

Random_Intensity_150_2a: 150 Cases, 150 Controls, 300 Peaks
 Brian T. Luke (lukeb@ncifcrf.gov)

BMDK Analysis

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

Peak	catboot	student	dtgini	dtinfg	nnfeat	chisq	kruswal	kolsmir	extreme	vartest
4			1	1					1	
6						4				
9		1					1			1
16									3	
43									3	
54						3				
82						5				
90									3	
93		4					4			4
97	5		4	3						
115	2							2		
117	4				3					
157									3	
162								4		
168			3	5	4			4		
169					5					
183									3	
210	1	5	2	4			5	1		5
218								4		
228						2				
237		3			2		3	4		3
246									2	
255				2						
260	3	2	5		1		2	2		2
276						1				

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 25 putative biomarkers.

Metric	1-ad	2-ad	3-ad	1-rd	2-rd	3-rd	2-cr	3-cr	2-sd	3-sd
Sens	60.0	61.3	62.0	60.0	59.7	64.3	63.4	59.7	62.2	69.6
Spec	56.7	64.0	63.3	55.8	65.7	64.1	62.2	66.9	61.2	56.5
%Undet	0.0	0.0	0.0	1.0	9.7	23.7	3.3	7.7	3.3	13.7
Quality	116.7	125.3	125.3	114.8	115.7	104.7	122.2	111.3	120.1	112.4

Sensitivity, specificity, %undetermined, and quality (sensitivity + specificity - %undetermined) for each of the best distance-dependent 6-nearest neighbor classifiers using any of the 25 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	60.0	61.3	62.0	60.0	None	None	63.4	None	62.2	None
Spec	56.7	64.0	63.3	55.8	None	None	62.2	None	61.2	None
%Undet	0.0	0.0	0.0	1.0	None	None	3.3	None	3.3	None
Quality	116.7	125.3	125.3	114.8	None	None	122.2	None	120.1	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 400 and ran for 800 generations. A decision node became a terminal node when it contained 1% (1 sample) or 4% (6 samples) of a given State.

Metric	1%		1%		4%		4%	
	1 st	200 th						
Sensitivity	76.0	72.7	88.0	74.7	67.3	62.7	69.3	66.0
Specificity	73.3	73.3	62.7	74.7	83.3	82.7	82.0	84.0
Quality	149.3	146.0	150.7	149.3	150.7	145.3	151.3	150.0

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 1000 and ran for 2000 generations with the requirement that there are at most 100 Case-cells and 100 Control-cells.

Metric	5-Features		5-Features		6-Features		6-Features		7-Features		7-Features	
	1 st	200 th										
Sens	100.0	100.0	81.3	74.7	100.0	100.0	82.7	78.0	100.0	100.0	87.3	80.0
Spec	83.3	74.7	100.0	100.0	82.7	78.7	100.0	100.0	84.7	80.7	100.0	100.0
Quality	183.3	174.7	181.3	174.7	182.7	178.7	182.7	178.0	184.7	180.7	187.3	180.0

(Last updated 4/21/07)