Random_Intensity_60_5a: 60 Cases, 60 Controls, 300 Peaks
Brian T. Luke (lukeb@ncifcrf.gov)

BMDK Analysis
27 peaks selected as putative biomarkers by the 10 methods within BMDK

Peak	catboot	student	dtgini	dtinfg	nnfeat	chisq	kruswal	kolsmir	extreme	vartest
6			3	2						
18								4		
29	5					2				
33	3									
51								4		
52						5				
90	4							4		
94					5		5			
103									3	
105	1									
114									3	
116			4	5						
118		4			4		4			4
128		5								5
140		3	4		1		3	3		3
157									2	
158				4					1	
159		1	1	3	3	4	1	1		1
162									3	
177						5				
184	2									
193		2			2		2	2		2
231			2	1						
239								4	3	
241						1			3	
260						2				
269									3	

Peaks used in each of the best distance-dependent 6-nearest neighbor classifiers

Peak	1-ad	2-ad	3-ad	1-rd	2-rd	3-rd	2-cr	3-cr	2-sd	3-sd
6			X		X					
18										
29										
33						X		X		X
51										
52										
90										
94						X		X		X
103										
105										
114										
116			X				X		X	
118										
128										
140		X				X		X		X
157										
158										
159										
162	X			X			X		X	
177					X					
184										
193										
231			X							
239										
241		X								
260										
269										

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 27 putative biomarkers.

Metric	1-ad	2-ad	3-ad	1-rd	2-rd	3-rd	2-cr	3-cr	2-sd	3-sd
Sens	78.3	71.2	73.7	77.6	71.4	67.6	70.2	73.1	70.2	74.0
Spec	73.3	66.7	65.0	71.2	68.5	78.0	72.2	74.5	71.7	72.5
%Undet	0.0	0.8	2.5	2.5	14.2	30.0	7.5	14.2	8.3	15.8
Quality	151.7	137.0	136.2	146.3	125.8	115.6	134.9	133.4	133.5	130.7

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 27 putative biomarkers with the caveat that %Undetermined cannot exceed 5.0%.

Metric	1-ad	2-ad	3-ad	1-rd	2-rd	3-rd	2-cr	3-cr	2-sd	3-sd
Sens	78.3	71.2	73.7	77.6	None	None	62.7	None	56.1	None
Spec	73.3	66.7	65.0	71.2	None	None	61.4	None	73.7	None
%Undet	0.0	0.8	2.5	2.5	None	None	3.3	None	5.0	None
Quality	151.7	137.0	136.2	146.3	None	None	120.8	None	124.8	None

Fingerprint Analysis

Sensitivity, specificity and quality (sensitivity + specificity) for the best and 200th best decision tree constructed from any of the 300 peak intensities. The evolutionary programming search used a population size of 200 and ran for 400 generations. A decision node became a terminal node when it contained 1% (no samples) or 4% (2 samples) of a given State.

Metric	1%		1%		4%		4%	
	1st	200th	1st	200th	1st	200th	1st	200th
Sensitivity	83.3	82.5	81.7	83.3	83.3	83.3	78.3	78.3
Specificity	88.3	87.5	88.3	85.0	88.3	85.0	91.7	86.7
Quality	171.7	170.0	170.0	168.3	171.7	168.3	170.0	165.0

Sensitivity, specificity and quality (sensitivity + specificity) for the best and 200th best medoid classifier algorithm in each of the two runs using 5-, 6-, and 7-peak intensities from the set of 300. The evolutionary programming search used a population size of 400 and ran for 800 generations with the requirement that there are at most 40 Case-cells and 40 Control-cells.

Metric	5-Features		5-Features		6-Features		6-Features		7-Features		7-Features	
	1st	200th	1st	200th	1st	200th	1st	200th	1st	200th	1st	200th
Sens	100.0	100.0	90.0	83.3	100.0	100.0	91.7	86.7	100.0	100.0	93.3	90.0
Spec	93.3	81.7	100.0	100.0	93.3	86.7	100.0	100.0	95.0	90.0	100.0	100.0
Quality	193.3	181.7	190.0	183.3	193.3	186.7	191.7	186.7	195.0	190.0	193.3	190.0

(Last updates 4/21/07)