<u>PDF Version</u>

 $Random_Intensity_60_4a: 60 \ Cases, 60 \ Controls, 300 \ Peaks \\ Brian \ T. \ Luke \ (\underline{lukeb@ncifcrf.gov})$

BMDK Analysis

22 peaks selected as putative biomarkers by the 10 methods within BMDK

Peak	catboot	student	dtgini	dtinfg	nnfeat	chisq	kruswal	kolsmir	extreme	vartest
27		3			1		3	3		3
28									3	
36			4			1		3		
43	5	2			2		1	3		2
82	1									
83				3		5			1	
99			1	1				1		
117					3			3		
126						3				
127	4									
139						5			1	
146			5	5						
166									3	
209									3	
220			2	2						
234						2				
245									3	
247		1	3	4	5	3	2	1		1
257		5					5			5
262	3	_					_	_		_
269	_	4			4		4	_		4
286	2	_	_					_	_	

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
27										
28								Χ		
36									Χ	Χ
43						Χ				
82								Χ		
83			Χ						Χ	Χ
99			Χ			Χ				Χ
117	Χ	Χ		Χ	Χ		Χ			
126						Χ				
127										
139										
146			X					X		
166										
209										
220										
234										
245										
247										
257										
262					X		Х			
269		Χ								
286										

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 22 putative biomarkers.

Metric	1-ad	2-ad	3-ad	1-rd	2-rd	3-rd	2-cr	3-cr	2-sd	3-sd
Sens	68.3	77.6	77.1	69.6	68.6	77.5	66.1	73.1	68.6	70.6
Spec	61.7	63.3	75.0	60.3	76.9	76.2	71.9	71.2	73.7	78.7
%Undet	0.0	1.7	16.7	5.0	14.2	31.7	5.8	13.3	10.0	18.3
Quality	130.0	139.3	135.4	125.0	131.4	122.0	132.2	130.9	132.3	131.0

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 22 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	68.3	77.6	71.2	69.6	None	None	75.0	None	75.4	None
Spec	61.7	63.3	63.3	60.3	None	None	53.7	None	57.9	None
%Undet	0.0	1.7	0.8	5.0	None	None	5.0	None	5.0	None
Quality	130.0	139.3	133.7	125.0	None	None	123.7	None	128.3	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.

Matria	1%		19	%	49	%	4%		
Metric	1 st	200 th	1^{st}	200 th	1^{st}	200 th	1 st	200 th	
Sensitivity	93.3	90.0	88.3	83.3	91.7	85.0	83.3	83.3	
Specificity	83.3	83.3	88.3	88.3	85.0	88.3	88.3	85.0	
Quality	176.7	173.3	176.7	171.7	176.7	173.3	171.7	168.3	

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	
Metric	1 st	200 th	1^{st}	200^{th}	1 st	200 th						
Sens	100.0	100.0	90.0	83.3	100.0	100.0	93.3	85.0	100.0	100.0	91.7	86.7
Spec	93.3	83.3	100.0	100.0	93.3	85.0	100.0	100.0	95.0	90.0	100.0	100.0
Quality	193.3	183.3	190.0	183.3	193.3	185.0	193.3	185.0	195.0	190.0	191.7	186.7

(Last updated 4/21/07)