Supplemental material

Technical details about FLAB segmentation and textural features computation.

FLAB automatic delineation

The Fuzzy Locally Adaptive Bayesian (FLAB) algorithm computes a probability of belonging to a given "class" (e.g. tumor with high or moderate uptake, and background) for each voxel within a 3D Region of interest (ROI) containing the tumor and its surrounding physiological background. This probability is iteratively estimated by taking into account the voxel's intensity with respect to the statistical distributions (characterized by their mean and variance) within the ROI, and its spatial correlation with neighbouring voxels.

The method proceeds in three steps:

1. Initialization: set the number of classes (2 or 3) depending on the heterogeneity of the tumor. Use Fuzzy C-means clustering to provide initial estimate of means and variances of each class. Means and variance of fuzzy transitions levels between each class are linearly calculated based on the two classes. Set equally distributed prior locally adaptive probabilities for all voxels.
2. Iterative estimation: use the stochastic expectation maximization algorithm to estimate the means and variances of each class and derived probabilities (Gaussian distributions) as well as the prior locally adaptive probabilities using a 3×3×3 sliding cube. Update the means and variances of fuzzy transitions levels accordingly. Stop when stopping criteria is reached (less than 1% change in the parameters between two successive iterations).
3. Assign a class or fuzzy level to each voxel by selecting the one that maximizes its posterior probability according its value with respect to mean and variance of each class or fuzzy level multiplied by its prior locally adaptive probability.
4. Associate all voxels classified as a fuzzy level to their closest class to obtain a 2-class or 3-class segmentation.

For additional mathematical and algorithmic implementation details, we refer the reader to previous works (24, 25).

Heterogeneity quantification through texture analysis

Once the tumor has been delineated, textural features are computed on the volume of interest.
There are three major steps involved:

1. Resampling the voxels' values into a smaller range of values. This allows the calculation of the features, by reducing the size of the matrices (see next step) and reducing the impact of noise. A previous study has shown that this resampling step may have an impact on the reproducibility of parameters, and

that a value of 64 (compared to 8, 16, 32, 128) was the best compromise to represent the range of SUV values (21).

2. Construction of normalized co-occurrences matrices along each of the 13 directions of the volume of interest and size-zone matrices (for which all directions are considered within a single matrix) which are defined as the following:

    a. Co-occurrence matrices: C is a 64×64 square matrix in which C(i,j) is the number of transitions between a voxel with intensity i and a voxel with a intensity j.

    b. Size-zone matrices: Z is a matrix with 64 lines and n columns with n the size of the largest homogeneous area within the tumor. Consequently, Z(i,j) denotes the number of areas with an intensity i and a size j. Homogeneous areas are constructed by identifying contiguous voxels with equal intensity after the resampling step.

3. Calculation of the features of interest.
    a. On the one hand, entropy (E), dissimilarity (D) and homogeneity (H) are "local" heterogeneity features that are calculated on each of the 13 co-occurrences matrices corresponding to each direction, followed by averaging over the 13 matrices in order to obtain a single value per tumor:

$$E = -\sum_i \sum_j C(i,j) \log\big(C(i,j)\big)$$

$$D = \sum_i \sum_j |i-j| C(i,j)$$

$$H = \sum_i \sum_j \frac{1}{1+|i-j|} C(i,j)$$

    b. On the other hand, size-zone variability (SZV), high intensity emphasis (HIE) and zone percentage (ZP) are calculated on size-zone matrices as follows:

$$SZV = \frac{1}{\Omega} \sum_{j=1}^{L} \left[\sum_{i=1}^{D} \frac{Z(i,j)}{i^2} \right]^2$$

$$HIE = \frac{1}{\Omega} \sum_{i=1}^{D} \sum_{j=1}^{L} i^2 . Z(i,j)$$

$$ZP = \Omega / S$$

Ω and S are respectively the number of homogeneous areas and the number of voxels. L is the size of the largest homogeneous area and D the resampling value (64 in the present case).

For additional mathematical details, we refer the reader to previous works (19, 21,22).

Supplemental Table 1: Inter-observer's agreement according to weighted kappa test for the 3-level heterogeneity scale

	Observer A			
Observer B	1	2	3	Total (%)
1	10	8	0	18 (17.6)
2	3	16	10	29 (28.4)
3	0	7	48	55 (53.9)
Total (%)	13 (12.7)	31 (30.4)	58 (56.9)	102 (100)

Weighted Kappa[a]	0.636
Standard error	0.060
95% CI	0.518 to 0.753

[a]Linear weights

Supplemental Table 2: Inter-observer agreement according to weighted kappa test for the 5-level heterogeneity scale (A stands for “diffuse” and B for “localized”)

	Observer A					
Observer B	1	2A	2B	3A	3B	Total (%)
1	9	5	3	0	0	17 (16.7)
2A	0	2	7	0	3	12 (11.8)
2B	3	0	7	0	7	17 (16.7)
3A	0	1	0	1	4	6 (5.9)
3B	1	0	6	0	43	50 (49.0)
Total	13 (12.7)	8 (7.8)	23 (22.5)	1 (1.0)	57 (55.9)	102 (100)

Weighted Kappa[a]	0.582
Standard error	0.060
95% CI	0.465 to 0.699

[a]Linear weights

Supplemental Table 3: Correlation between H_{visu} and each heterogeneity feature, according to spearman rank coefficient (n=102)

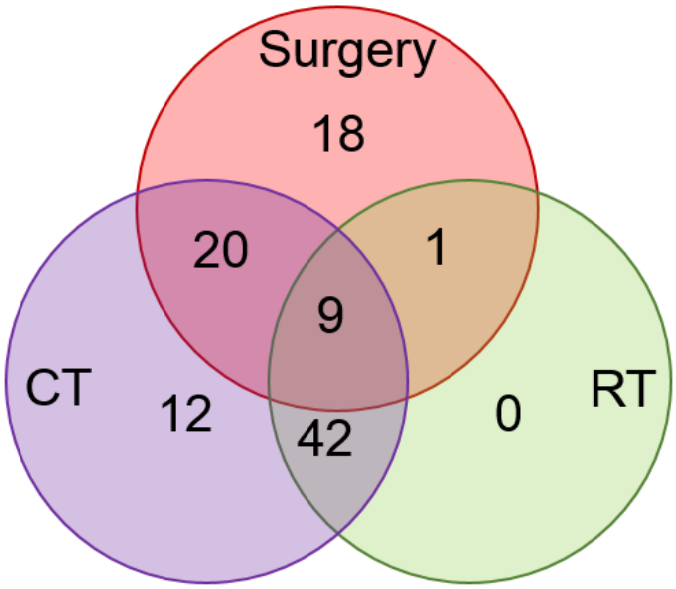
Parameter	Spearman rank correlation with H_{visu} rho	p-value
SUV_{COV}	-0.22	0.03
CH_{AUC}	-0.07	0.49
E	0.60	<0.0001
H	0.59	<0.0001
D	-0.59	<0.0001
SZV	-0.60	<0.0001
ZP	-0.44	<0.0001
HIE	-0.20	0.04

Supplemental Table 4: OS Kaplan-Meier analysis (n=102)

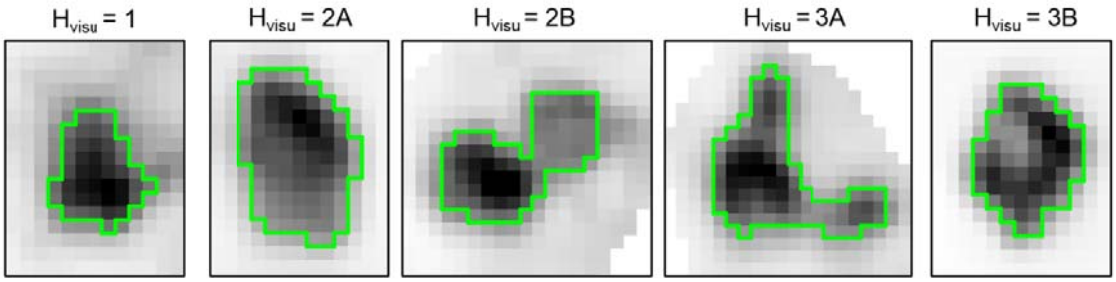
Parameters		Cut-off	Nb of events (nb of patients)	HR	HR 95% CI	p-value
Clinical	Surgery	=no	24 (48) / 38 (54)	2.00	1.22 to 3.29	0.006
	Age	>60	21 (41) / 41 (61)	1.78	1.10 to 2.95	0.03
	Sex	=male	09 (23) / 53 (79)	2.29	1.23 to 3.98	0.02
	Smoker	=yes	10 (16) / 52 (86)	1.06	0.55 to 2.06	0.8
	Histology	=adenocarcinoma	28 (49) / 34 (53)	1.27	0.78 to 2.08	0.3
	Stage	=III	23 (48) / 39 (54)	2.26	1.37 to 3.71	0.001
SUV and volumetric	SUV_{max}	>9.3	24 (52) / 38 (50)	2.74	1.57 to 4.77	0.009
	SUV_{mean}	>5.5	25 (54) / 37 (48)	2.73	1.57 to 4.76	0.007
	MATV	>35	25 (57) / 37 (45)	3.23	1.96 to 5.56	<0.0001
	TLG	>190	27 (58) / 35 (44)	2.94	1.66 to 5.20	0.001
Global heterogeneity	SUV_{COV}	≤0.24	29 (54) / 33 (48)	1.73	1.04 to 2.88	0.4
	CH_{AUC}	≤0.36	07 (18) / 55 (84)	2.14	1.12 to 4.07	0.9
Visual heterogeneity	H_{visu}	>1	12 (21) / 50 (81)	1.20	0.67 to 2.15	1
	H_{visu}	>2	24 (44) / 38 (58)	1.47	0.90 to 2.41	1
Local heterogeneity (TF)	E	>0.74	24 (55) / 38 (47)	3.14	1.85 to 5.35	<0.0001
	H	>0.37	19 (43) / 43 (59)	2.44	1.48 to 4.00	0.008
	D	≤0.57	11 (30) / 51 (72)	2.92	1.75 to 4.87	0.007
Regional heterogeneity (TF)	HIE	≤0.59	05 (16) / 57 (86)	2.29	1.18 to 4.47	0.8
	SZV	≤0.08	26 (58) / 36 (44)	3.08	1.79 to 5.26	<0.0001
	ZP	≤0.32	18 (45) / 44 (57)	3.02	1.83 to 4.98	<0.0001

Supplemental Table 5: RFS Kaplan-Meier analysis (n=48)

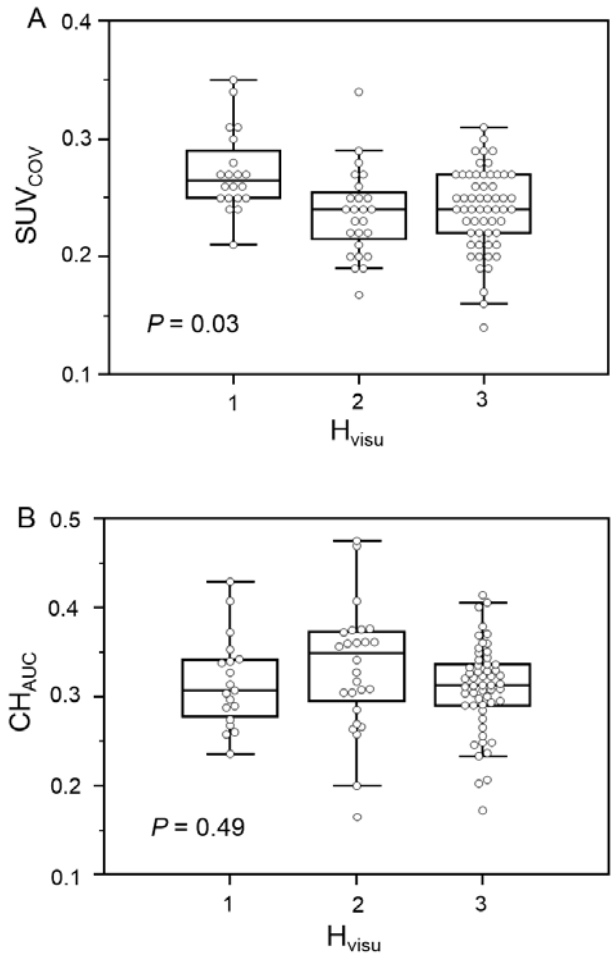
Parameters		Cut-off	Nb of events (nb of patients)	HR	HR 95% CI	p-value
Clinical	Age	>51	00 (04) / 32 (44)	1.09	0.52 to 2.10	0.9
	Sex	=male	06 (10) / 26 (38)	1.69	0.78 to 3.65	0.2
	Smoker	=yes	03 (07) / 29 (41)	2.51	1.07 to 5.92	0.1
	Histology	=adenocarcinoma	15 (23) / 17 (25)	1.17	0.58 to 2.33	0.7
	Stage	=III	23 (37) / 09 (11)	1.78	0.72 to 4.39	0.1
SUV and volumetric	SUV_{max}	>5.7	12 (20) / 20 (28)	2.52	1.00 to 6.29	1
	SUV_{mean}	>6.0	19 (31) / 13 (17)	2.23	0.93 to 5.32	0.6
	MATV	>35	20 (35) / 12 (13)	3.61	1.34 to 9.70	0.001
	TLG	>190	22 (36) / 10 (12)	3.24	1.15 to 9.15	0.03
Visual heterogeneity	H_{visu}	>1	09 (13) / 23 (35)	1.33	0.64 to 2.76	1
	H_{visu}	>2	14 (25) / 18 (23)	2.20	1.07 to 4.50	0.3
Global heterogeneity	SUV_{COV}	≤0.17	30 (46) / 02 (02)	0.87	0.36 to 2.08	1
	CH_{AUC}	≤0.37	05 (11) / 27 (37)	2.97	0.71 to 2.85	0.8
Local heterogeneity (TF)	E	>0.72	18 (32) / 14 (16)	3.26	1.36 to 7.86	0.004
	H	>0.29	09 (20) / 23 (28)	3.52	1.75 to 7.10	0.005
	D	≤0.57	09 (21) / 23 (27)	4.01	1.97 to 8.16	0.001
Regional heterogeneity (TF)	HIE	≤0.56	02 (08) / 30 (40)	3.70	1.55 to 8.83	0.7
	SZV	≤0.09	18 (32) / 14 (16)	3.26	1.36 to 7.85	0.004
	ZP	≤0.17	23 (39) / 09 (09)	3.69	1.14 to 11.96	0.004



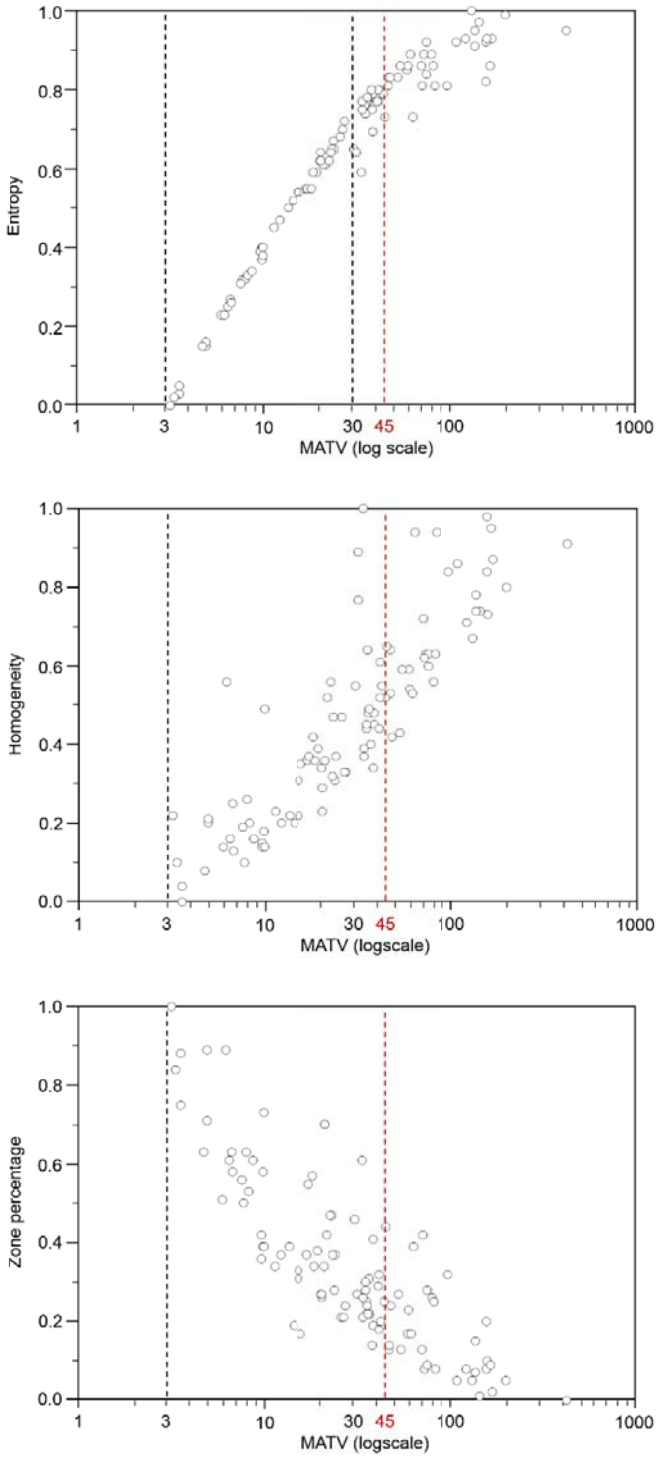
Supplemental Figure 1: treatment combinations in the cohort.



Supplemental Figure 2: examples of tumors with visual heterogeneity assessment according to the five-value scale.



Supplemental Figure 3: distributions of A) SUV_{COV} and B) CH_{AUC} according to H_{visu} score on the 3-level scale.



Supplemental Figure 4: distributions of entropy, homogeneity and zone percentage with respect to tumor metabolically active tumor volume (MATV, log-scale), showing the correlations between heterogeneity parameters and tumor volume.