes the histological patterning
    - Age, Origin etc.
    - 10 in total



Training Study Model

- Final Model is a Weibull model with:

  Scale = a0 + a1*PSABSL

  – A boring model!

Training Study Model

- We can use the linear predictors to develop a simple classification model

$$Scale = a0 + a1*PSABSL$$

- Assess the linear predictor within a ROC analysis
  - Bootstrap to suggest a suitable threshold (median value) that maximises both sensitivity and specificity (youdens index) to take forward into test set aka "Magic Number"

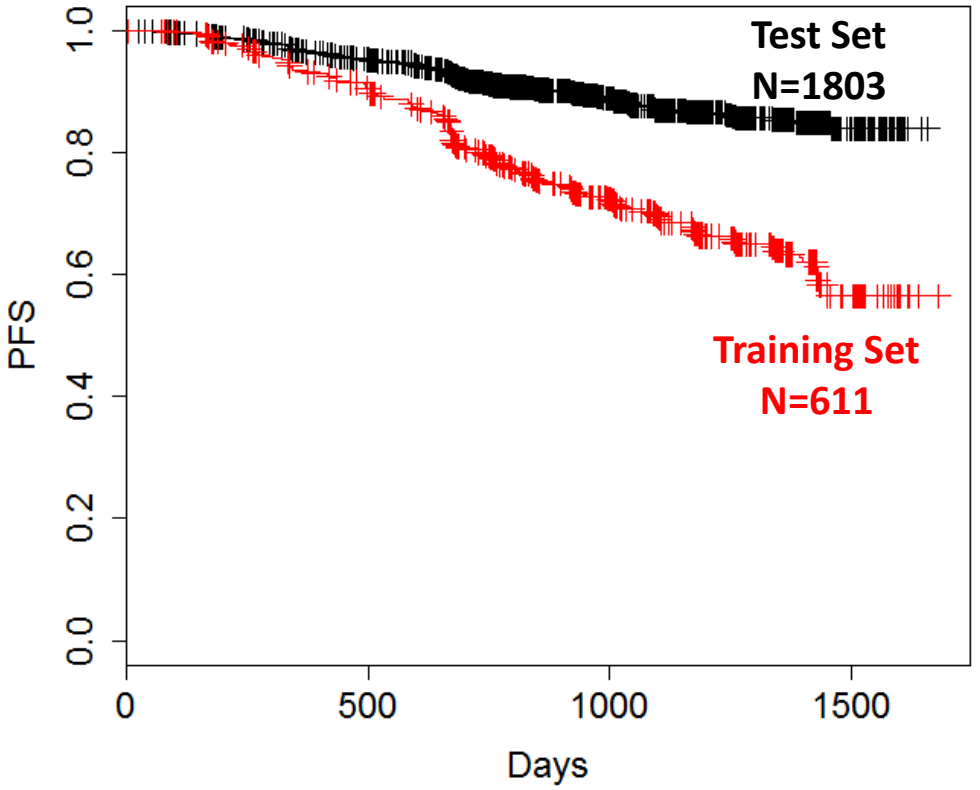
Training Set	
Sensitivity	Specificity
0.72 (0.54-0.89)	0.62 (0.41-0.77)

Test Set Study

- Placebo arm of a PhIII study in locally invasive non-metastatic prostate cancer:
  - n=1803
  - 188 events
- Test set is slightly different to the training set

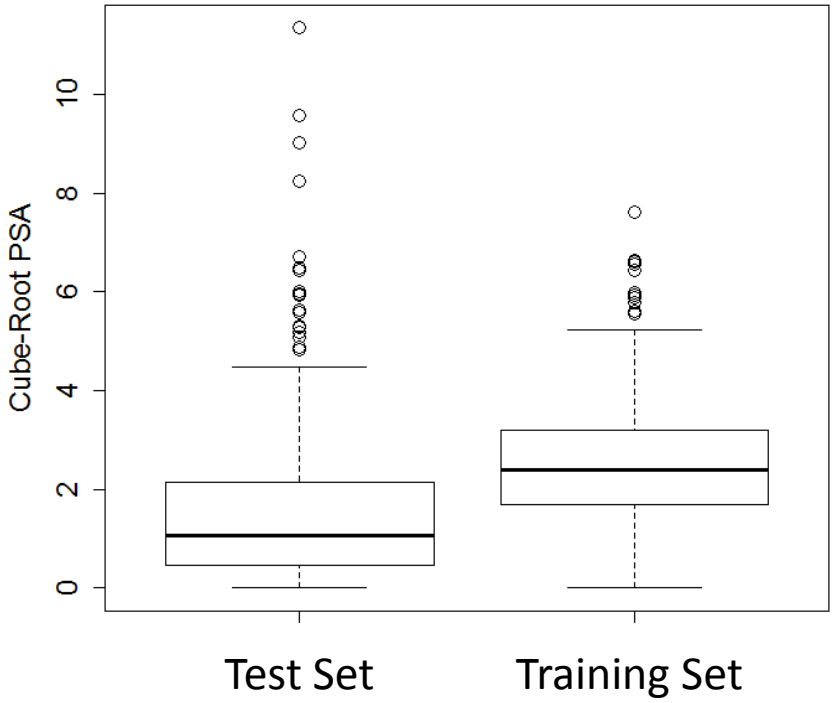
  HR = 2.8 (2.3-3.4)

- Should be a good test
  - Predict HR?
  - Classification of patients?



Test Set Study - Results

- Observed HR = 2.8
- Prediction:
    - Mean HR = 2.9
- It worked! It's hilarious! This is why:



Test Set Study - Results

"Magic" value performance

- Application of the median threshold from the training set on training set
- Scale the magic value by the change in predicted weibull scale ratio– helps to maintain classifier performance
- The results are pretty impressive!

Training Set			Test Set		
Sensitivity	Specificity	Youdens Index	Sensitivity	Specificity	Youdens Index
0.69	0.63	0.32	0.69	0.62	0.31

Negative Arm			Positive Arm		
n	Obs.	Exp.	n	Obs.	Exp.
1035	59	112.4	766	129	75.6

Other way round

- Build model on the data-set with less proportion of events and test in the one with more
  - Final model: Weibull with covariates: PSABSL, DIFF, TUMCLS and HISTGD
    - Different to the model from the other study
  - Observed HR = 2.8 (2.3-3.4)
  - Prediction HR = 3.4 (3.2-3.6)
    - Not amazing but not terrible!
  - The "Magic" number did that work…

Other way round

- Appears to work well!

Training Set			Test Set		
Sensitivity	Specificity	Youdens Index	Sensitivity	Specificity	Youdens Index
0.75	0.61	0.36	0.65	0.60	0.25

Negative Arm			Positive Arm		
n	Obs.	Exp.	n	Obs.	Exp.
317	63	105.2	285	117	74.8

Summary

- Hazard Ratio prediction were pretty good
  - Simple ideas have value
  - The final model was different with either data-set
    - Bound to be as there were very few events in one of the studies
- Performance of the classifier was good too
  - "Magic" values can work
    - Use the linear predictor of the survival model to scale the threshold – a simple idea!
    - Nothing wrong with mixing non-parametric and parametric approaches
- The setting is important
- Not looked at PSA kinetics yet

Our field likes quotes…