

Tumor Suppressor p53 Gene in a Controversial Breast Cancer Population – Development of Methodology and Analysis of Mutations via Bayesian Network Algorithms

Robin, H. (a), Truong, S. (b), Patten, P. (b), Wu, L. (b), Gallarda, J. (b), Eberhardt, J. (c), Clark R. (a), Muller W. (a), Sadeghi S. (a), Kam J. (a)

(a) Sharp Memorial Hospital, San Diego, CA; (b) Roche Molecular Systems, Pleasanton, CA; (c) DecisionQ Corporation, Kentfield, CA

Introduction

- The p53 gene is widely implicated in breast cancer pathogenesis and may have potential for therapeutic decision making (1-4)
- Aims:
 - Compare p53 mutation status with histologic findings and treatment outcomes,
 - Evaluate prognostic significance of p53 mutations
 - Determine role of p53 mutation testing for small, HER2 negative breast cancers
- Breast cancer specimens evaluated for standard predictive and prognostic biomarkers
- Treatment and outcomes data evaluated
- Tumor specimens characterized for p53 mutations with GeneChip technology
- DecisionQ FasterAnalytics™ machine learning Bayesian networks generated a predictive outcomes model

Methods – Data Collection and Analysis

- 64 T1 (under 2 cm), HER2 negative breast cancer specimens collected at Sharp Memorial Hospital in San Diego, CA
- Specimens analyzed for prognostic and predictive biomarkers using Ventana Benchmark and the Applied Imaging Ariol 50 automated imaging platforms (See Table 1)
- Treatment and outcomes data collected from patient records with median follow up 3.5 years
- p53 gene mutations were identified using GeneChip technology
- Region 550-580 mutations were censored, GC rich region reduces reliability of analysis
- Data analyzed with Bayesian Belief Network encoding probabilities of all variables in domain
- Bayesian Networks constructed using DecisionQ FasterAnalytics™, a machine learning program that uses a set of heuristics to generate hypothetical models (5)
- Step-wise modeling process was used to reduce analogs and confounding variables
- Network validated using a train-and-test cross-validation methodology
 - Data set randomized seven times into a 95% training set and 5% test set
 - Output of the test sets was used to produce ROC curves and predictive values

Methods – GeneChip Analysis

- One 10m section was used for DNA extraction. The tissue samples were heated at 98°C in 200L of Lysis Buffer (10 mM Tris pH 8.6, 50 mM KCl, 0.1% Triton X-100, 1 mM EDTA) for 30 min. 20L of Proteinase K Solution (2 mg/mL Proteinase K, 10 mM Tris pH8.6) were added after the Lysis Buffer cooled down.
- The tissue samples were then incubated at 68°C for 60 min, followed by heat denaturation at 95°C for 10 min. The crude lysate was used for PCR at 10L per reaction. Six PCR reactions were performed for each sample. One was used for p53 GeneChip PCR Amplification of Exon 2-11 of human p53 gene, and the others were used to amplify Exon 4, 5, 8, 10 or 11 alone. All six PCR reactions were performed in a solution of 10 mM Tris-HCl pH8.3, 50 mM KCl, 2.5 mM MgCl₂, 200M dATP, 200M dCTP, 200M dGTP, 200M dTTP, 10 units of AmpliTaq-Gold, and 10 pairs of primers for Exons 2-11 or 1 pair of primers (Exon 4, 5, 8, 10 or 11). The 40L of amplicon from the Exons 2-11 PCR reaction was pooled with 10L of amplicon from each of Exon 4, 5, 8, 10 and 11 for one p53 GeneChip Assay (Affymetrix).
- The post-PCR p53 GeneChip assay was performed according to the manufacture's instruction. A total of 90L amplicon were fragmented with Dnase I, labeled with Fluorescein-ddATP by way of a terminal transferase reaction and hybridized to a p53 GeneChip Array. Fluorescently labeled fragmented DNA samples were washed and allowed to bind to complementary oligonucleotide probes.
- Hybridized probe arrays were then scanned by the GeneArray Scanner (HP G2500).
- The intensity data were analyzed with the p53 GeneChip Mixture Detection Algorithm. The mutations were detected with a GeneChip Score equal to or greater than 13 (Ref).

The AmpliChip combines the gold standards in PCR and microarray technology

AmpliChip p53 Workflow

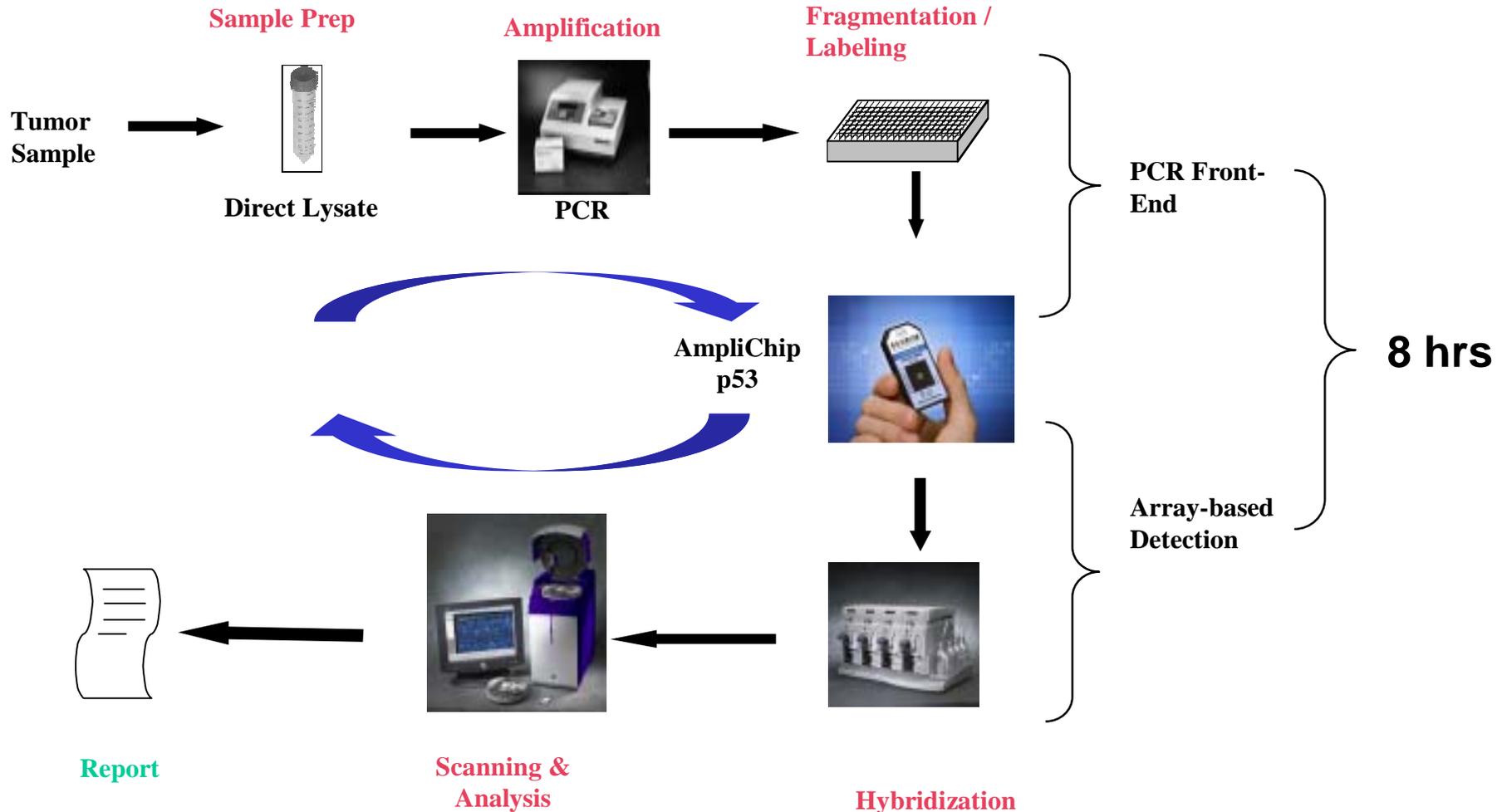


Table 1 – Data Elements and Preparation

<u>Variable</u>	<u>Description</u>	<u>Classifications</u>
Tumor Type	Tumor classification	Infiltrating ductal, Infiltrating lobular, Other
Differentiation	Differentiation of tumor cells	Well, Moderate, Poor
Estrogen/ Progesterone	Expression of Estrogen and Progesterone receptors	Positive, Negative
Ploidy	Chromosomal ploidy	Diploid, Not Diploid
p53 IHC	Expression of p53 protein in the nucleus	Negative, Positive
Ki67 Proliferative Index	Measure of tumor cell proliferation	2 bins
HER2 Receptor	Expression of HER2 receptor measured by IHC	0, 1+
S-Phase%	Percentage of cells in synthesis phase	2 bins
Mortality	Mortality of patient during follow-up period	Alive, Deceased
Chemo	Chemo administered	Yes or No
Exon 4, 5, 6, 7, 8	Exon on which mutation found	Yes or No
Disease Recurrence	Recurrence within follow-up period	Yes or No

- Variables included for modeling but eliminated in the step-wise modeling process: age, tumor size, lymph node involvement, combined Nottingham histologic grade, hormone and radiation therapy, nuclear grade, time to progression, and multiple tumors

Table 2 – Distribution of Markers in Cohort by Disease Recurrence

	Recurrent Disease					
	No		Yes		<i>Recurrence Rate</i>	
	<i>n</i>	<i>% of Total</i>	<i>n</i>	<i>% of Total</i>		
ER/PR Negative	15	27.3%	3	42.9%	16.7%	
ER/PR Positive	40	72.7%	4	57.1%	9.1%	
Ki67 > 7	19	45.2%	4	80.0%	17.4%	
Ki67 < 7	23	54.8%	1	20.0%	4.2%	
S-Phase% > 5	18	45.0%	4	80.0%	18.2%	
S-Phase% < 5	22	55.0%	1	20.0%	4.3%	
p53 IHC Negative	17	37.0%	3	50.0%	15.0%	
p53 IHC Positive	29	63.0%	3	50.0%	9.4%	
Nottingham 1	28	52.8%	2	28.6%	6.7%	
Nottingham 2	22	41.5%	3	42.9%	12.0%	
Nottingham 3	3	5.7%	2	28.6%	40.0%	

- Specimens with high S-Phase %, Ki67, Combined Nottingham Histologic Grade, and Estrogen/Progesterone Negative have increased risk for recurrence of disease

Table 3 – Distribution of p53 Mutations in Cohort by Disease Recurrence

	Recurrent Disease				
	No		Yes		<i>Recurrence Rate</i>
	<i>n</i>	<i>% of Total</i>	<i>n</i>	<i>% of Total</i>	
Exon 4 Mutation	3	6.8%	0	0.0%	0.0%
Exon 4 WT	41	93.2%	6	100.0%	12.8%
Exon 5 Mutation	2	5.3%	3	50.0%	60.0%
Exon 5 WT	36	94.7%	3	50.0%	7.7%
Exon 6 Mutation	1	2.0%	1	14.3%	50.0%
Exon 6 WT	48	98.0%	6	85.7%	11.1%
Exon 7 Mutation	3	6.0%	0	0.0%	0.0%
Exon 7 WT	47	94.0%	7	100.0%	13.0%
Exon 8 Mutation	9	20.5%	1	16.7%	10.0%
Exon 8 WT	35	79.5%	5	83.3%	12.5%
Exon 10 Mutation	0	0.0%	0	0.0%	NA
Exon 10 WT	44	100.0%	6	100.0%	12.0%
Exon 11 Mutation	0	0.0%	0	0.0%	NA
Exon 11 WT	46	100.0%	6	100.0%	11.5%

- The incidence of disease recurrence by p53 mutation demonstrates that mutations of Exon 5 and 6 have a significantly higher incidence of recurrence

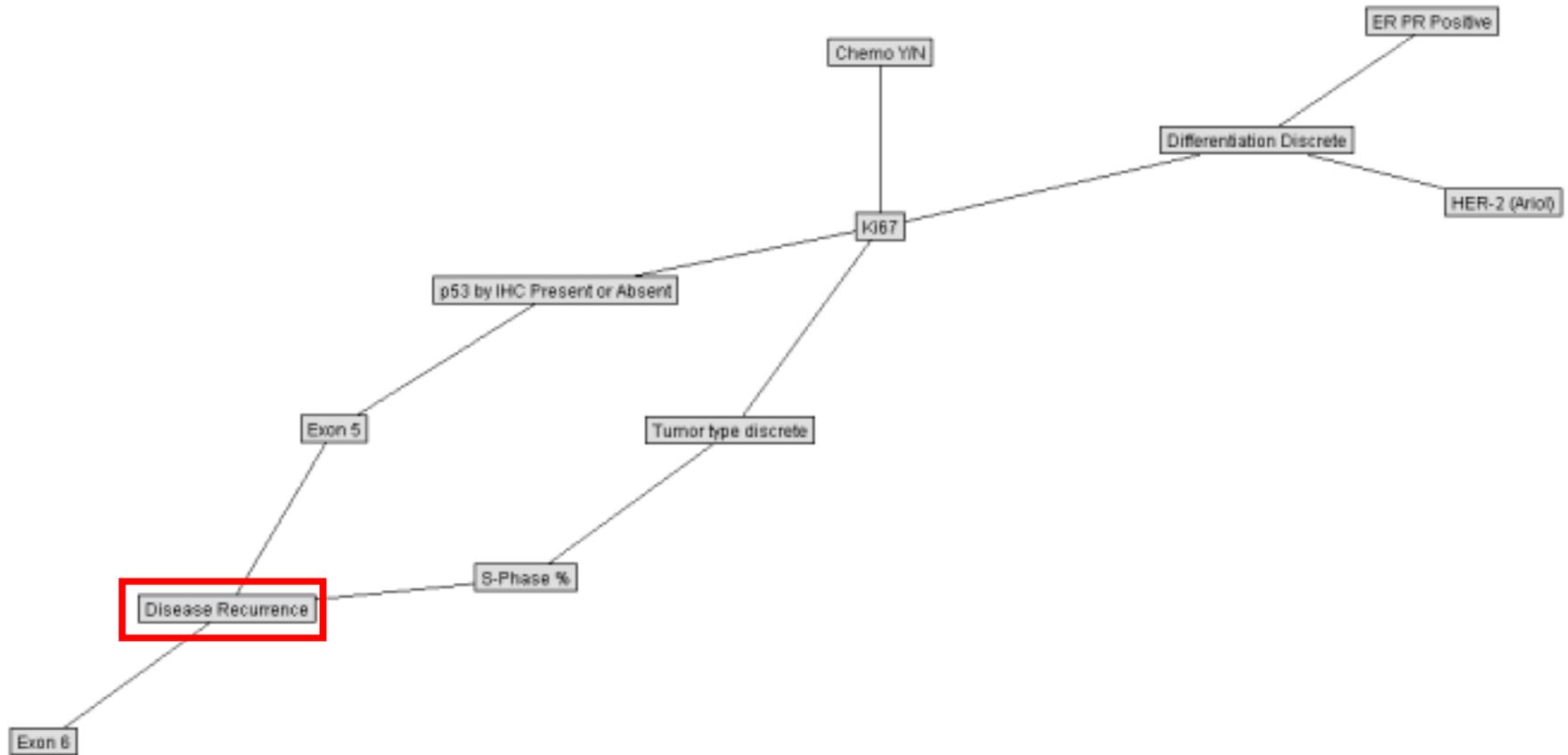
Table 4 – p53 by IHC and mutations on Exons 5 and 6

	p53 IHC				
	Negative		Positive		<u>% p53 IHC Present</u>
	<u>n</u>	<u>%</u>	<u>n</u>	<u>%</u>	
Any Mutation	8	50.0%	8	27.6%	50.0%
All WT	8	50.0%	21	72.4%	72.4%

	p53 IHC				
	Negative		Positive		<u>% p53 IHC Present</u>
	<u>n</u>	<u>%</u>	<u>n</u>	<u>%</u>	
Exon 5 Mutation	4	26.7%	0	0.0%	0.0%
Exon 5 WT	11	73.3%	24	100.0%	68.6%
Exon 6 Mutation	0	0.0%	2	6.5%	100.0%
Exon 6 WT	17	100.0%	29	93.5%	63.0%

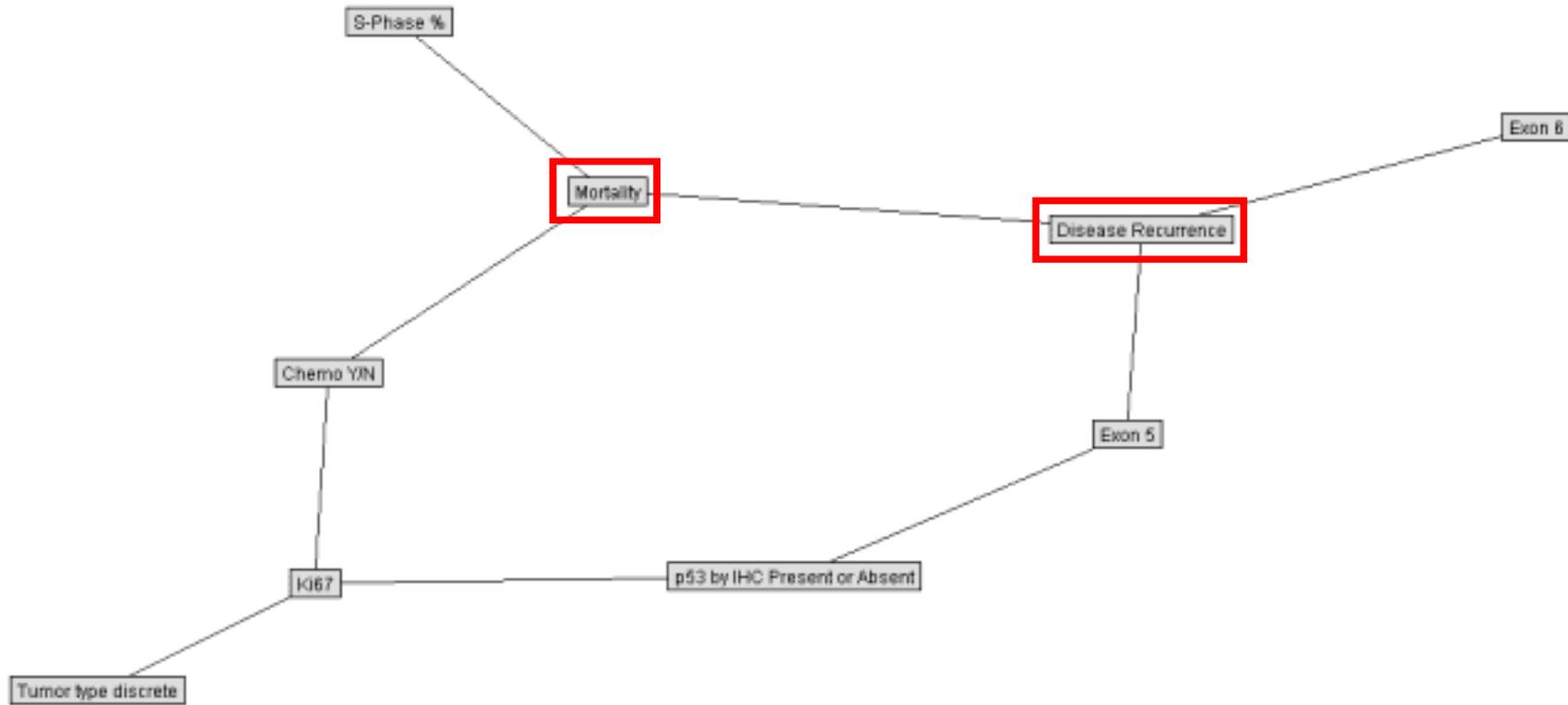
- p53 mutations as detected by IHC appear more frequently in tumors with wild type p53 gene and were not identified in 100% of Exon 5 mutations

Figure 1 – Bayesian Model Focused on Recurrence of Disease



- Recurrent disease predicted by mutations on Exons 5 and 6 and high S-Phase %
- p53 by IHC negatively correlated with mutations on Exon 5

Figure 2 – Bayesian Model Using Disease Recurrence and Mortality



- S-Phase %, chemotherapy, and disease recurrence are key predictors of mortality

Table 5 – Cross-Validation Results

- Cross validation was performed on the two models in Figure 1 and Figure 2
- We exploited the flexibility of a probabilistic model to improve our predictive values
- Mortality as a variable has poor validation results, impairing the clinical utility of this variable in this study
- Using the recurrence only model and the 20% probability threshold, the predictive value for recurrence of disease was 83.3%

Predictive Value for Each Model by Outcome at Different Thresholds (% Threshold Negative/% Threshold Positive)

<u>Model</u>	<u>No Recurrence</u>	<u>Recurrence</u>	<u>Alive</u>	<u>Dead</u>
<i>Recurrence and Mortality Model</i>				
PV 50/50	70.0%	100.0%	77.8%	33.3%
PV 80/20	92.9%	85.7%	80.0%	27.3%
PV 90/10	87.5%	46.2%	83.3%	26.7%
<i>Recurrence Only Model</i>				
PV 50/50	70.0%	NA	NA	NA
PV 80/20	86.7%	83.3%	NA	NA
PV 90/10	80.0%	54.5%	NA	NA

Table 6 – Predicted Probabilities of Disease Recurrence

Expected Probability of case	Input Drivers			Target Probability	
	Exon 5	Exon 6	S-Phase %	Disease Recurrence	
				No	Yes
56.2%	No	No	Up to 5	97.9	2.1
2.3%	No	Yes	Up to 5	85.3	14.7
30.0%	No	No	5 plus	82.5	17.5
2.9%	Yes	No	5 plus	55.4	44.6
5.0%	Yes	No	Up to 5	48.2	51.8
2.4%	No	Yes	5 plus	36.7	63.3
0.4%	Yes	Yes	5 plus	13.2	86.8
0.8%	Yes	Yes	Up to 5	10.3	89.7

- Using S-Phase % and mutation status, the model permits the calculation of the expected probability of disease recurrence

Discussion

- Analysis and modeling indicate that recurrence of disease in this population can be predicted using knowledge of mutations on Exons 5 and 6, and the percentage of cells in synthesis phase (S-Phase %)
- Exon 5 correlates with p53 as measured by IHC in an inverse relationship – the absence of p53 as measured by IHC indicates a likelihood of mutation on Exon 5 (**Ref 6?**)
- p53 by IHC was not a predictor of recurrent disease in this study
- The strong cross-validation of the models relying on Exon 5 and 6 status indicates the importance of a molecular diagnostic in this patient population
- Using prognostic and predictive biomarkers, molecular characterization, and Bayesian Networks the probability of disease recurrence can be calculated

Conclusion

- In making decisions about adjuvant therapy in a controversial breast cancer population an accurate, quantitative measure of risk assessment is essential
- Mutations on Exon 5, Exon 6, and high S-Phase % are key predictive markers of disease recurrence
- p53 mutations as measured by IHC do not predict disease recurrence
- By encoding our findings in a Bayesian probabilistic network, we can risk stratify a *de novo* population
- Cross validation analysis demonstrates that these models are effective classifiers of disease recurrence
- This approach provides us with an effective quantitative risk assessment tool for use in controversial breast cancer patients
- Given these results, we are evaluating a prospective validation of this methodology in a larger study population using Roche AmpliChip p53

References

- (1) Bull SB, Ozcelik H, Pinnaduwege D, Blackstein ME, Sutherland DA, Pritchard KI, Tzontcheva AT, Sidlofsky S, Hanna WM, Qizilbash AH, Tweeddale ME, Fine S, McCready DR, Andrulis IL., The combination of p53 mutation and neu/erbB-2 amplification is associated with poor survival in node-negative breast cancer. *J Clin Oncol.* 2004 Jan 1;22(1):86-96.
- (2) Rahko E, Blanco G, Soini Y, Bloigu R, Jukkola A., A mutant TP53 gene status is associated with a poor prognosis and anthracycline-resistance in breast cancer patients., : *Eur J Cancer.* 2003 Mar;39(4):447-53.
- (3) Fisher, B., Costantino, J., Redmond, C., et al., A randomized clinical trial evaluating tamoxifen in the treatment of patients with node-negative breast cancer who have estrogen-receptor-positive tumors. *N Engl J Med* 1989;320;479-84
- (4) Henderson, IC, Patek, AJ. The relationship between prognostic and predictive factors in the management of breast cancer. *Breast Cancer Res Treat* 1998;52;261-88
- (5) Moraleda, J. New Algorithms, Data Structures, and User Interfaces for Machine Learning of Large Datasets with Applications. Doctoral Dissertation, Leland Stanford, Jr. University, Palo Alto, CA, December 2003
- (6) Ref – p53 by IHC