

Random_Intensity_150_4a: 150 Cases, 150 Controls, 300 Peaks
 Brian T. Luke (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
1									2	
33						5			4	
38						1			4	
39						5				
64		3	1	1		4	3	2		3
67	4									
69	2	4					4	2		4
72									4	
76			3	2						
85									1	
92						1				
150		5	5	5	3		5	4		5
154					2					
161			4	4					2	
166			5	5						
186	3									
226	5	2			1		2	4		2
232					4					
262	1	1	2	3			1	1		1
269									4	
271						3				
298					5					

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	60.7	60.0	62.0	61.3	64.0	69.8	60.0	64.4	60.0	71.1
Spec	57.3	63.3	63.3	56.6	62.3	67.9	65.1	65.0	64.4	60.9
%Undet	0.0	0.0	0.0	1.7	7.7	24.0	3.0	10.3	3.0	12.3
Quality	118.0	123.3	125.3	116.2	118.7	113.7	122.1	119.0	121.4	119.7

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	60.7	60.0	62.0	61.3	None	None	60.0	None	60.0	None
Spec	57.3	63.3	63.3	56.6	None	None	65.1	None	64.4	None
%Undet	0.0	0.0	0.0	1.7	None	None	3.0	None	3.0	None
Quality	118.0	123.3	125.3	116.2	None	None	122.1	None	121.4	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	67.3	72.0	66.0	69.3	85.3	86.0	78.7	75.0
Specificity	84.7	78.0	83.3	78.7	70.0	68.0	74.7	76.0
Quality	152.0	150.0	149.3	148.0	155.3	154.0	153.3	151.0

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	80.0	74.7	100.0	100.0	85.3	78.7	100.0	100.0	85.3	81.3
Spec	80.7	75.3	100.0	100.0	82.0	78.7	100.0	100.0	84.0	80.7	100.0	100.0
Quality	180.7	175.3	180.0	174.7	182.0	178.7	185.3	178.7	184.0	180.7	185.3	181.3

(Last updated 4/21/07)