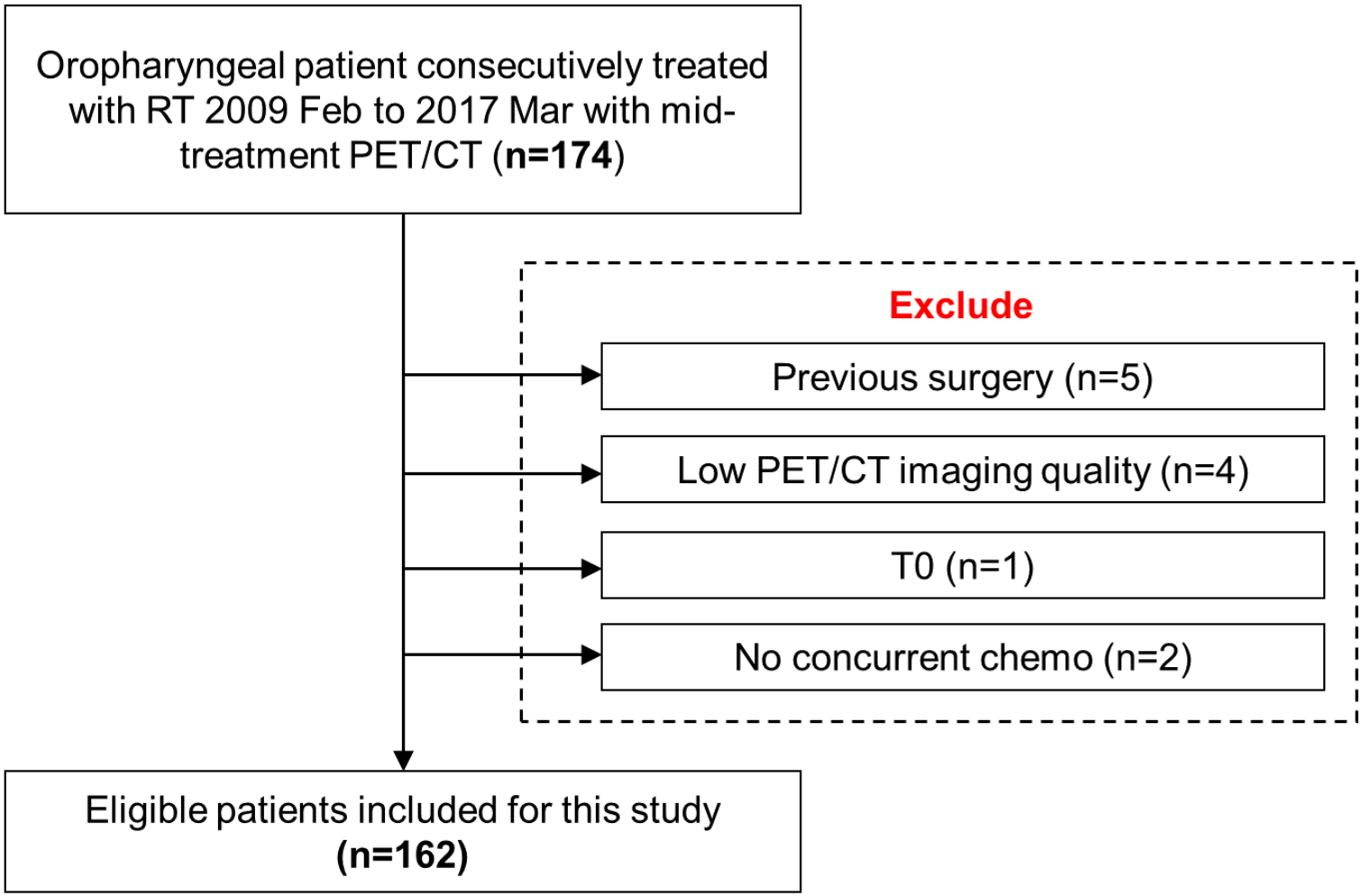
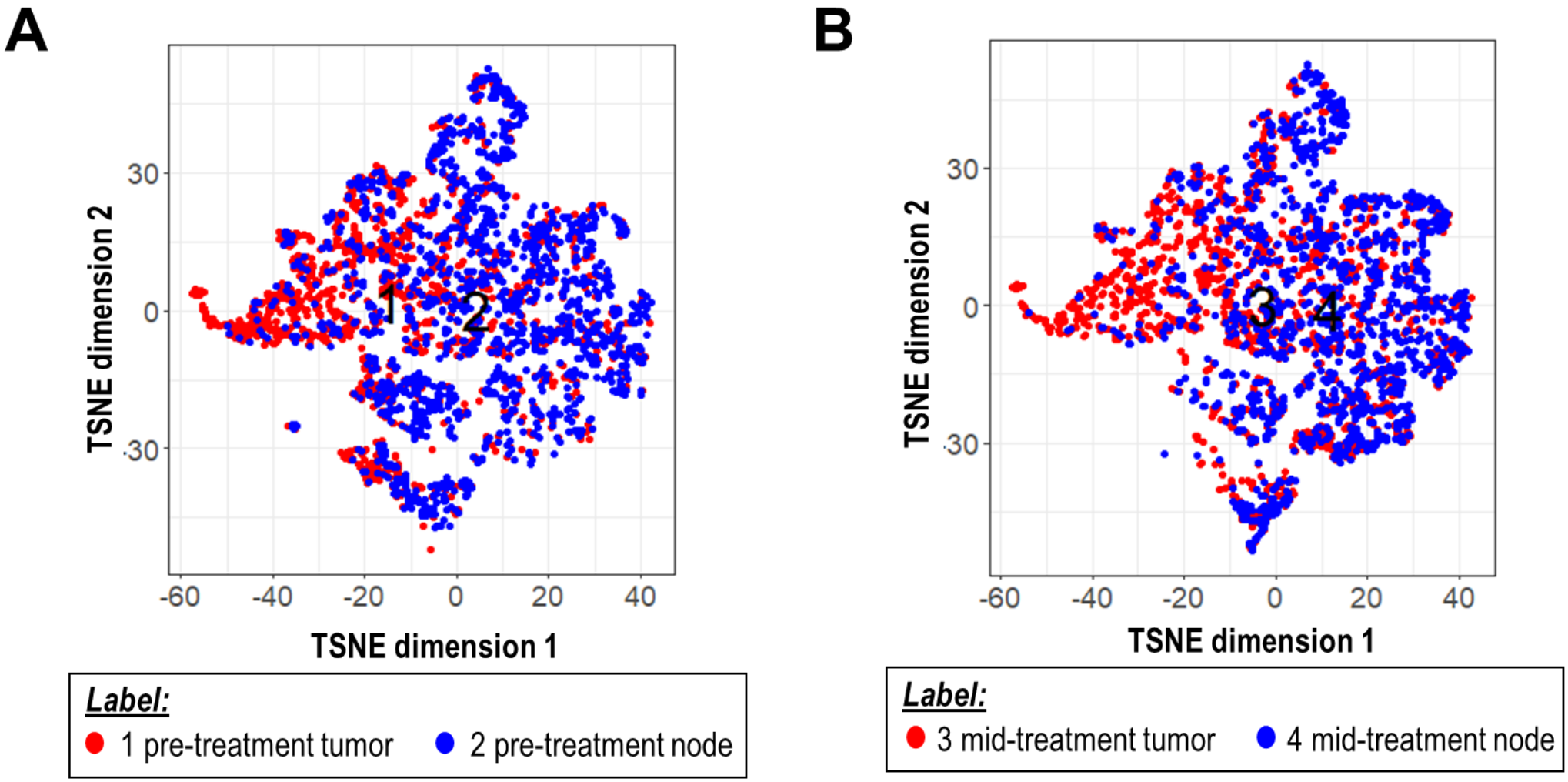
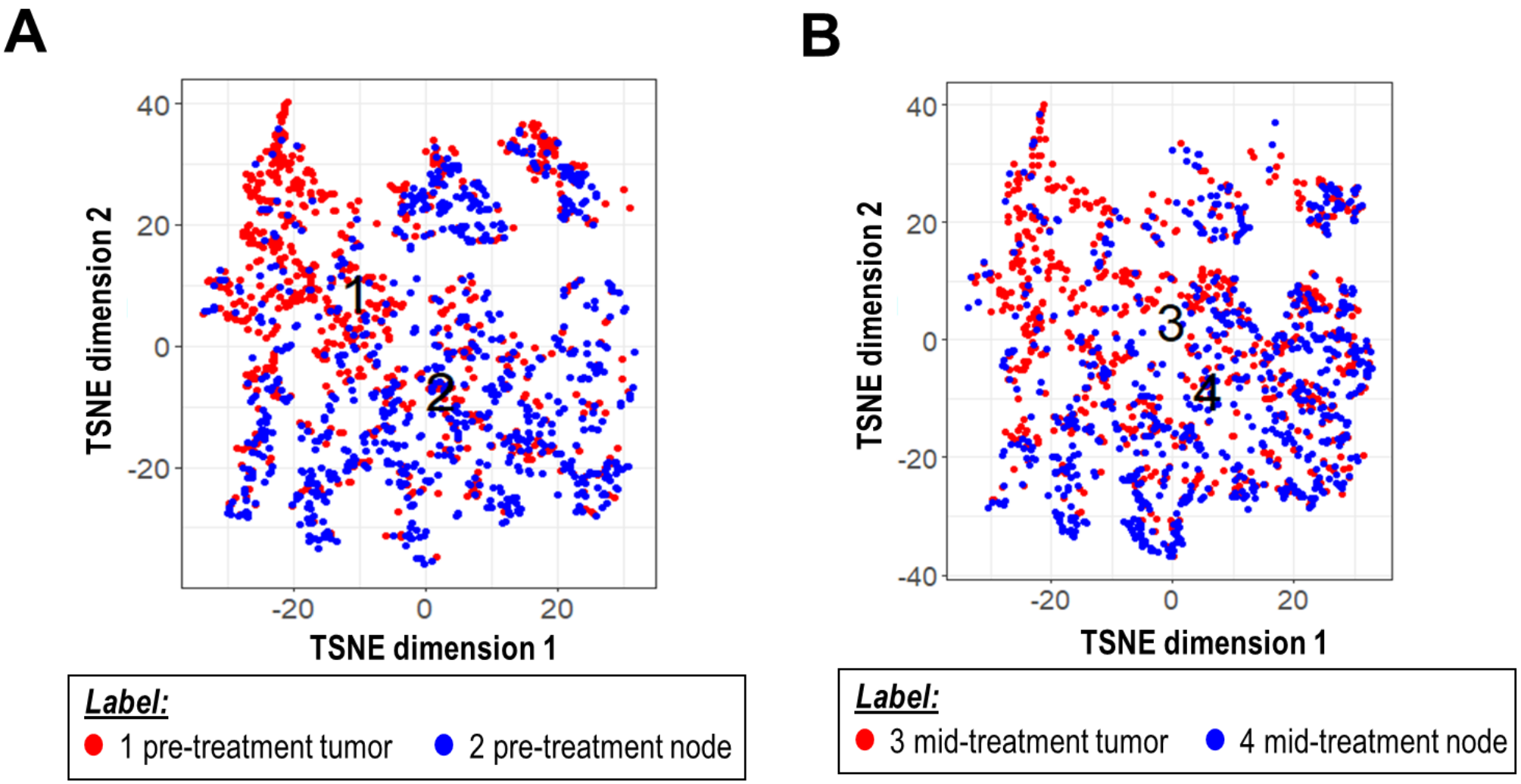


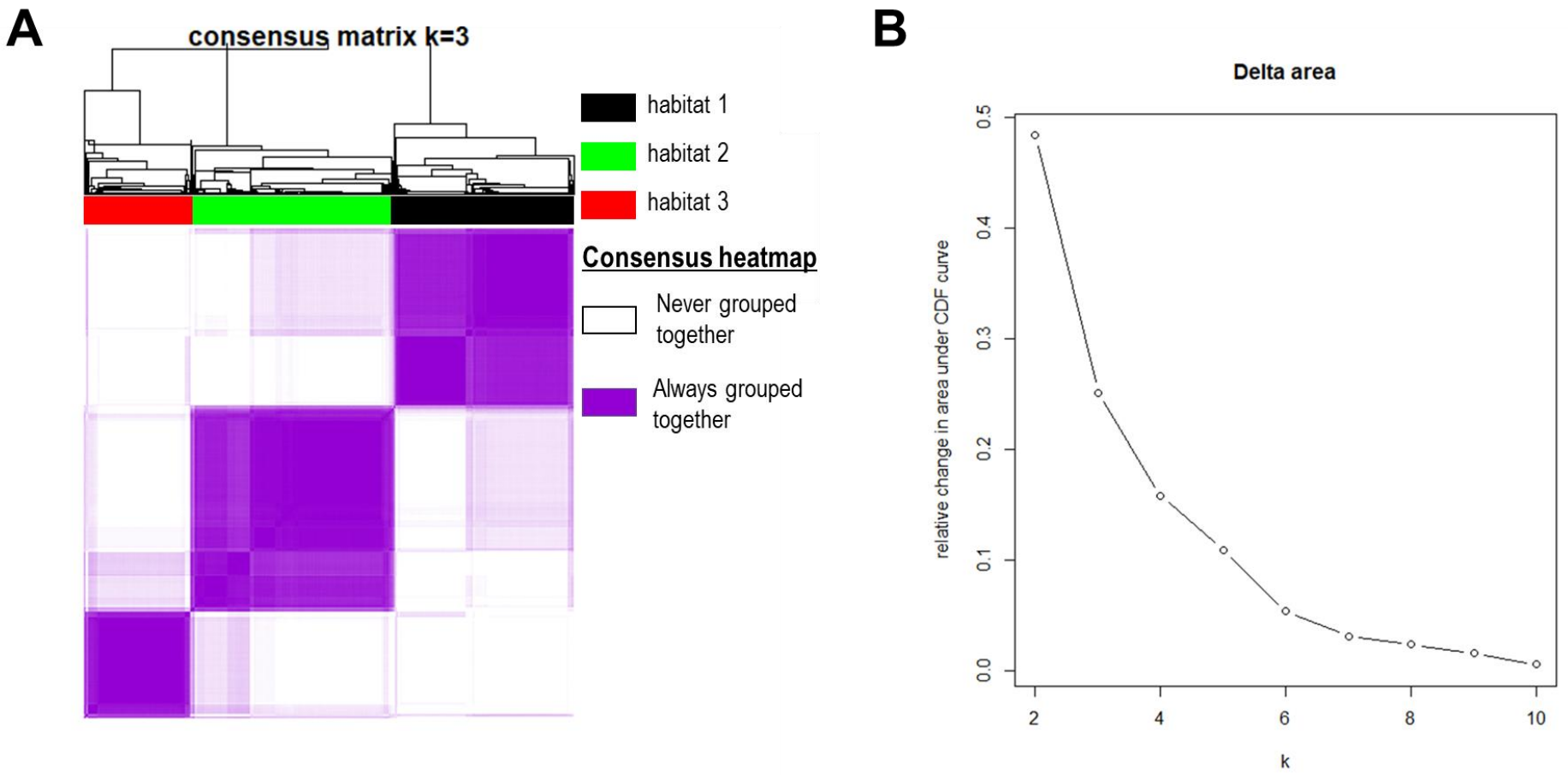
Supplemental Figure 1. Flowcharts of detailed patient's selection for the proposed study.



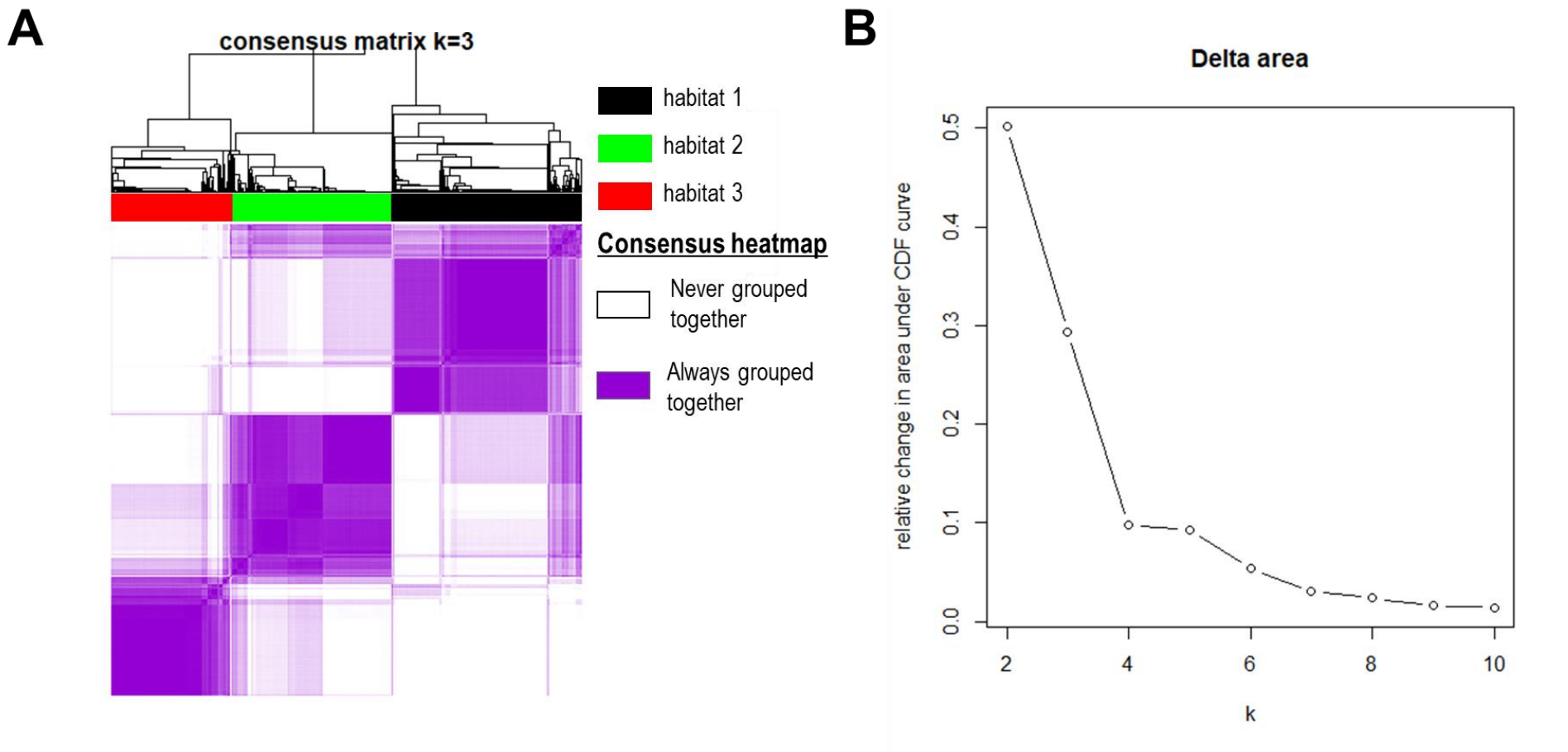
Supplemental Figure 2. t-SNE plot of superpixels from tumor (red) and lymph node (blue) of training cohort based on A) pre-treatment FDG-PET and CT imaging, and B) mid-treatment FDG-PET and CT imaging.



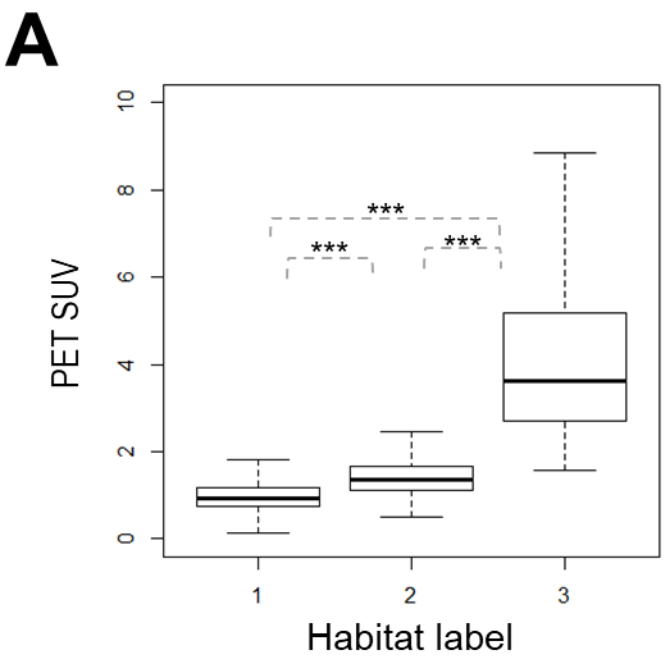
Supplement Figure 3. t-SNE plot of superpixels from tumor (red) and lymph node (blue) of validation cohort based on A) pre-treatment FDG-PET and CT imaging, and B) mid-treatment FDG-PET and CT imaging.

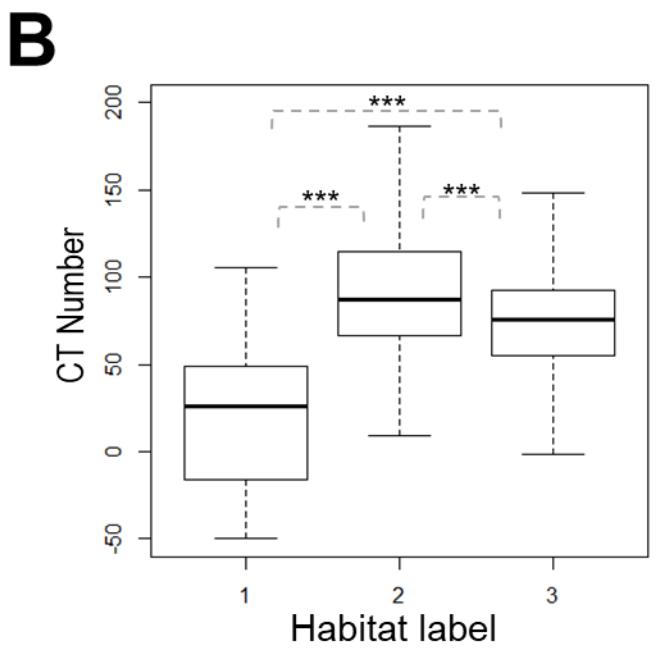


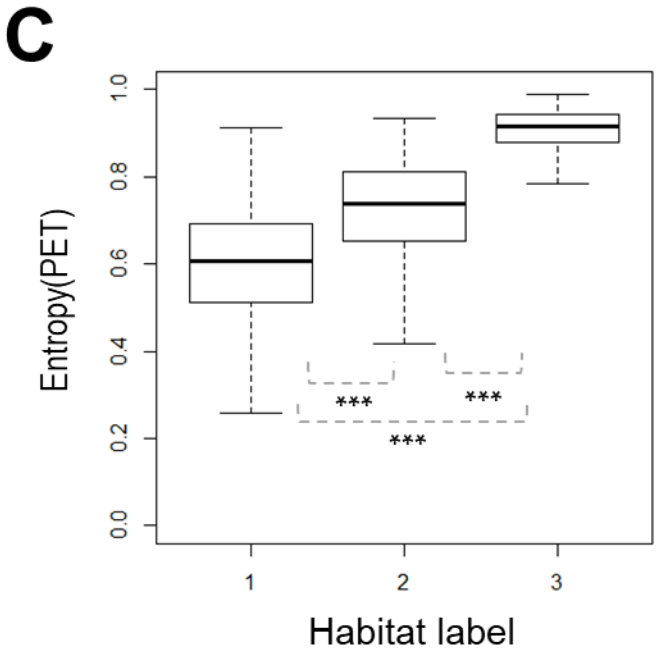
Supplement Figure 4. For training cohort, A) the consensus matrix, B) the delta CDF curve for all clusters. The consensus matrix represented as a heat map for k = 3. Patients are both rows and columns, and consensus values range from 0 (never clustered together, white) to 1 (always clustered together, dark violet). The matrix is ordered by consensus-clustered groups, depicted as a dendrogram above the heat map. The delta curve depicts the CDF progression graph, plotting the relative change in area under CDF curve, comparing k with k+1. The goal is to select the largest k that induced the smallest incremental change in the area under curve. Abbreviation: CDF = cumulative distribution function.

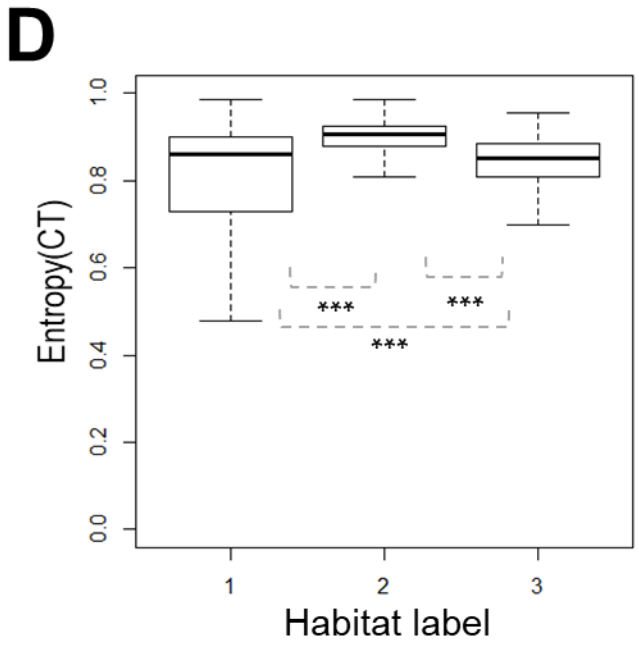


Supplemental Figure 5. For validation cohort, A) the consensus matrix, B) the delta CDF curve for all clusters. The consensus matrix represented as a heat map for k = 3. Patients are both rows and columns, and consensus values range from 0 (never clustered together, white) to 1 (always clustered together, dark cyan). Abbreviation: CDF = cumulative distribution function.

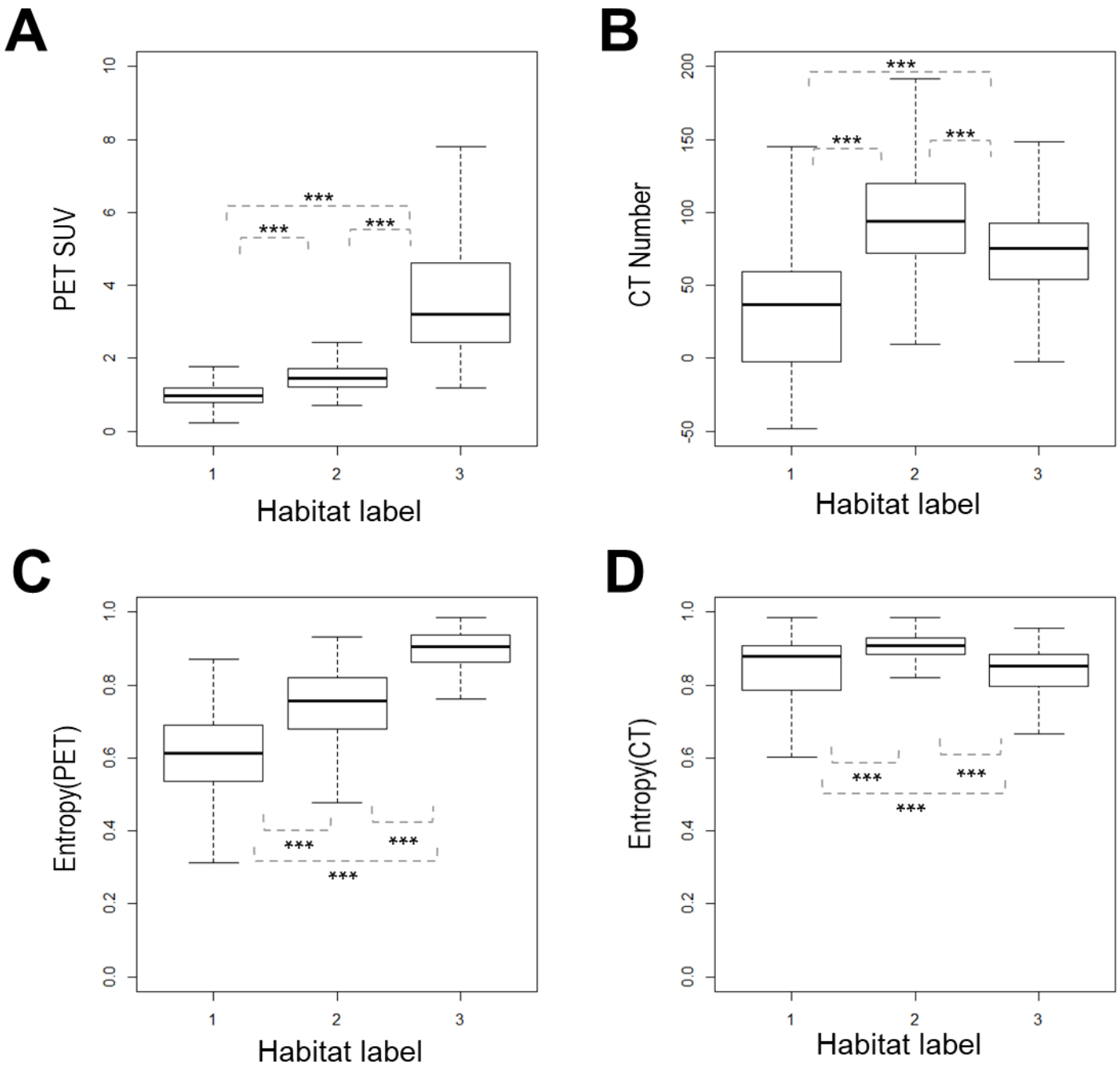




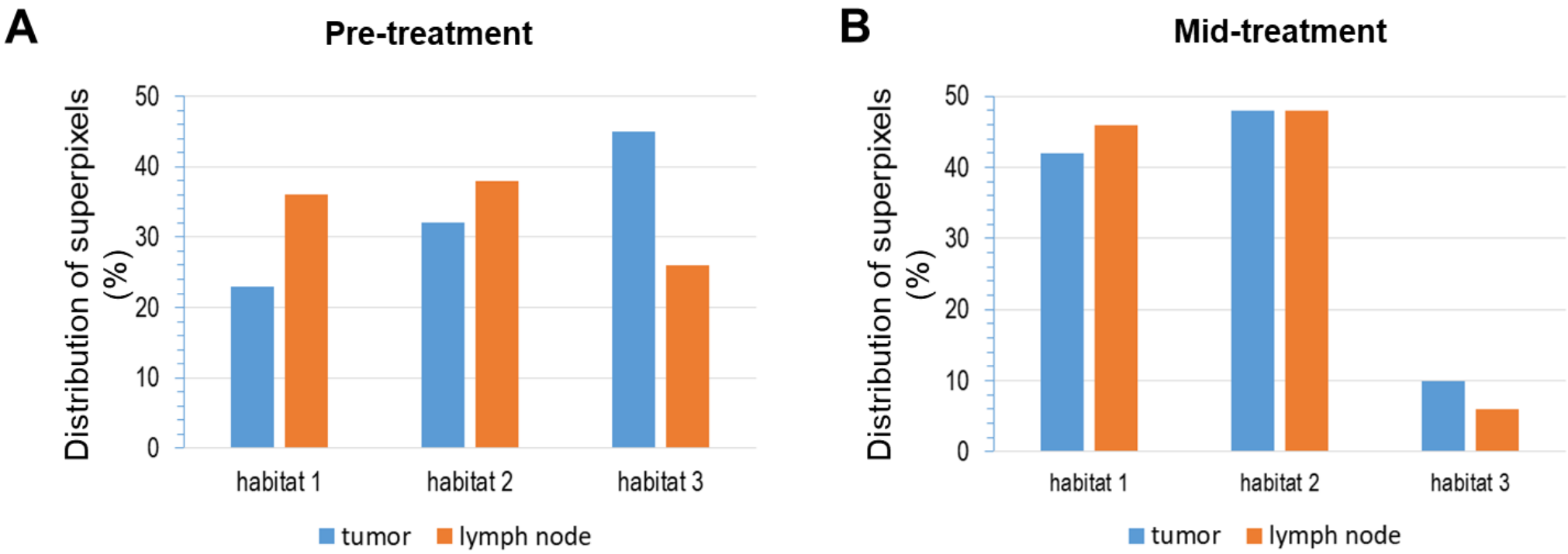




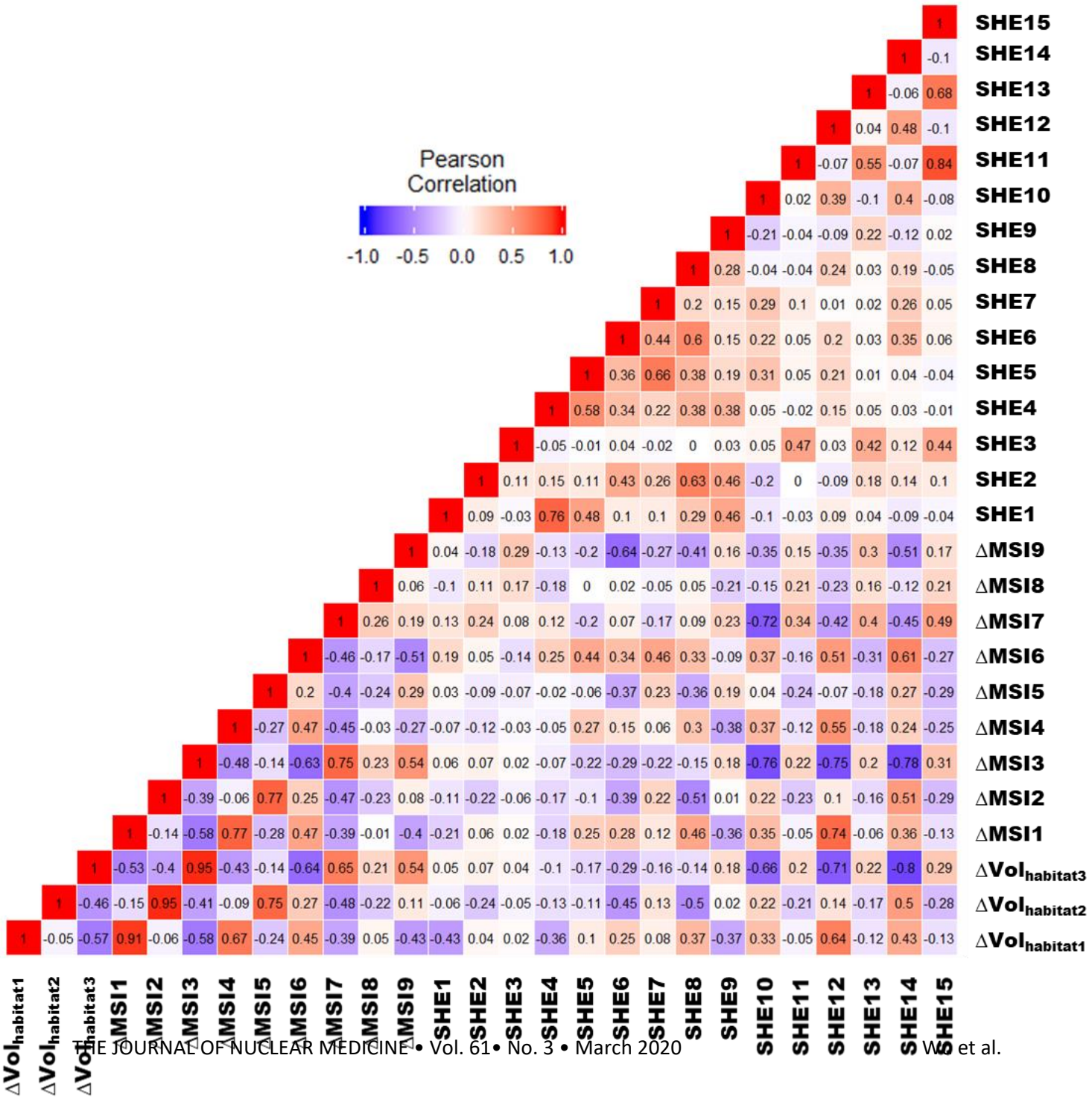
Supplemental Figure 6. Box-and-whisker plots show distribution of four PET/CT imaging parameters, including A) PET SUV, B) CT number, C) entropy of PET SUV, and D) entropy of CT number for three intratumoral habitats based on the training cohort. PET = positron emission tomography, SUV = standardized uptake value, CT = computed tomography. P values were obtained with Student t test. * = P < .05, ** = P < .001, *** = P < .0001.



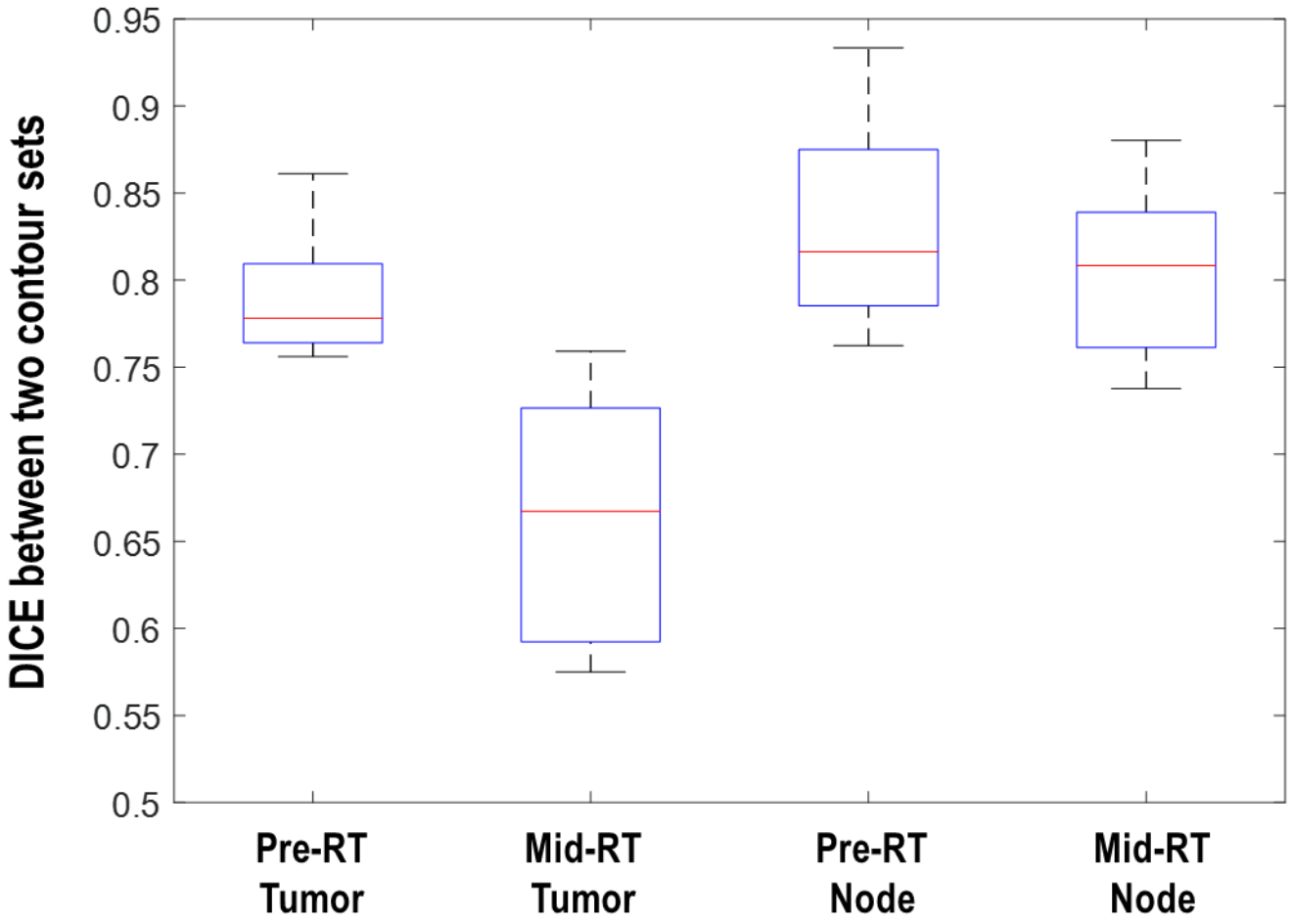
Supplemental Figure 7. Box-and-whisker plots show distribution of four PET/CT imaging parameters, including A) PET SUV, B) CT number, C) entropy of PET SUV, and D) entropy of CT number for three intratumoral habitats based on the validation cohort. PET = positron emission tomography, SUV = standardized uptake value, CT = computed tomography. P values were obtained with Student t test. * = P < .05, ** = P < .001, *** = P < .0001.



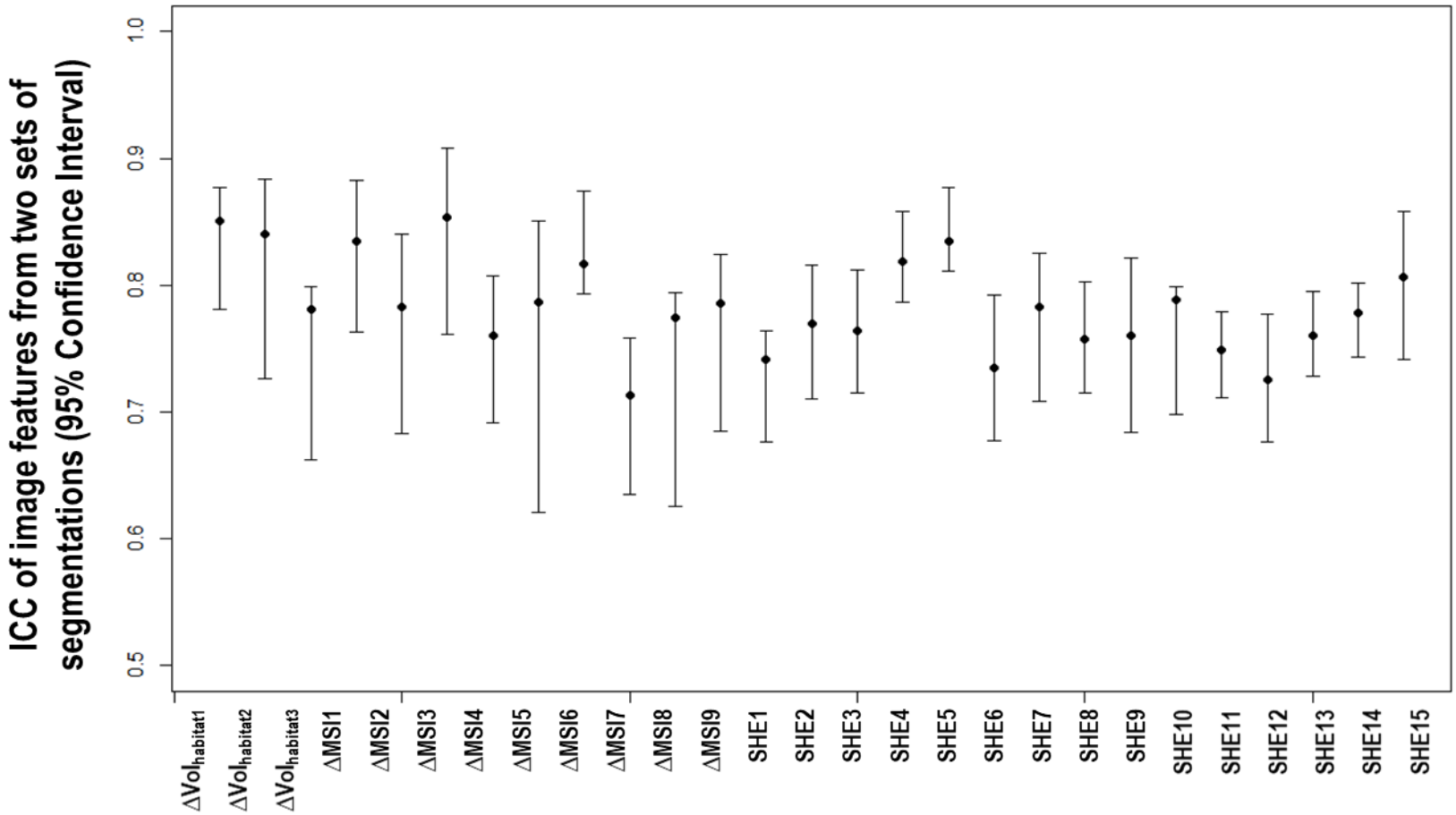
Supplemental Figure 8. Distribution of three habitat in tumor and lymph node regions for A) pre-treatment and B) mid-treatment



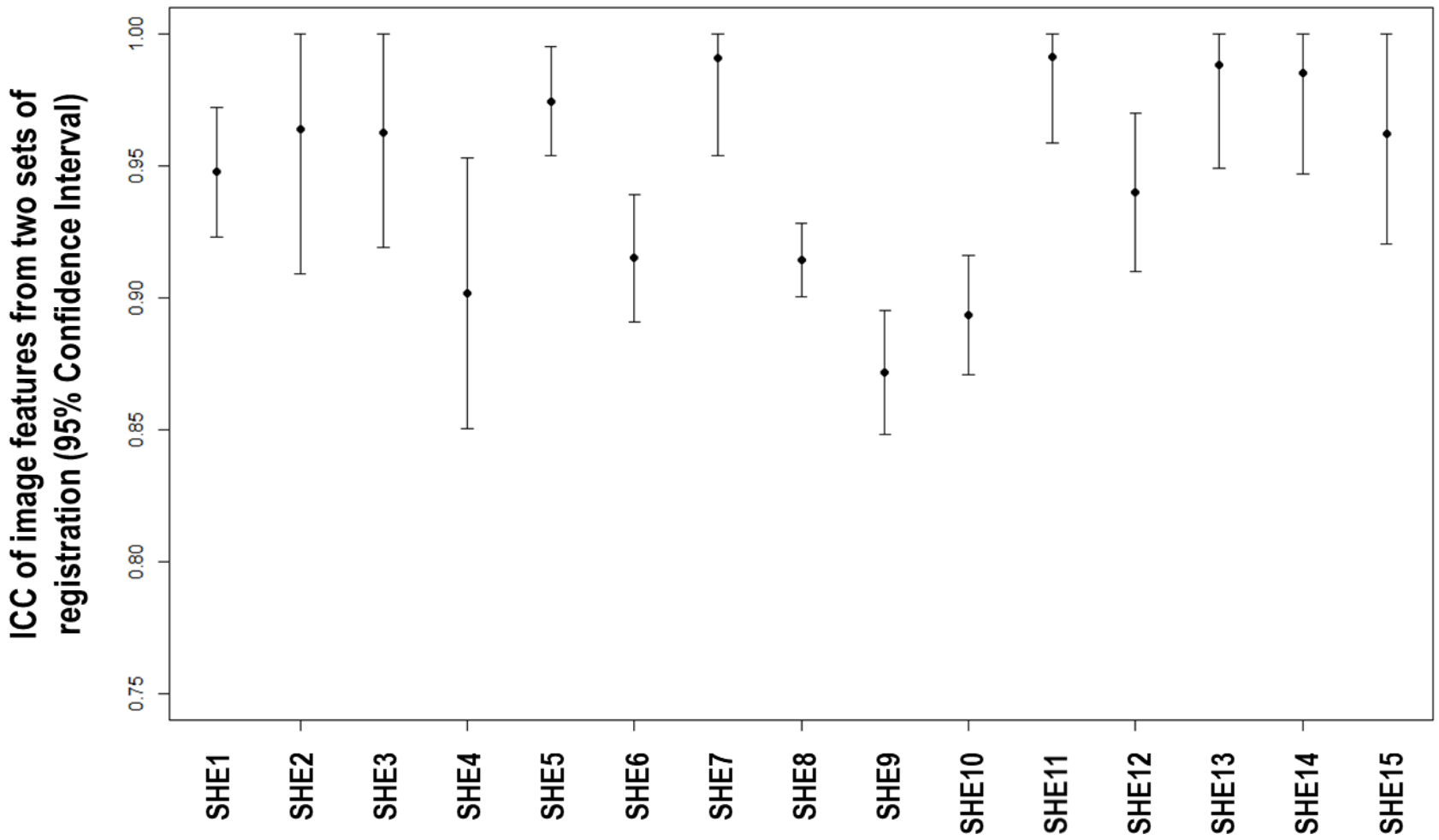
Supplemental Figure 9. Pair-wise Pearson's correlation of 27 quantitative CT imaging features (see definition in Supplemental Table 1).



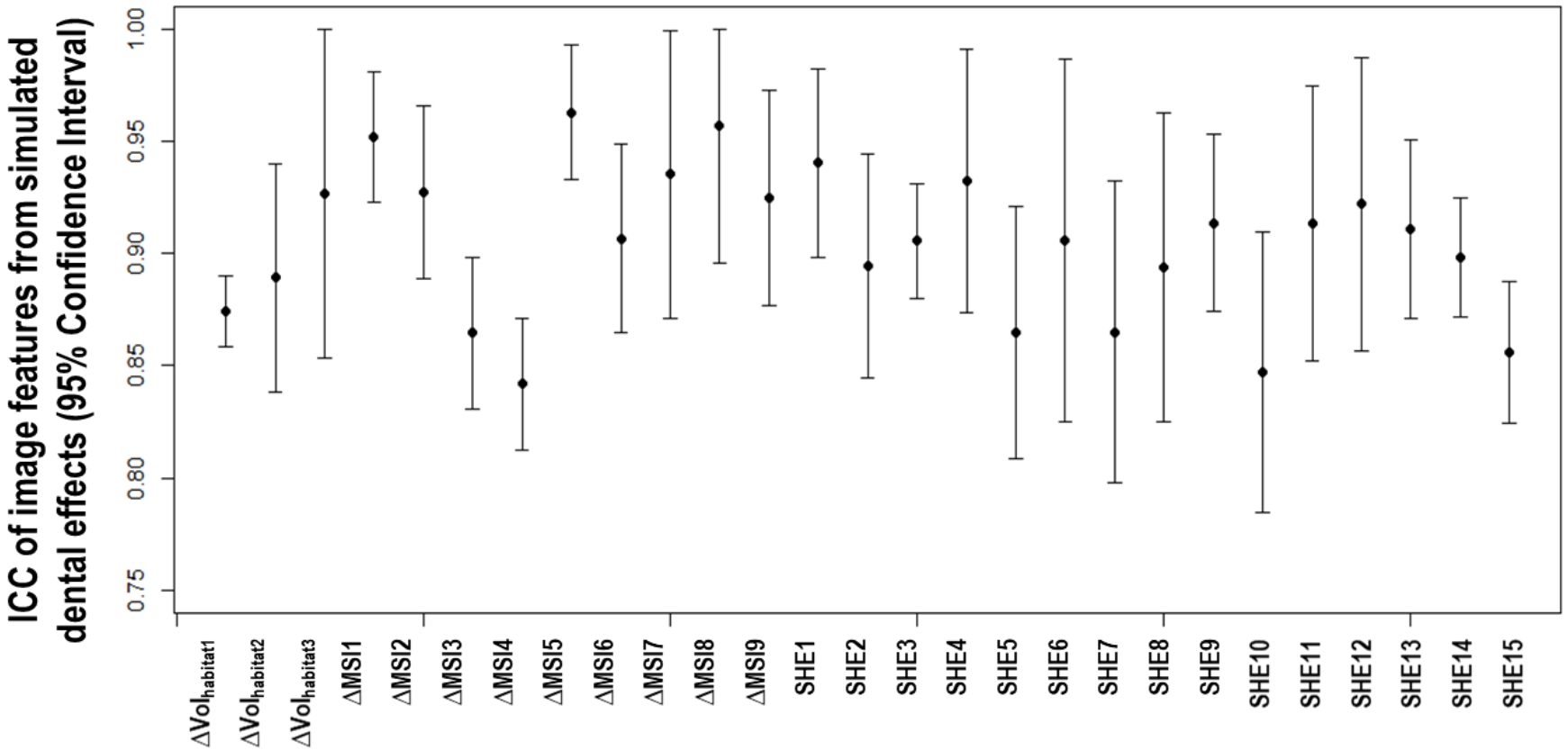
Supplemental Figure 10. The Dice coefficients of between contours by two radiation oncologists for tumor or node from pre-treatment and mid-treatment CT. Pre-RT = pre-treatment radiation therapy, Mid-RT = mid-treatment radiation therapy.



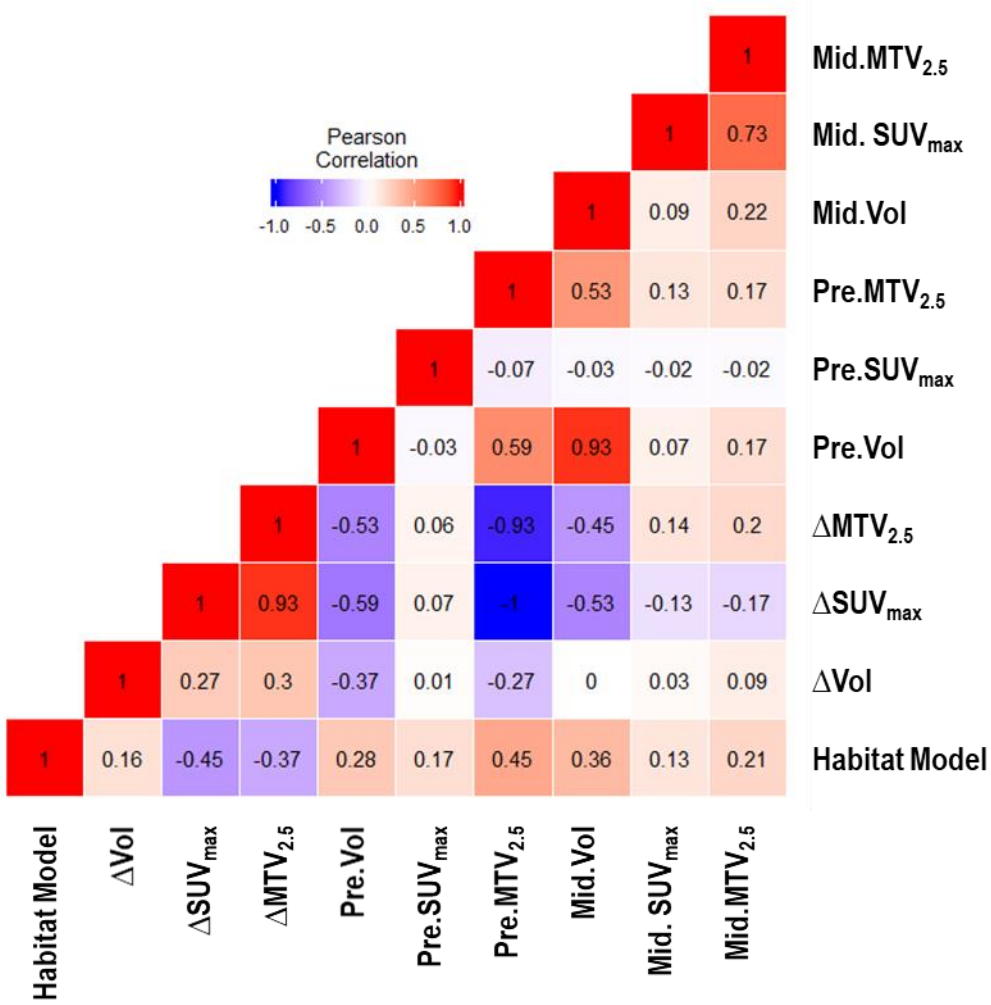
Supplemental Figure 11. Intraclass correlation coefficient (ICC) of proposed imaging features computed from two sets of 3D contours obtained from two board certified radiation oncologists.



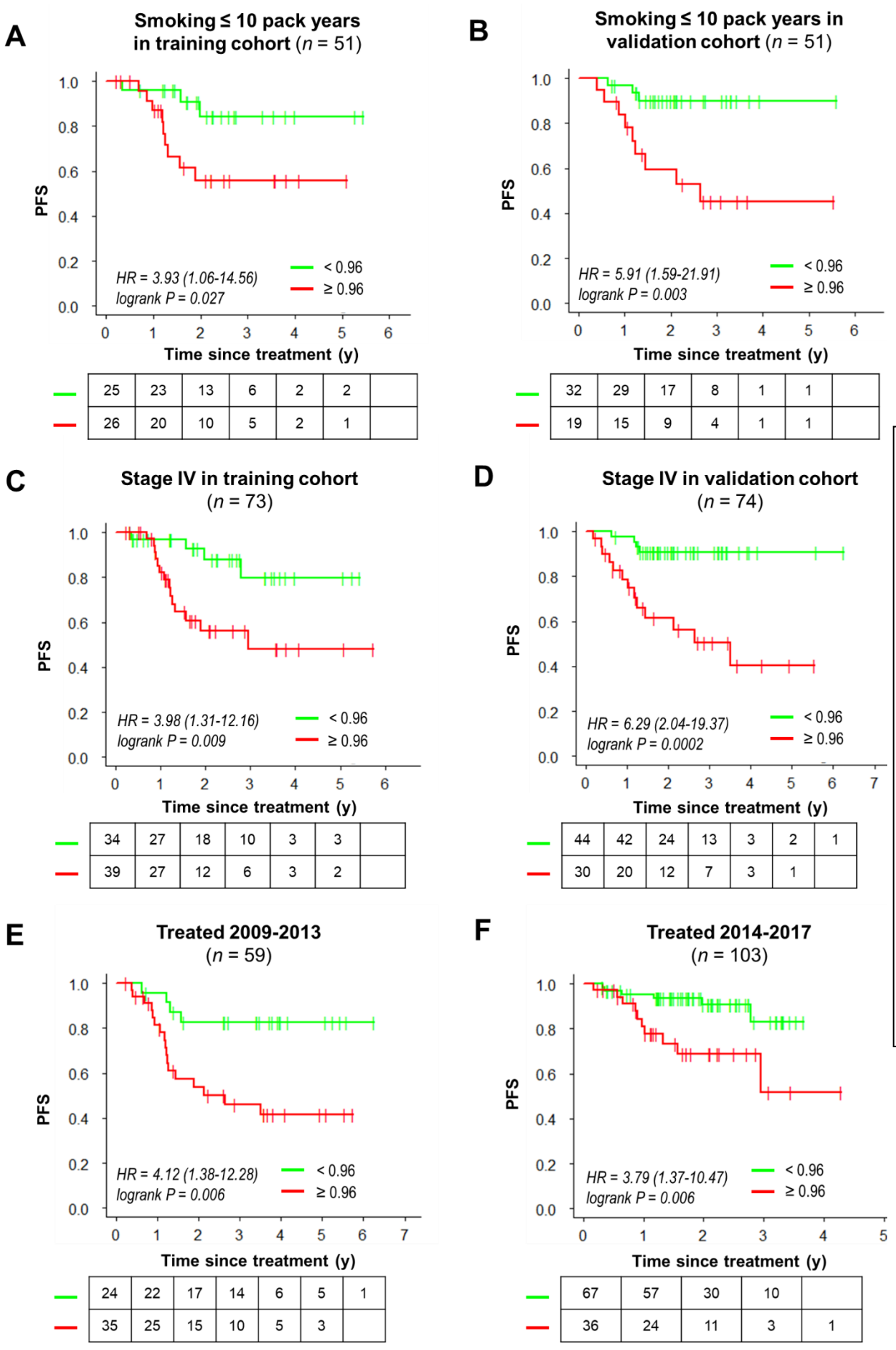
Supplemental Figure 12. Intraclass correlation coefficient (ICC) of proposed spatio-temporal habitat evolution features computed two sets of registration settings.



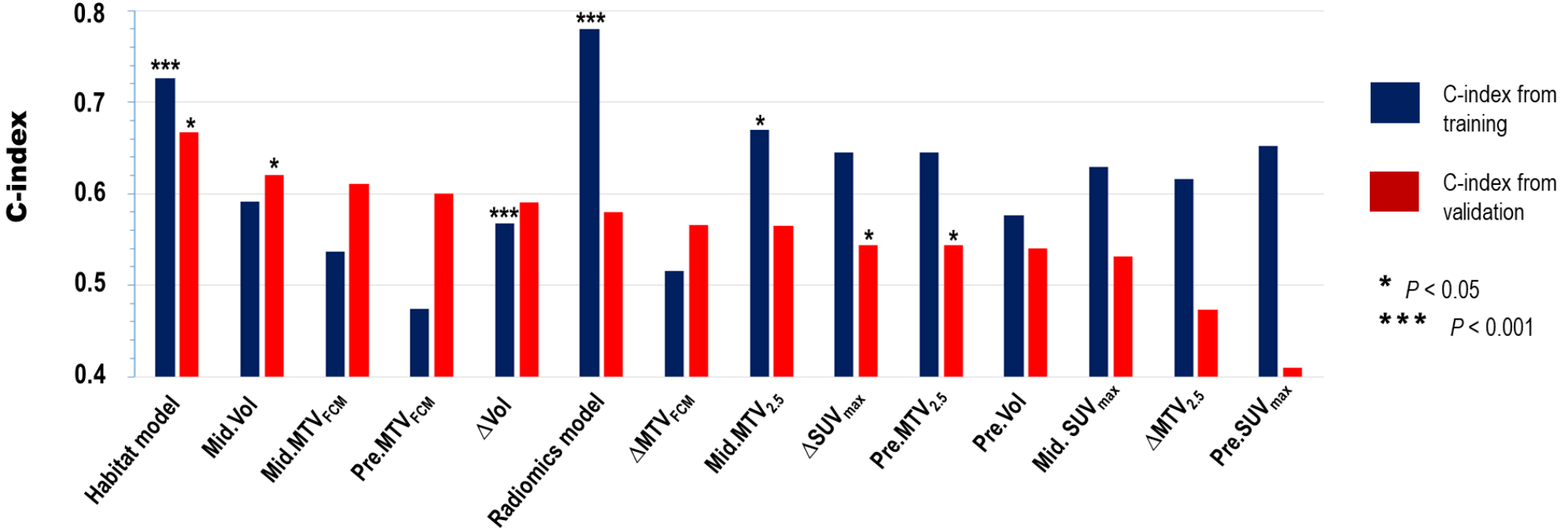
Supplemental Figure 13. Intraclass correlation coefficient (ICC) of proposed imaging features computed from whole 3D contours and excluded randomly selected slices, to assess their robustness with dental artifacts.



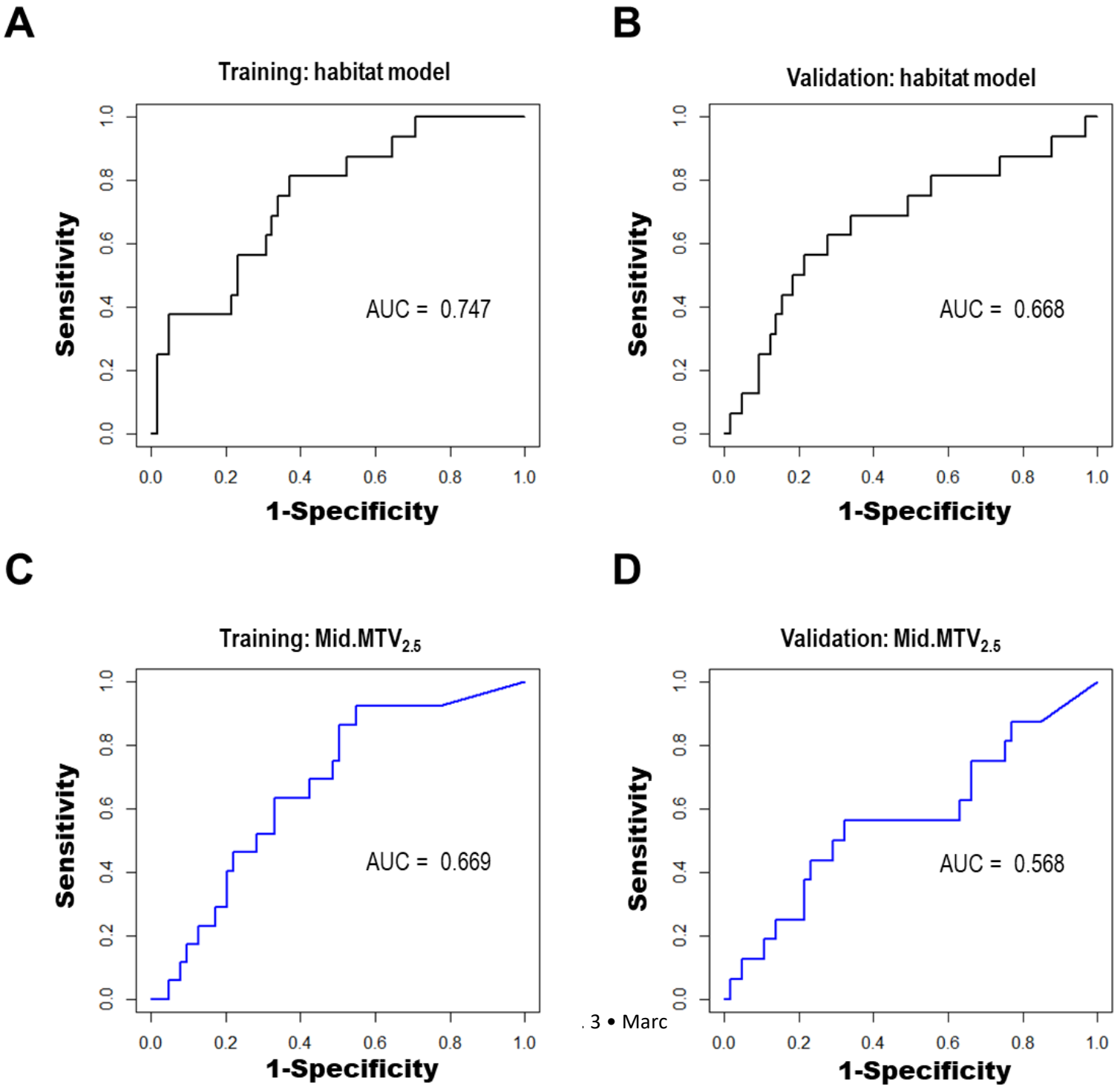
Supplemental Figure 14. Pair-wise Pearson's correlation of 10 imaging features including the final Cox model (Habitat Model) as well as 9 conventional quantitative PET and CT imaging features.



Supplemental Figure 15. Kaplan-Meier curves of progression free survival (PFS). Patients are stratified by median risk score according to the proposed habitat evolution-based imaging signature in the training cohort. Plots are for: A) smoking ⩽ 10 pack years subgroup from the training set, B) smoking ⩽ 10 pack years subgroup from the validation set, C) stage IV in the training set, D) stage IV in the validation set, E) patients treated in 2009-2013, and F) patients treated in 2014-2017. HR = hazard ratio.



Supplemental Figure 16. Accuracy of predicting PFS as measured by C-index for the habitat-based imaging signature, compared with all conventional imaging features from PET and CT images extracted respectively at pre-treatment, mid-treatment or change (mid - pre). The comparison was carried out separately in training and validation cohorts. The features were ranked by C-index of validation. Abbreviations: MTV = metabolic tumor volume, SUV = standardized uptake value, Δ = mid - pre.



Supplemental Figure 17. ROC curves of predicting 2-year progression-free survival, where patients are stratified according to the risk scores of proposed habitat evolution-based imaging model for A) the training set and B) the validation set, as well as the MTV from mid-treatment PET for C) the training set and D) the validation set.

Supplemental Table 1. Details of the investigated imaging features from PET and CT, including habitat related features and conventional imaging features.

Habitat related features		
Type	**Details**	**No.**
Change for individual three habitat regions burden	$\Delta Vol_{habitat\ i}$ where i = 1,2,3, Δ is defined as mid - pre	3
Change for connected habitat region burden via multiregional spatial interaction (MSI) Matrix	a) Habitat 1↔Habitat 1 (ΔMSI1); b) Habitat 2 ↔Habitat 2 (ΔMSI2); c) Habitat 3 ↔Habitat 3 (ΔMSI3); d) Habitat 1↔Peritumor Parenchyma (ΔMSI4); e) Habitat 2 ↔ Peritumor Parenchyma (ΔMSI5); f) Habitat 2 ↔Habitat 1 (ΔMSI6); g) Habitat 3 ↔ Peritumor Parenchyma (ΔMSI7); h) Habitat 3 ↔Habitat 1 (ΔMSI8); i) Habitat 3 ↔Habitat 2 (ΔMSI9)	9
Habitat region progression via spatio-temporal habitat evolution (SHE) matrix	a) Pre-RT Habitat 1→Mid-RT Habitat 1 (SHE1); b) Pre-RT Habitat 2→Mid-RT Habitat 2 (SHE2); c) Pre-RT Habitat 3→Mid-RT Habitat 3 (SHE3); d) Pre-RT Habitat 1→Mid-RT Peritumor Parenchyma (SHE4); e) Pre-RT Peritumor Parenchyma→Mid-RT Habitat 1 (SHE5); f) Pre-RT Habitat 2→Mid-RT Peritumor Parenchyma (SHE6); g) Pre-RT Peritumor Parenchyma→Mid-RT Habitat 2 (SHE7); h) Pre-RT Habitat 2→Mid-RT Habitat 1 (SHE8); i) Pre-RT Habitat 1→Mid-RT Habitat 2 (SHE9); j) Pre-RT Habitat 3→Mid-RT Peritumor Parenchyma (SHE10); k) Pre-RT Peritumor Parenchyma→Mid-RT Habitat 3 (SHE11); l) Pre-RT Habitat 3→Mid-RT Habitat 1 (SHE12); m) Pre-RT Habitat 1→Mid-RT Habitat 3 (SHE13); n) Pre-RT Habitat 3→Mid-RT Habitat 2 (SHE14); o) Pre-RT Habitat 2→Mid-RT Habitat 3 (SHE15);	15
Conventional features		
Type	**Details**	**No.**
Tumor and node burden from CT	a) Pre.Vol: volume from Pre-RT CT; b) Mid.Vol: volume from Mid-RT CT; c) ΔVol_{CT}: volume change	3
Metabolic activity of tumor and node from PET	Pre-RT PET: a) $Pre.SUV_{max}$; b) $Pre.MTV_{2.5}$; c) $Pre.MTV_{FCM}$ Mid-RT PET: d) $Mid.SUV_{max}$; e) $Mid.MTV_{2.5}$; f) $Mid.MTV_{FCM}$ Change: g) ΔSUV_{max}; h) $\Delta MTV_{2.5}$; i) ΔMTV_{FCM};	9

Radiomics signature	We applied pyradiomics to extract first order statistics (n=18), shape-based 3d (n=14), gray level co-occurrence matrix (n=24), gray level run length matrix (n=16), gray level size zone matrix (n=16), gray level dependence matrix (n=14), and neighboring gray tone difference matrix (n=5) from baseline, mid-treatment CT. Similarly we computed radiomics feature for PET SUV, including first order statistics (n=18), gray level co-occurrence matrix (n=24), gray level run length matrix (n=16), gray level size zone matrix (n=16), gray level dependence matrix (n=14), and neighboring gray tone difference matrix (n=5). Moreover, the changes of these radiomics feature (Δ = mid - pre) are also computed. This results in totally 321 features from CT and 279 features from PET SUV. Given the 600 radiomics features, LASSO and Cox regression was used to build the signature within the training cohort.	1

Supplemental Table 2. Details of the four habitat related features in the trained Cox model for predicting PFS.

Feature	Coefficient	HR (95% CI)	*P*-value
ΔMSI4	0.35	1.41 [1.03-1.94]	.030
ΔMSI8	-0.15	0.86 [0.70-1.06]	.153
SHE3	0.38	1.47 [1.13-1.91]	.004
SHE7	0.26	1.30 [1.04-1.62]	.019

Supplemental Table 3. Details of the four radiomics features in the trained Cox model for predicting PFS.

Feature	Coefficient	HR (95% CI)	*P*-value
Original_firstorder_Median (ΔCT)	-0.67	0.34 – 0.77	0.001
Original_ngtdm_Contrast (PET Mid-RT)	0.4	1.01 – 2.39	0.045
Original_glrlm_LongRunLowGrayLevelEmphasis (PET Pre-RT)	-34.38	0 – 14.98	0.068
Original_shape_Sphericity ((ΔCT)	0.89	1.36 – 4.32	0.003