BAP1 Dataset Information

July 2023

Table 1: Patient demographics and tumor volume statistics.

	Total (n=131)	**BAP1 Status [+] (n=60)**	**BAP1 Status [-] (n=71)**
Sex			
Male	84	42	42
Female	47	18	29
Age			
Median	69	69	69
Range	21-90	51-90	21-81
Tumor volume [mm^3]			
Median	12697	11045	14485
Range	1323 - 98413	1323 - 98413	2134 - 66561

Table 2: Image acquisition characteristics.

	Total (n=131)	**BAP1 Status [+] (n=60)**	**BAP1 Status [-] (n=71)**
Pixel Size [mm]			
Median	0.72	0.71	0.73
Range	0.56–1.07	0.57–1.07	0.56–0.95
Slice Thickness [mm]			
Median	3	3	3
Range	1–5	1–5	1–5
kVp [kV]			
Median	120	120	120
Range	80–140	100–120	80–140
Scanner Manufacturer			
GE	65	31	34
Philips	43	20	23
Toshiba	10	4	6
Siemens	13	5	8
Reconstruction Kernel			
GE: Standard	64	31	33
GE: Chest	1	0	1
Philips: B	42	20	22
Philips: C	1	0	1
Toshiba: FC13	5	1	4
Toshiba: FC14	1	1	0
Toshiba: FC18	4	2	2
Siemens: B30f	3	0	3
Siemens: B31f	1	0	1
Siemens: B40f	1	0	1
Siemens: B31s	1	0	1
Siemens: B35s	1	1	0
Siemens: Bf39f	1	0	1
Siemens: Br36f	1	0	1
Siemens: Br40d	1	1	0
Siemens: I31f	2	2	0
Siemens: I41f	1	1	0

Table 1: Comparisons of the three classification models: gradient boosting, extra trees, and random forest. P-values comparing the differences in AUC values were calculated using the DeLong test, with their corresponding confidence intervals (CIs), prior to multiple comparisons correction. Significance levels (α) and widths of the CIs were adjusted based on multiple comparisons.

Comparison	**p-value for ΔAUC**	α	**CI of ΔAUC**
Gradient boosting versus extra trees	0.0013*	0.05	97.5% CI: [0.0068 0.038]
Gradient boosting versus random forest	0.3917	0.017	98.3% CI: [0.014 0.040]
Extra trees versus random forest	0.32	0.025	95% CI: [-0.0041 0.013]

*Statistical significance was achieved after correcting for multiple comparisons using Bonferroni-Holm corrections.

Table 1: Texture features selected for the three classifiers across the 100 repeated cross validations and the selection frequency each feature was chosen for classification.

Transformation	Class	Feature	Selection Frequency
LoG ($\sigma = 1.5$ mm)	GLSZM	High Gray Level Zone Emphasis	100%
LoG ($\sigma = 1.55$ mm)	GLCM	Cluster Prominence	100%
LoG ($\sigma = 0.75$ mm)	GLSZM	High Gray Level Zone Emphasis	95%
Wavelet (bior1.1–HL)	GLDZM	High Gray Level Emphasis	95%
LoG ($\sigma = 0.75$ mm)	GLCM	Correlation	92%