Prostate Cancer Active Surveillance

Author Alfonso Lam

Boutros Lab
Jonsson Comprehensive Cancer Center

June 29, 2020

UCLA Health

1 Introduction

1.1 Study Design and Objectives

In this study, a group of individuals with prostate cancer is on Active Surveillance (AS). AS is a combination of blood tests, imaging and biopsies. It intends to monitor the progression of the disease, but the current prediction of the aggressiveness of prostate cancer is not perfect and needs improvement. What features do the best job in predicting aggressive disease? In what order (sequence) should different biomarkers be run to maximize sensitivity and specificity, while minimizing cost?

1.2 Data Description

Filename	Number of Individuals	Number of Columns
MRI_DOD_Biomarkers_Database_Boutros−2020.04.20.xlsx	123	53

2 Exploratory Data Analysis

Some basic information from the individuals are available in the original dataset. Figure 1 shows some of the basic features the patients have.

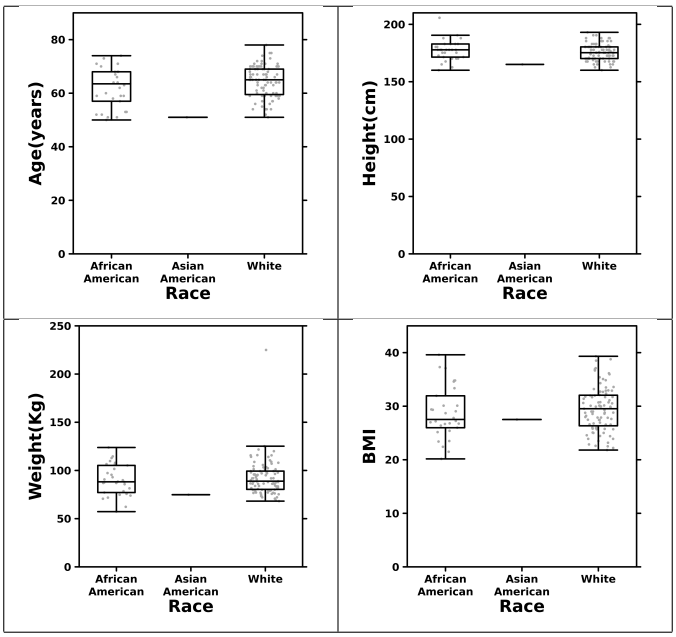

Figure 1: Basic information from individuals grouped by race: African-American, Asian-American, and White

Figure 2 shows the prostate volume distributed in the individuals. These individuals are patients that were administered blood test, urine test, MRI, biopsies, etc. They are currently on AS and not all of them did all the tests.

In the original dataset, most of those tests report numerical quantities or scores, and others reporting binary responses (0 or 1, etc). As part of the exploratory data analysis, Table 1 summarizes the data in the tests analyzed in this report.

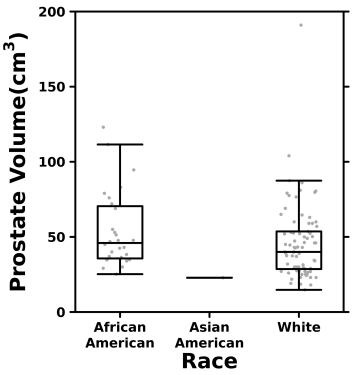

Figure 2: Prostate volume for the three different groups of subjects.

Test	Minimum	First Quarter	Median	Mean	Third Quarter	Maximum
ADC normal Signal	3.00	17.00	32.50	45.57	50.75	264.00
ADC lesion Signal	0.00	1.00	18.50	77.09	64.00	919.00
RSI lesion Signal	16.00	34.25	55.00	53.93	71.75	94.00
RSI normal Signal	6.00	15.00	31.00	33.15	45.50	77.00
PCA3	1.20	4.05	6.25	6.96	8.67	30.40
T2ERG	0.17	0.52	0.79	0.93	1.10	3.87
MiPS Cancer Risk	1.73	7.73	12.84	29.05	19.09	1561.00
MiPS High Grade Cancer Risk	3.20	9.70	13.65	14.96	17.80	41.30
PSA Hybrid	5.40	27.60	36.55	41.63	54.38	101.20
free PSA	1.00	1.00	1.00	1.55	2.00	4.00
PSA Density	0.003952	0.011972	0.018269	0.021220	0.025472	0.121317
p2PSA	1.00	8.00	24.00	33.79	52.00	98.00
Percent Free PSA	1.00	1.00	2.00	1.67	2.00	3.00
PHI	0.00	0.00	0.00	0.17	0.00	1.00
PHI Density	0.0750	0.5371	0.8927	1.0839	1.4006	3.8899
SOCPSA	0.00	0.00	0.00	0.19	0.00	2.00
TNFa Average	0.00	0.00	0.00	0.12	0.00	1.00
Genetic Ancestry	0.00	0.00	0.00	0.36	0.00	5.00
Genetic Risk Score	0.00	0.00	0.00	0.07	0.00	4.00
Genetic Risk Category	1.07	2.16	2.89	2.76	3.24	4.58
Global Screening Array	23.40	833.11	1031.24	994.27	1189.62	1736.73
GSA Positives	-1.53	495.67	637.40	650.40	780.48	1202.47
BRCA Mutation	3.60	29.05	42.04	54.43	54.77	848.73
Mutation 1	1.50	8.80	16.98	17.23	23.57	41.19
Mutation 2	1.18	3.77	5.56	6.89	8.35	60.00

Table 1: Descriptive statistics of the tests administered to the subjects.

Some of these tests are related to other tests or are combination of simpler tests. To quantify these relations, Spearman's correlation between tests and subjects features are computed. The result is shown in Figure 3. The information from the heatmap will be useful to reduce the number of tests used for the sequential testing.

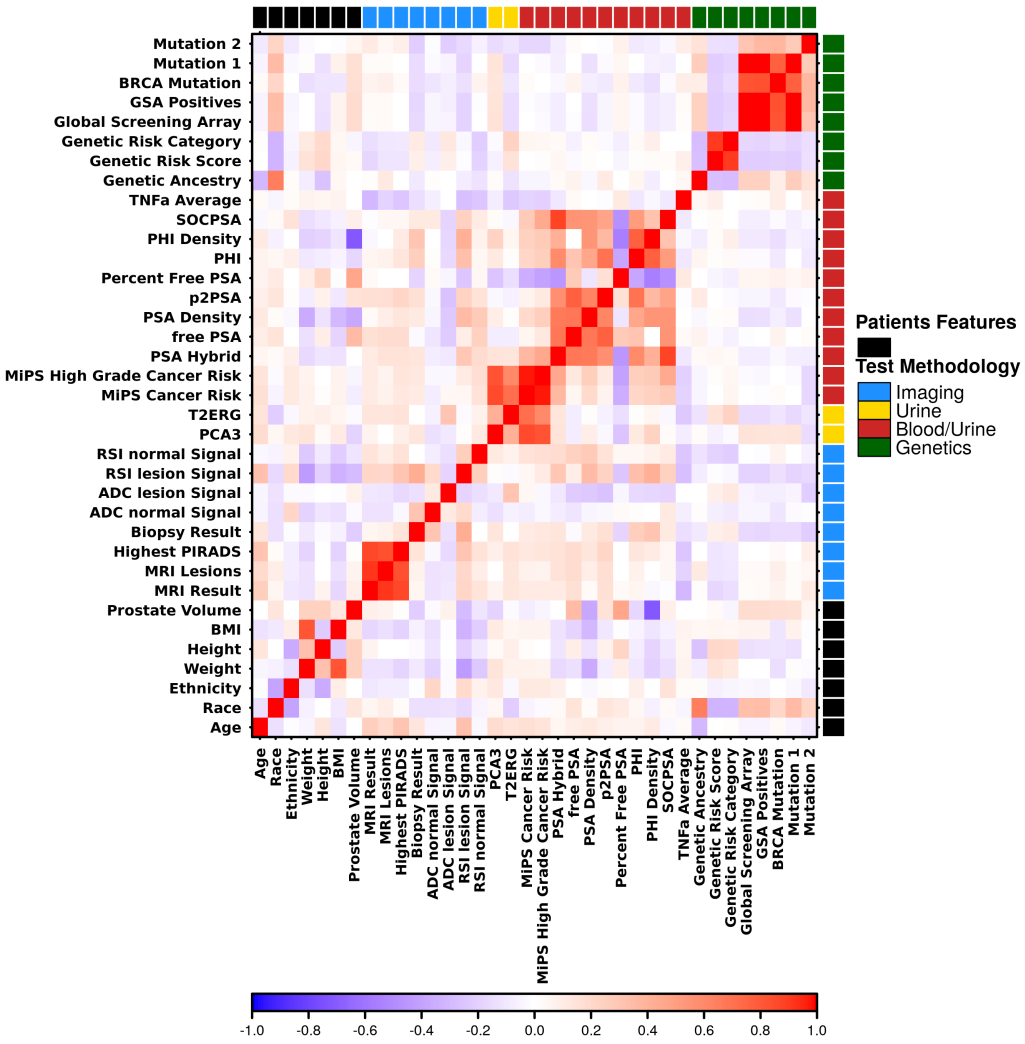

Figure 3: Correlations between features of the subjects and tests performed in the subjects.

By definition, a patient on AS has prostate cancer. Thus, a sequence of tests are performed during the AS in order to monitor the progression of the disease. At some point in time, the doctor updates the biopsy information (`BiopsyUpgraded`). It is a categorical variable (0=No, 1=Yes) and will be used as the gold standard to determine a relation between the `BiopsyUpgraded` with the different biomarkers used in the AS. Those biomarkers are shown in the Appendix section.

For each test, its performance is evaluated by plotting the Precision-Recall (PR) Curve. The PR curves for all the tests are displayed in Figure 4. From that figure, most of the test have in general low precision, except `RSIlesionSignal` where particularly shows a larger area under the curve in comparison to other tests. Figure 5 shows the ROC curves of the tests. Several of those tests are on the diagonal line or even below, which indicates they have low or no predictive power. Figure 6 summarizes the areas under the curve for the PR-curves and ROC-curves.

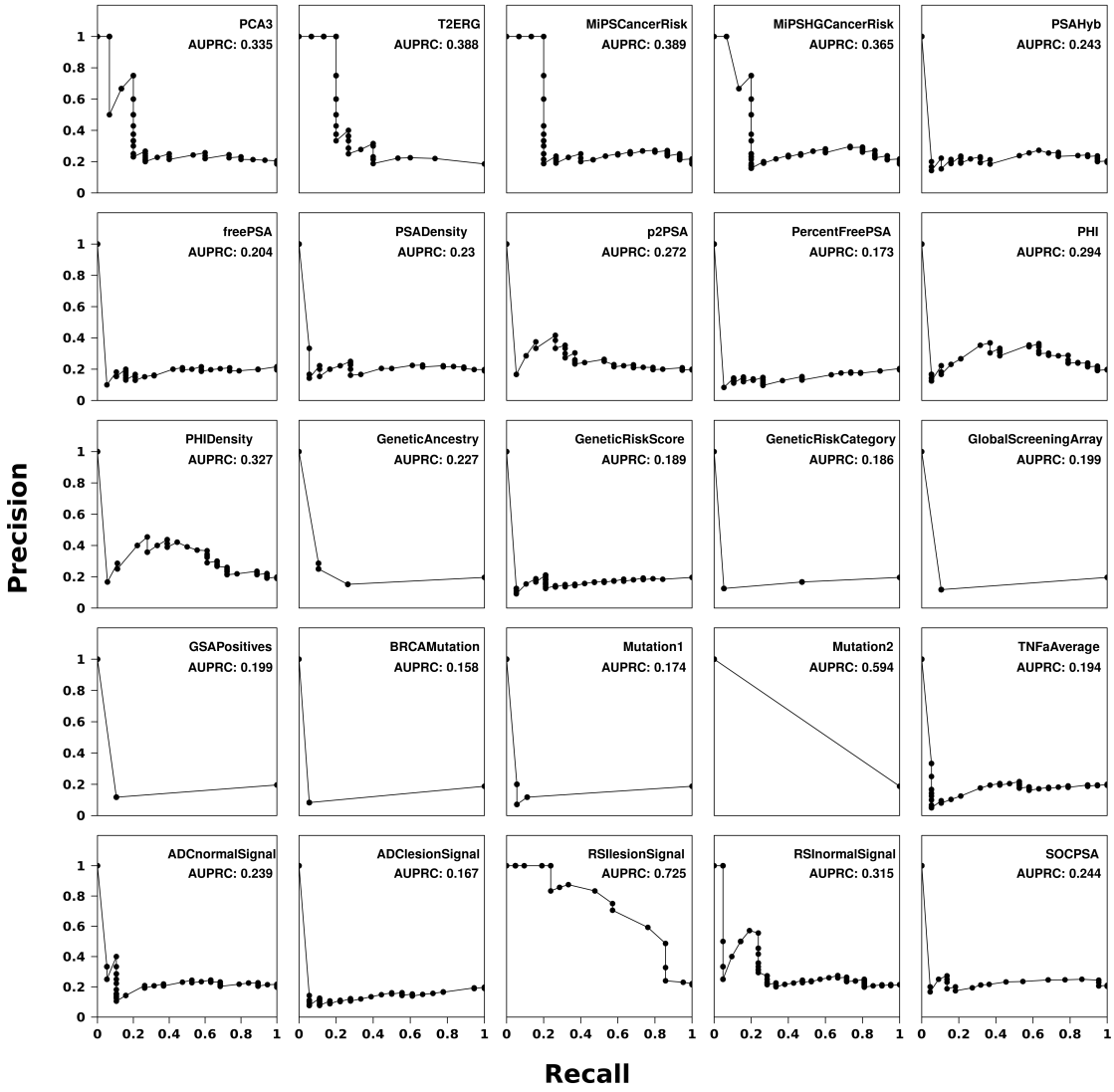

Figure 4: Precision-Recall Curves of the tests used in AS.

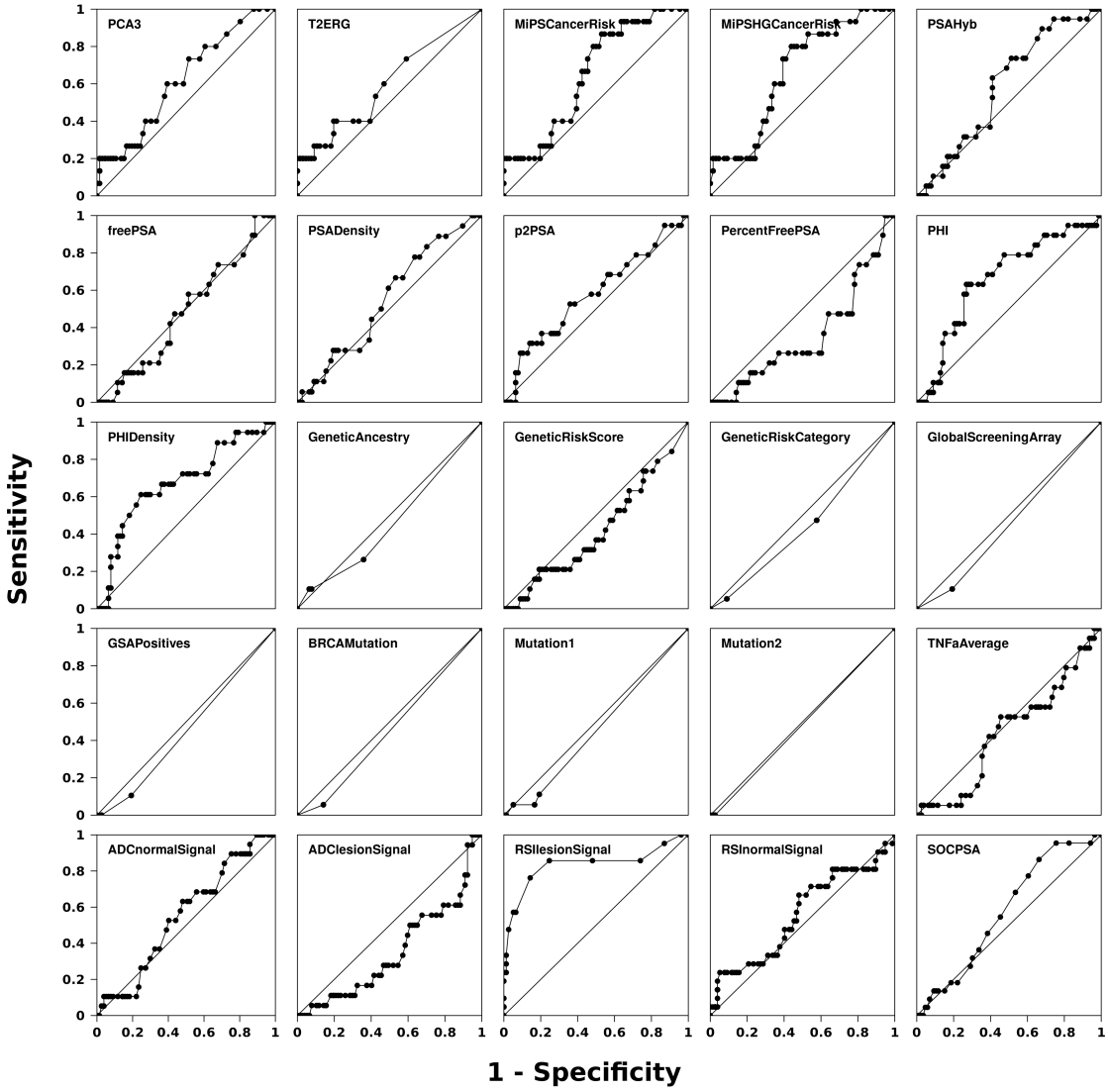

Figure 5: ROC Curves of the tests used in AS.

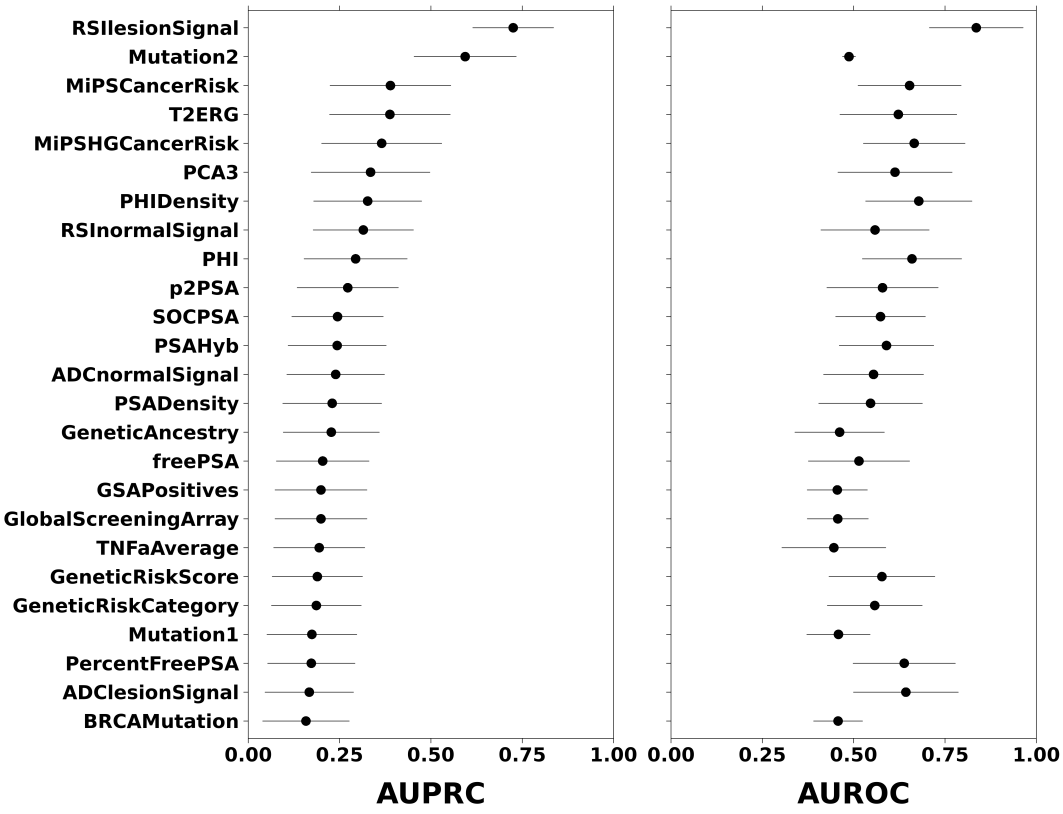

Figure 6: Forest plot for AUPRC and AUROC Curves of the tests used in AS.

3 Sequential Testing

Sequential testing was performed to determine under what condition a patient that did a test A needs to do test B. Patients on AS do a set of tests over a period of time to monitor the progression of their cancer. The idea behind is to find an optimal sequence of tests such that the overall sensitivity and specificity are maximized at lower cost, and also reduce the number of tests for the entire procedure.

For this sequential testing, leave-one-out cross-validation was performed in each test, their optimal thresholds were determined (See Methods) to build the confusion matrix, and subsequently all variables such F_1-scores, sensitivity and specificity were reported in Table 2.

Test	Threshold	Sensitivity	Specificity
RSI lesion Signal	57	0.73	0.92
PHI	45	0.58	0.71
p2PSA	13	0.50	0.58
MiPS High Grade Cancer Risk	30	0.80	0.55
PCA3	27	0.73	0.48
MiPS Cancer Risk	47	0.87	0.45
RSI normal Signal	16	0.73	0.45
Genetic Ancestry	3.00	0.92	0.29
PSA Hybrid	4.00	1.00	0.27

SOCPSA	4.00	1.00	0.25
ADC normal Signal	829	1.00	0.23
T2ERG	1.00	0.93	0.21
freePSA	0.35	1.00	0.10
Genetic Risk Score	3.0	1.00	0.10
ADC lesion Signal	323	0.92	0.03
Percent Free PSA	6.2	1.00	0.03
TNFa Average	1.41	1.00	0.03
BRCA Mutation	0.00	1.00	0.00
Genetic Risk Category	1.00	1.00	0.00
Global Screening Array	0.00	1.00	0.00
GSA Positives	0.00	1.00	0.00
Mutation 1	0.00	1.00	0.00
Mutation 2	0.00	1.00	0.00

Table 2: Optimal threshold for each test determined by leave-one-out cross-validation

To make a sequential testing, the process can be illustrated as follows: A group of N subjects did test A. Using the threshold determined by the cross-validation, a portion of these subjects, N-X, is classified as positive (true positive and false negative) and X subjects are considered negative (true negative and false positive). Thus, the positive cases (true and false) will go to the next test.

As previously mentioned, some of those tests performed to the subjects have strong correlation with other tests (See Figure 3). That will be useful to reduce the number of tests included in the sequential testing. Health care provider could choose a test based on pre-defined criteria such as specificity, cost, invasive/non-invasive, easy to administer, etc.

The simplest sequential testing consists of a pair of tests. Using the outputs from the cross-validation analysis, all the tests were paired each other, and their respective overall sensitivity/specificity were evaluated. The top biomarkers with high sensitivity and specificity are: `RSIlesionSignal`, `PHI`, `p2PSA`, `MiPSHGCancerRisk`, `PCA3`, `MiPSCancerRisk`, and `RSInormalSignal`. The results for all of the pairs of tests are shown as heatmaps in Figure 7. The heatmaps were created based on the following criteria: 1) To make a fair comparison, only subject who did all the tests administered in this study were considered. 2) For the overall sensitivity/specificity, Equations 1 and 2 were used.

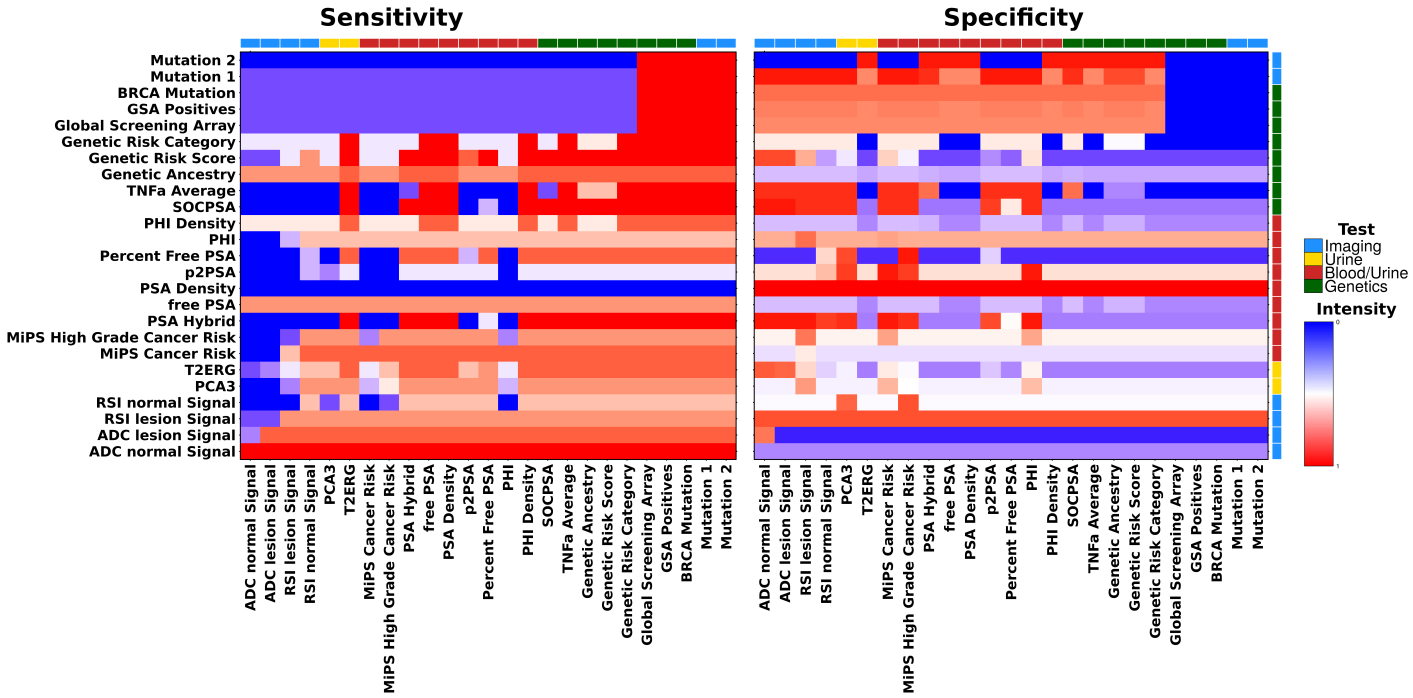

Figure 7: Overall Sensitivity/Specificity for Pairs of Tests. Notice that the order of the tests changes the outcome

Tests `PCA3`, `MiPSCancerRisk`, and `MiPSHGCancerRisk` are highly correlated each other and therefore, for sequential testing purposes, one can be chosen. Based on the principles from sequential testing (See Methods), `MiPSHGCancerRisk` would be the chosen test. Also, it is important to notice that this test is a combination of multiple tests which includes `PCA3`.

Thus, as a first attempt to make a sequential testing could be as follows:

- RSIlesionSignal
- PHI
- p2PSA
- MiPSHGCancerRisk
- RSInormalSignal

This is a configuration above is a set of imaging/blood/urine tests. To visualize the sequential testing for this set, the tree in Figure 8 highlights the process. At the topf of the tree, the confusion matrix for `RSIlesionSignal` is computed. The positive outcomes from `RSIlesionSignal` (TP =10 and FP=5), will go to the next test, `PHI`. It determines TP=7 and FP=2. And this procedure continues until the TP and TN cannnot be reduced. At that moment, the procedure stops and `RSINormalSignal` was not necessary to implement.

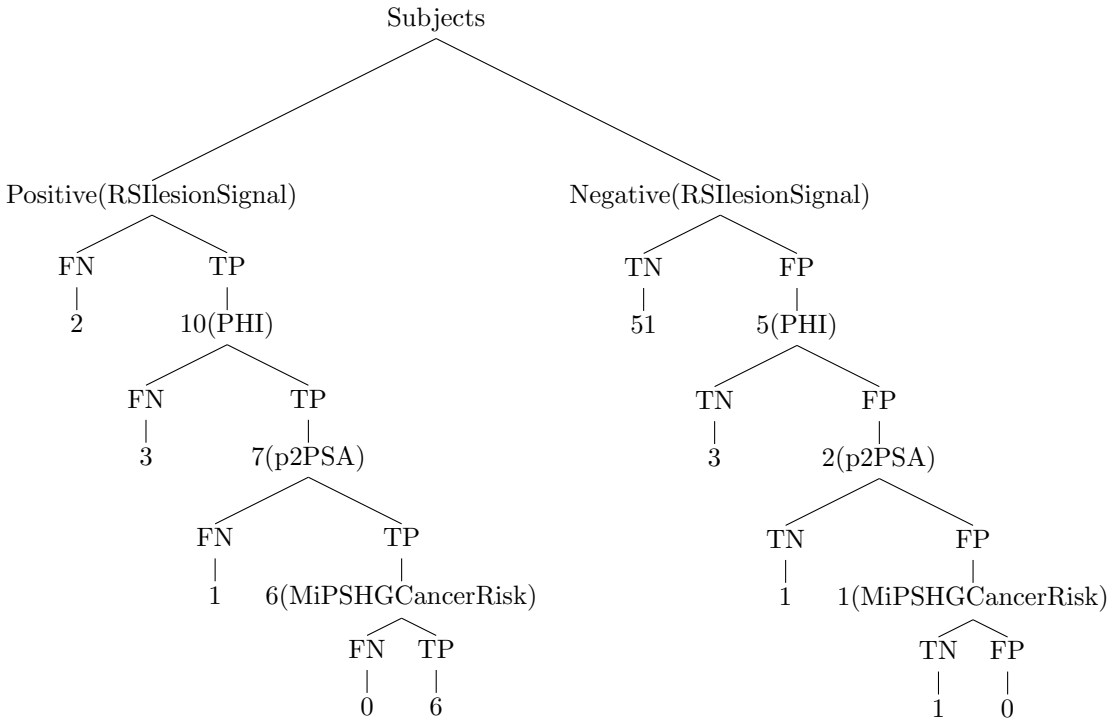

Figure 8: RSIlesionSignal-PHI-p2PSA-MiPSHGCancerRisk

A relevant detail is that `PHI` test is a combination of three PSA forms: total PSA, free PSA, and p2PSA. In that case, this first approach could be reduced even further, and a simplified tree is shown in Figure 9. By reducing the number of tests, it helps reduce cost and anxiety/stress from the patient.

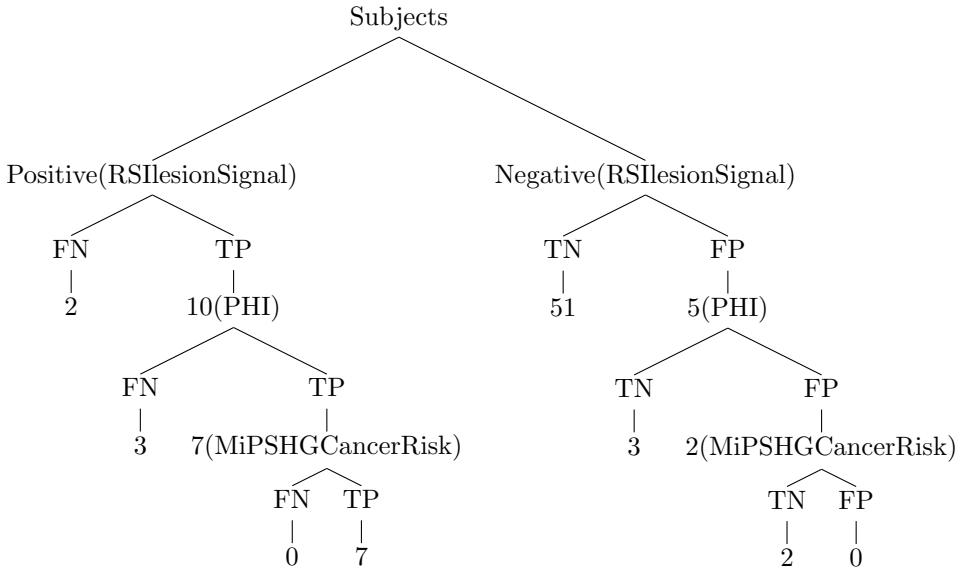

Figure 9: Sequential test after removing p2PSA: RSIlesionSignal-PHI-MiPSHGCancerRisk. One step is saved with this sequence with better overall sensitivity.

4 Survival Analysis

In this section, survival analysis was performed. From the original data, the time field `DaysDxToUpgrade` was initially considered. It contains data from patients whose `BiopsyUpgraded` have a value of 1. The remaining empty rows (patients with `BiopsyUpgraded` equals to 0) were replaced with `DaysBxToLastReview`. Considering the top biomarkers used described in the previous section, the Kaplan-Meier survival plots are shown below:

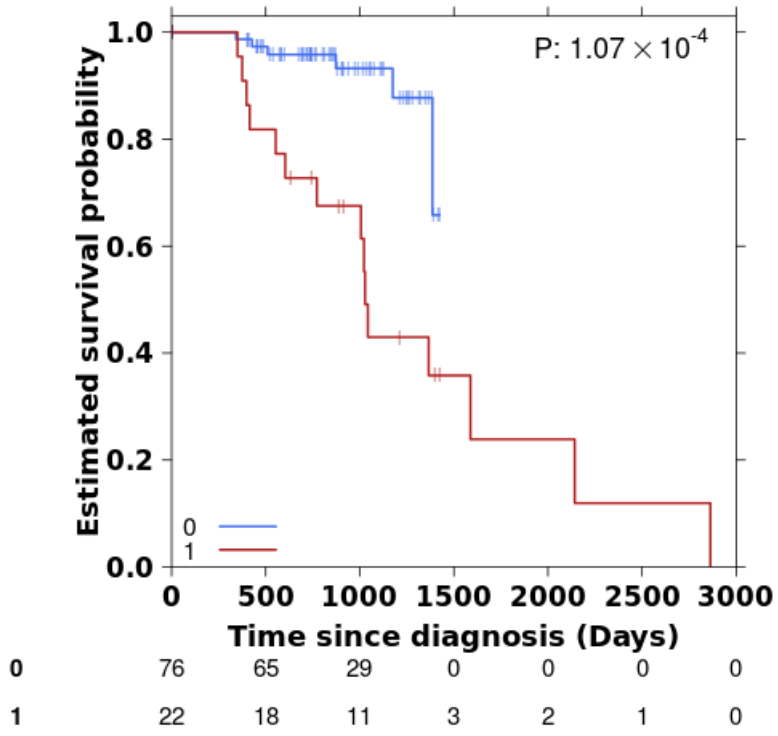

Figure 10: Kaplan-Meier Plot for RSI lesion Signal

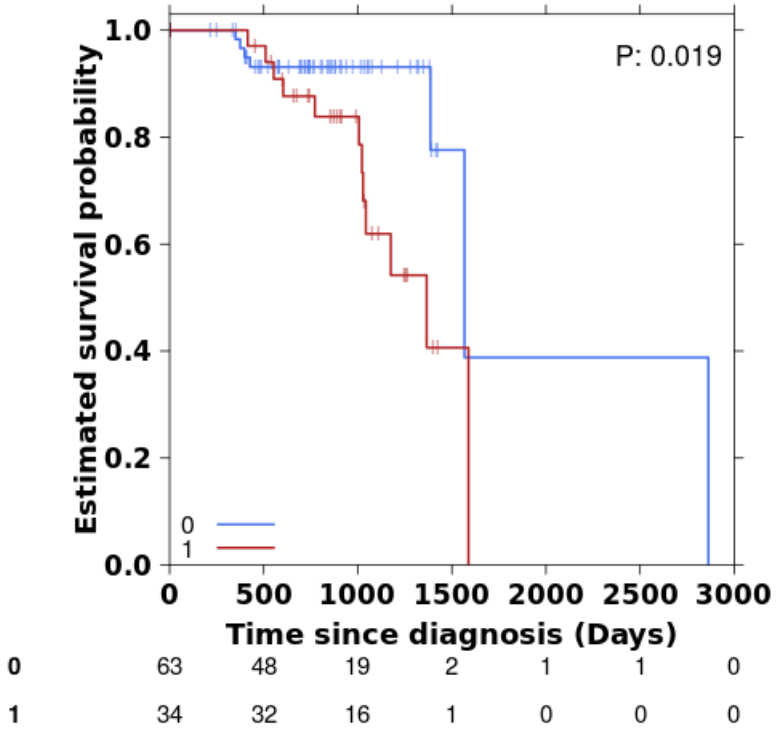

Figure 11: Kaplan-Meier Plot for PHI

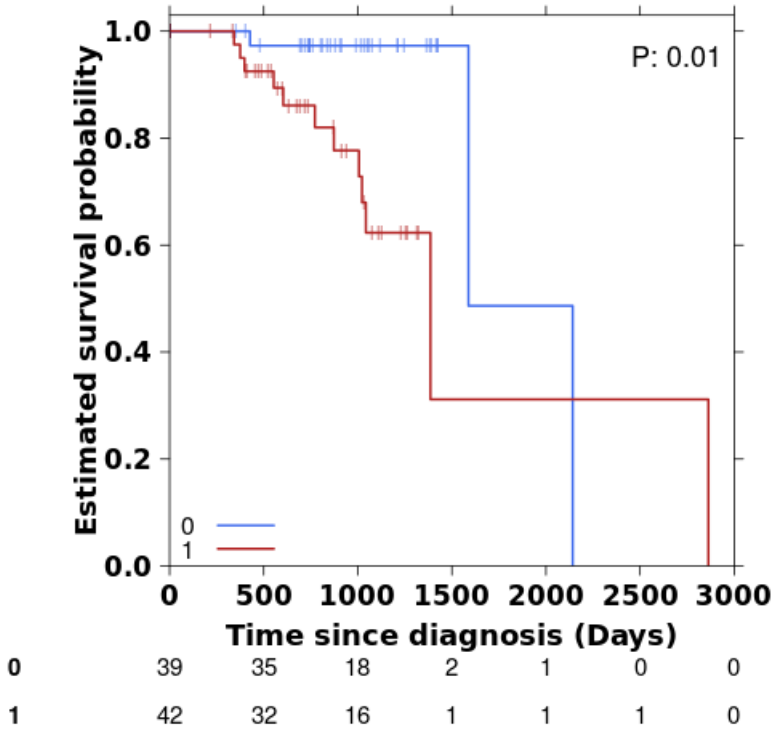

Figure 12: Kaplan-Meier Plot for MiPS High Grade Cancer Risk

5 Next Steps

Systematic descriptive analysis of cohorts.

6 Methods

6.1 Precision-Recall (PR) Curves

Recall or Sensitivity is the ability of the test to correctly mark all positive responses as positive. Precision is the ability of the test not to wrongly label a negative sample as positive. To make the PR curve of a test, the confusion matrix is computed based on the scores of the test and the gold standard data (in this study, `BiopsyUpgraded`). Each point of the PR curve (Recall, Precision) comes from a confusion matrix at a given threshold. The threshold is applied to the scores of the test to determine which scores are predicted as positive (1) and negative (0). Using these predicted values, the elements of the confusion matrix are computed and other variables such as accuracy, precision, sensitivity and specificity, F_1-score can be easily obtained. The area under the Precision-Recall curve (AUPRC) can be computed using the trapezoid rule. A similar procedure is employed to make the ROC curve.

6.2 Leave-One-Out Cross-Validation

Leave-One-Out Cross-Validation is a method that trains a model by taking one point out of the dataset as a test data and the remaining points are used to train the model. Once the optimal model is found, the model is evaluated using the test data. And this process is done repetively such that all points will be used as test data. In this study, the model is the confusion matrix with the optimal threshold. The criterion to determine that threshold is based on F_1-score.

6.3 Sequential Tests

Serial or sequential testing is useful in some clinical situations to potentially avoid the need for a subsequent test. For the case of two tests, A and B, the computation of the sensitivity and specificity are given by:

$$\text{Sensitivity(A,B)} = \text{Sensitivity}_A.\text{Sensitivity}_B \tag{1}$$

$$\text{Specificity(A,B)} = \text{Specificity}_A + [1 - \text{Specificity}_A].\text{Specificity}_B \tag{2}$$

From these equations, it is clear that more tests performed, overall sensitivity will decline and overall specificity will increase. Sequential testing looks for an increase in the positive predictive value, *i.e* , determine that a subject with a condition truly has the condition. Thus, in each stage the test has to report positive. Otherwise, the procedure stops.

Based on Equation 2, a good guess to make a sequential testing is to choose a test with a high sensitivity. That makes the first term dominant. The reason behind to choose a test by specificity is that the test with high specificity will reduce the group of subjects to go to the next test by excluding the negative cases from the group. This can be seen from the specificity formula: $tn/(tn + fp)$. The maximum value is reached when $fp = 0$. Thus, the list of positive subjects to go to the next test will be smaller.

7 Appendix

Current spreadsheet (04/20/2020) has the following columns:

Column Name	Description
BiopsyUpgraded	Research study biopsy increased cancer grade from most recent previous biopsy result
PCA3	Urinary biomarker
T2ERG	Urinary biomarker
MiPSCancerRisk	Urinary/Blood biomarker using PSA, PCA3, T2ERG (Mi-Prostate Score Cancer Risk)
MiPSHighGradeCancerRisk	Urinary/Blood biomarker using PSA, PCA3, T2ERG (Mi-Prostate Score High Grade Cancer Risk, Gleason Score 7)
PSAHyb	Blood biomarker (Access Hybritech PSA assay)
freePSA	Blood biomarker (amount of unbound PSA)
p2PSA	Blood biomarker ([-2]proPSA, precursor of PSA)
PercentFreePSA	Blood biomarker (percentage of unbound PSA)
PHI	Blood biomarker (Prostate Health Index)
GeneticAncestry	Genetic biomarker (Scores: 1:European, 2:African, 3:East Asian, 4:Native American)
GeneticRiskScore	Genetic biomarker (Scores: 1:Low, 2:Normal, 3:High)
GeneticRiskCategory	Genetic biomarker (Scores: 1, 2 and 3)
GlobalScreeningArray	Genetic biomarker (Scores: 0 and 1)
GSAPositives	Genetic biomarker (Scores: 0 and 1)
BRCAMutation	Genetic biomarker (BRCA1 or BRCA2 gene mutation) (0: Negative, 1:Positive)
Mutation1	Genetic biomarker (if Global Screening Array positive: 0:Negative 1:BRCA1, 2:BRCA2, 3:ATM, 4:MLH1, 5:PMS2)
Mutation 2	Genetic biomarker (if Global Screening Array positive: 0:Negative 1:BRCA1, 2:BRCA2, 3:ATM, 4:MLH1, 5:PMS2)
RSInormalSignal	Imaging biomarker (Continuous values)
RSIlesionSignal	Imaging biomarker (Continuous values)
ADCnormalSignal	Imaging biomarker (Continuous values)
ADClesionSignal	Imaging biomarker (Continuous values)
TNF average	
SOCPSA	

Table 3: Test used fo this study

8 References

[1] James, G. *et al.*. An Introduction to Statistical Learning. Springer. 7th Printing.

[2] C. P'ng. *et al.*. BPG: Seamless, Automated and Interactive Visualization of Scientific Data. *BMC Bioinformatics* 20, 42 (2019).