

Supplemental Table S1. Radiomics Model Performances Generated by Different Classifier

Prediction of RECIST response using top 10 LASSO features on 57 patients			
Classifier Type	Accuracy	Sensitivity	Specificity
Support Vector Machine	71.93%	75.68%	65.00%
Decision Tree	75.44%	78.38%	70.00%
Random Forest	77.19%	81.08%	70.00%
XGBoost	94.74%	97.30%	90.00%
Prediction of irRECIST response using top 10 LASSO features on 57 patients			
Classifier Type	Accuracy	Sensitivity	Specificity
Support Vector Machine	73.68%	75.76%	70.83%
Decision Tree	70.18%	72.73%	66.67%
Random Forest	77.19%	81.82%	70.83%
XGBoost	94.74%	93.94%	95.83%

Supplemental Table S2. Results of Normality Test Looking for Distribution of Variables

Top 10 LASSO features from RECIST model		Top 10 LASSO features from irRECIST model		Demographic Variable	
Feature	P-value	Feature	P-value	Age	P-value=0.00215
P_F7	0.57112	P_F89	0.88205		
P_F84	0.23476	P_F187	0.13816		
P_F269	0.23184	P_F269	0.57944		
P_F289	0.3061	P_F289	0.35358		
PV_F35	0.38787	PV_F24	0.53648		
PV_F89	0.91328	PV_F81	0.10991		
PV_F109	0.13542	PV_F270	0.21483		
Ar_F152	0.05004	Ar_F152	0.43135		
Ar_F154	0.63007	ArV_F36	0.28886		
Ar_F238	0.38352	ArV_F247	0.37607		

Supplemental Table S3A. Multivariate Analysis on the 10 Features (Taken from Patients Assessed by RECIST) with Clinical Variables

Demographic Variable	MANOVA test results					
	Df	Pillai	approx F	num Df	den Df	Pr(>F)
No. of Prior Therapies [> 2 or <= 2] Residuals	1	0.14697	0.79251	10	46	0.636
	55					
Race Residuals	4	0.49823	0.65448	40	184	0.9433
	52					
Gender Residuals	1	0.28036	1.7921	10	46	0.08905
	55					
Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1						

Supplemental Table S3B: Multivariate Analysis on the 10 Features (Taken from Patients Assessed by irRECIST) with Clinical Variables

Demographic Variable	MANOVA test results					
	Df	Pillai	approx F	num Df	den Df	Pr(>F)
No. of Prior Therapies [> 2 or <= 2] Residuals	1	0.24446	1.4883	10	46	0.1745
	55					
Race Residuals	4	0.31563	0.39407	40	184	0.9996
	52					
Gender Residuals	1	0.31292	2.095	10	46	0.0444*
	55					
Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1						

Df stands for the degrees of freedom.
Pillai stands for the value of the test statistic in this test.
approx F stands for an approximated F statistic to this test.
num Df stands for the numerator degrees of freedom of the F-ratio corresponding to this test.
den Df stands for the denominator degrees of freedom of the F-ratio corresponding to this test.
Pr(>F) stands for the p-value of this test.

Supplemental Table S4A: Univariate Analysis on the 10 Features (Taken from Patients Assessed by RECIST) with Clinical Covariates

Feature	One-way ANOVA test			Pearson Correlation Coefficient [r]
	Gender	Race	Number of Prior Therapies [> 2 or ≤ 2]	Age at Initial Diagnosis, in years
P_F7	F(1,55) = 0.069, p = 0.793	F(4,52) = 0.980, p = 0.426	F(1,55) = 0.253, p = 0.617	0.103
P_F84	F(1,55) = 4.309, p = 0.043*	F(4,52) = 0.330, p = 0.856	F(1,55) = 0.854, p = 0.359	-0.146
P_F269	F(1,55) = 0.591, p = 0.445	F(4,52) = 0.045, p = 0.995	F(1,55) = 0.722, p = 0.399	-0.002
P_F289	F(1,55) = 0.767, p = 0.385	F(4,52) = 0.280, p = 0.889	F(1,55) = 0.442, p = 0.508	0.012
PV_F35	F(1,55) = 0.229, p = 0.634	F(4,52) = 0.559, p = 0.693	F(1,55) = 0.035, p = 0.853	-0.084
PV_F89	F(1,55) = 0.555, p = 0.459	F(4,52) = 1.593, p = 0.190	F(1,55) = 0.030, p = 0.862	0.032
PV_F109	F(1,55) = 0.031, p = 0.862	F(4,52) = 0.963, p = 0.436	F(1,55) = 0.029, p = 0.865	0.012
Ar_F152	F(1,55) = 6.243, p = 0.015*	F(4,52) = 0.217, p = 0.928	F(1,55) = 0.049, p = 0.826	-0.039
Ar_F154	F(1,55) = 0.000, p = 0.988	F(4,52) = 0.993, p = 0.419	F(1,55) = 0.154, p = 0.696	0.240
Ar_F238	F(1,55) = 0.038, p = 0.846	F(4,52) = 0.341, p = 0.849	F(1,55) = 1.814, p = 0.183	0.048

Supplemental Table S4B: Univariate Analysis on the 10 Features (Taken from Patients Assessed by irRECIST) with Clinical Covariates

Feature	One-way ANOVA test			Pearson Correlation Coefficient [r]
	Gender	Race	No. of Prior Therapies [> 2 or ≤ 2]	Age at Initial Diagnosis, in years
P_F7	F(1,55) = 0.069, p = 0.793	F(4,52) = 0.980, p = 0.426	F(1,55) = 0.253, p = 0.617	0.103
P_F84	F(1,55) = 4.309, p = 0.043*	F(4,52) = 0.330, p = 0.856	F(1,55) = 0.854, p = 0.359	-0.146
P_F269	F(1,55) = 0.591, p = 0.445	F(4,52) = 0.045, p = 0.995	F(1,55) = 0.722, p = 0.399	-0.002
P_F289	F(1,55) = 0.767, p = 0.385	F(4,52) = 0.280, p = 0.889	F(1,55) = 0.442, p = 0.508	0.012
PV_F35	F(1,55) = 0.229, p = 0.634	F(4,52) = 0.559, p = 0.693	F(1,55) = 0.035, p = 0.853	-0.084
PV_F89	F(1,55) = 0.555, p = 0.459	F(4,52) = 1.593, p = 0.190	F(1,55) = 0.030, p = 0.862	0.032
PV_F109	F(1,55) = 0.031, p = 0.862	F(4,52) = 0.963, p = 0.436	F(1,55) = 0.029, p = 0.865	0.012
Ar_F152	F(1,55) = 6.243, p = 0.015*	F(4,52) = 0.217, p = 0.928	F(1,55) = 0.049, p = 0.826	-0.039
Ar_F154	F(1,55) = 0.000, p = 0.988	F(4,52) = 0.993, p = 0.419	F(1,55) = 0.154, p = 0.696	0.240
Ar_F238	F(1,55) = 0.038, p = 0.846	F(4,52) = 0.341, p = 0.849	F(1,55) = 1.814, p = 0.183	0.048

Supplemental Table S5: Univariate Association for Independent Association Between Radiomics Features and Treatment Response

Top 10 LASSO features from RECIST model		Top 10 LASSO features from irRECIST model	
Radiomics Feature	Welch Two Sample t-test	Radiomics Feature	Welch Two Sample t-test
P_F7	t = 0.75051, df = 31.4, p-value = 0.4585	P_F89	t = 2.7162, df = 44.815, p-value = 0.009348*
P_F84	t = 0.28333, df = 37.84, p-value = 0.7785	P_F187	t = 1.9374, df = 35.734, p-value = 0.06064
P_F269	t = 2.2598, df = 46.958, p-value = 0.02851*	P_F269	t = 2.1697, df = 49.883, p-value = 0.03482*
P_F289	t = 2.2383, df = 41.312, p-value = 0.03065*	P_F289	t = 2.3428, df = 46.87, p-value = 0.02344*
PV_F35	t = 2.4684, df = 19.589, p-value = 0.02292*	PV_F24	t = 1.8515, df = 29.119, p-value = 0.07426
PV_F89	t = 2.3159, df = 21.968, p-value = 0.03028*	PV_F81	t = 2.1749, df = 38.588, p-value = 0.03583*
PV_F109	t = 2.8168, df = 19.894, p-value = 0.01069*	PV_F270	t = -0.96974, df = 50.109, p-value = 0.3368
Ar_F152	t = -0.77189, df = 33.756, p-value = 0.4456	Ar_F152	t = -0.51419, df = 48.364, p-value = 0.6095
Ar_F154	t = 0.67226, df = 31.596, p-value = 0.5063	ArV_F36	t = 0.32664, df = 39.354, p-value = 0.7457
Ar_F238	t = -1.1571, df = 32.633, p-value = 0.2556	ArV_F247	t = 1.4169, df = 39.817, p-value = 0.1643

Supplemental Table S6: A univariate cox regression model looking for association of OS with each of the top 10 texture features from RECIST model.

Texture Feature	Beta	Wald test	p-value
P_F7	-0.65	0.57	0.45
P_F84	0.25	0.05	0.82
P_F269	2.7	7.6	0.0058*
P_F289	2.2	5.9	0.015*
PV_F35	3.9	7.4	0.0063*
PV_F89	5.4	5.6	0.018*
PV_F109	3.5	6	0.014*
Ar_F152	0.082	0.01	0.94
Ar_F154	1.2	2.5	0.11
Ar_F238	-2.5	4.1	0.044*

Supplemental Table S 7: A multivariate cox regression using significant texture features from univariate cox models of RECIST with overall survival.

N= 57, Number of events= 44							
	coef	exp(-coef)	se(coef)	lower .95	upper .95	z	Pr(> z)
P_F269	0.6866	0.5033	3.4334	0.002375	1662.318	0.2	0.8415
P_F289	1.7488	0.173975	3.2569	0.009711	3402.262	0.537	0.5913
PV_F35	5.0281	0.006551	2.5573	1.016074	22931.01	1.966	0.0493 *
PV_F89	-0.6446	1.905239	4.37	0.0001	2753.112	-0.148	0.8827
PV_F109	-2.0183	7.52534	2.9064	0.000446	39.574	-0.694	0.4874
Ar_F238	-1.7083	5.519334	1.4691	0.010177	3.226	-1.163	0.2449
Concordance = 0.668 (se = 0.055)							
Likelihood ratio test = 22.27 on 6 df, p=0.001							
Wald test = 13.84 on 6 df, p=0.03							
Score (log-rank) test = 15.61 on 6 df, p=0.02							
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1							

Supplemental Table S8: The results of log-rank test performed on each of the 10 radiomic texture feature from RECIST model with overall survival.

Texture Feature	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V	Chisq	df	p-value
P_F7								
High	21	15	16.8	0.184	0.299	0.3	1	0.58
Low	36	29	27.2	0.113	0.299			
P_F84								
High	37	29	27.1	0.134	0.356	0.4	1	0.55
Low	20	15	16.9	0.215	0.356			
P_F269								
High	37	30	24.2	1.41	3.28	3.3	1	0.07
Low	20	14	19.8	1.72	3.28			
P_F289								
High	21	18	9.56	7.46	10.2	10.2	1	0.0014*
Low	36	26	34.44	2.07	10.2			
PV_F35								
High	37	31	21.8	3.88	8.08	8.1	1	0.0045*
Low	20	13	22.2	3.81	8.08			
PV_F89								
High	37	31	23.3	2.56	5.76	5.8	1	0.016*
Low	20	13	20.7	2.88	5.76			
PV_F109								
High	37	30	26.7	0.412	1.06	1.1	1	0.3
Low	20	14	17.3	0.635	1.06			
Ar_F152								
High	16	11	16.8	2.01	3.43	3.4	1	0.064
Low	41	33	27.2	1.24	3.43			
Ar_F154								
High	41	33	29.4	0.447	1.38	1.4	1	0.24
Low	16	11	14.6	0.898	1.38			
Ar_F238								
High	16	10	18.2	3.68	6.51	6.5	1	0.011*
Low	41	34	25.8	2.59	6.51			

Supplemental Table S9: A univariate cox regression model looking for association of OS with each of the top 10 texture features from irRECIST model.

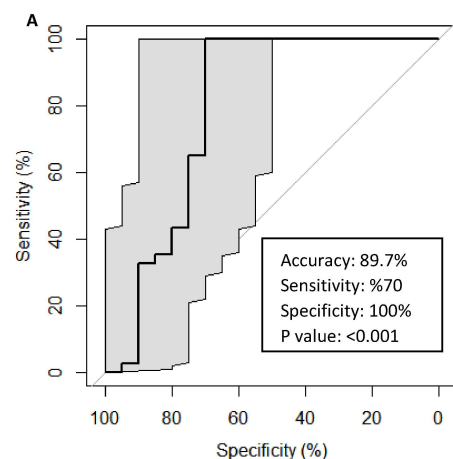
Texture Feature	Beta	Wald test	p-value
P_F89	0.9	1.2	0.27
P_F187	-0.27	0.11	0.74
P_F269	2.7	7.6	0.0058*
P_F289	2.2	5.9	0.015*
PV_F24	2.1	4.9	0.027*
PV_F81	6.1	9	0.0026*
PV_F270	-0.52	0.26	0.61
Ar_F152	0.082	0.01	0.94
ArV_F36	3.1	4	0.044*
ArV_F247	2.8	3.6	0.058

Supplemental Table S10: A multivariate cox regression using significant texture features from univariate cox models of irRECIST with overall survival.

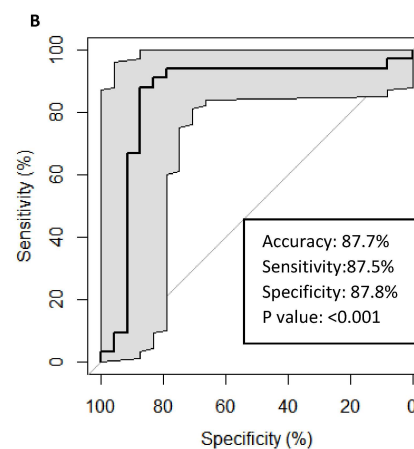
N= 57, Number of events= 44							
	coef	exp(-coef)	se(coef)	lower .95	upper .95	z	Pr(> z)
P_F269	0.9416	0.390005	3.5546	0.002417	2720.2	0.265	0.79109
P_F289	1.2886	0.275652	3.2142	0.006664	1975	0.401	0.68848
PV_F24	1.0991	0.333162	2.2067	0.039717	226.8	0.498	0.61843
PV_F81	5.8544	0.002867	2.1692	4.967427	24486.6	2.699	0.00696 **
ArV_F36	3.1774	0.041696	1.4094	1.514479	379.8	2.254	0.02417 *
Concordance = 0.701 (se = 0.046)							
Likelihood ratio test = 26.2 on 5 df, p=8e-05							
Wald test = 17.18 on 5 df, p=0.004							
Score (log-rank) test = 21.54 on 5 df, p=6e-04							
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1							

Supplemental Table S11: The results of log-rank test performed on each of the 10 radiomic texture feature from irRECIST model with overall survival.

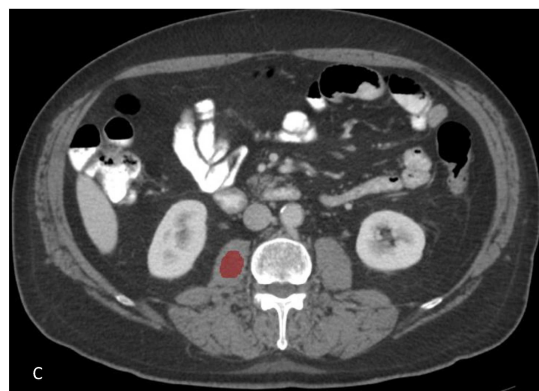
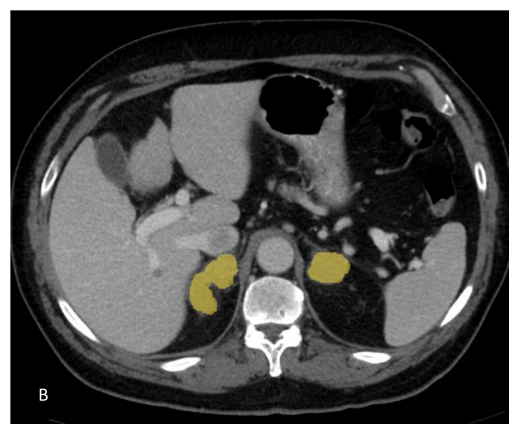
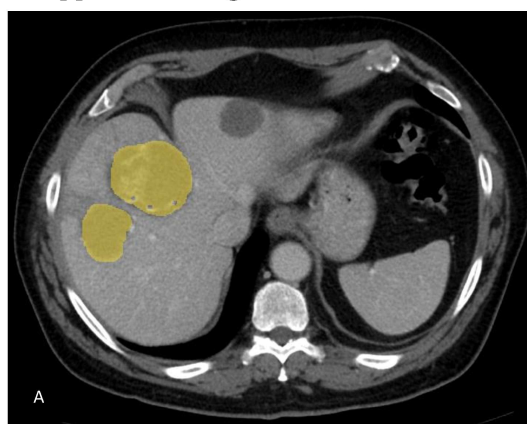
Texture Feature	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V	Chisq	df	p-value
P_F89								
High	37	30	26.2	0.54	1.35	1.4	1	0.24
Low	20	14	17.8	0.797	1.35			
P_F187								
High	37	30	30.5	0.00754	0.0248	0	1	0.87
Low	20	14	13.5	0.01701	0.0248			
P_F269								
High	37	30	24.2	1.41	3.28	3.3	1	0.07
Low	20	14	19.8	1.72	3.28			
P_F289								
High	21	18	9.56	7.46	10.2	10.2	1	0.0014*
Low	36	26	34.44	2.07	10.2			
PV_F24								
High	37	31	26	0.945	2.35	2.3	1	0.13
Low	20	13	18	1.37	2.35			
PV_F81								
High	37	31	22.3	3.4	7.24	7.2	1	0.0071*
Low	20	13	21.7	3.49	7.24			
PV_F270								
High	37	28	33.4	0.869	3.92	3.9	1	0.048*
Low	20	16	10.6	2.734	3.92			
Ar_F152								
High	16	11	16.8	2.01	3.43	3.4	1	0.064
Low	41	33	27.2	1.24	3.43			
ArV_F36								
High	41	33	30.7	0.168	0.569	0.6	1	0.45
Low	16	11	13.3	0.39	0.569			
ArV_F247								
High	41	34	29.6	0.658	2.07	2.1	1	0.15
Low	16	10	14.4	1.35	2.07			

Supplemental Figure S1.

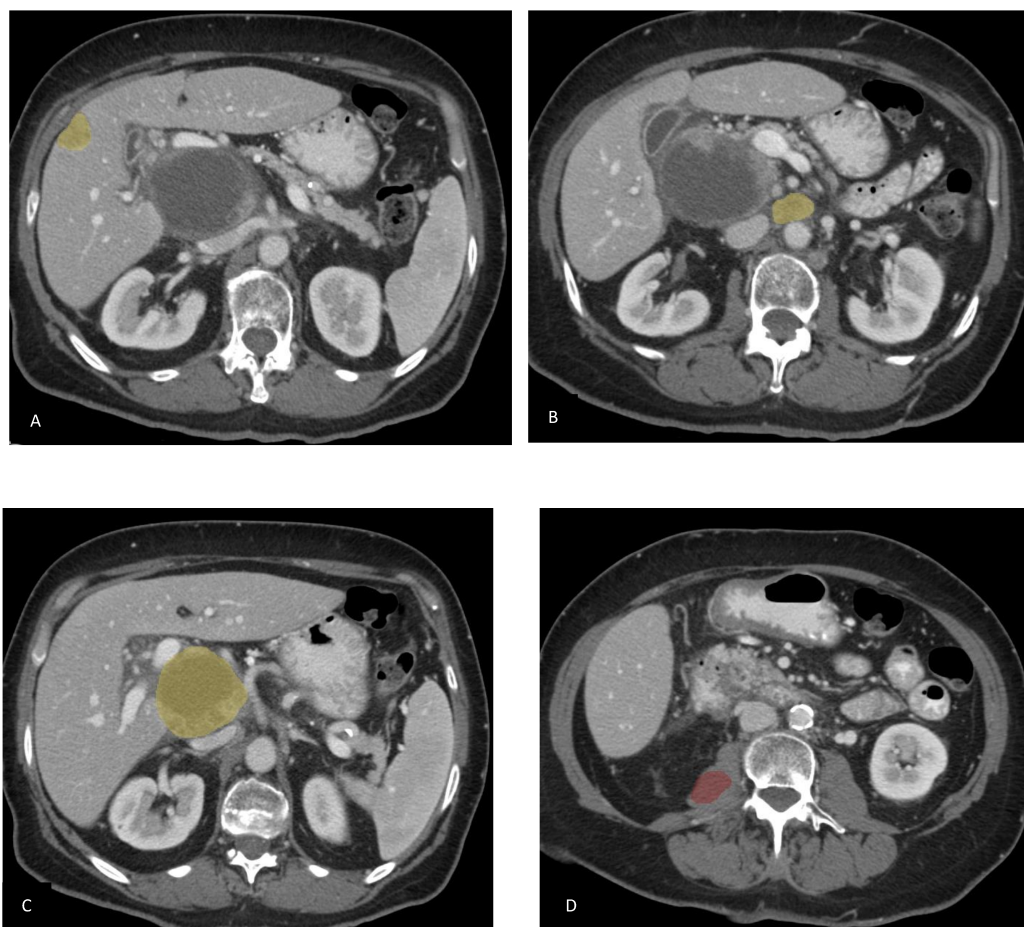
Supplemental Figure S1A. ROC curve representing the performance of the radiomic analysis when feature selection and LOOCV were performed simultaneously in patients assessed by RECIST



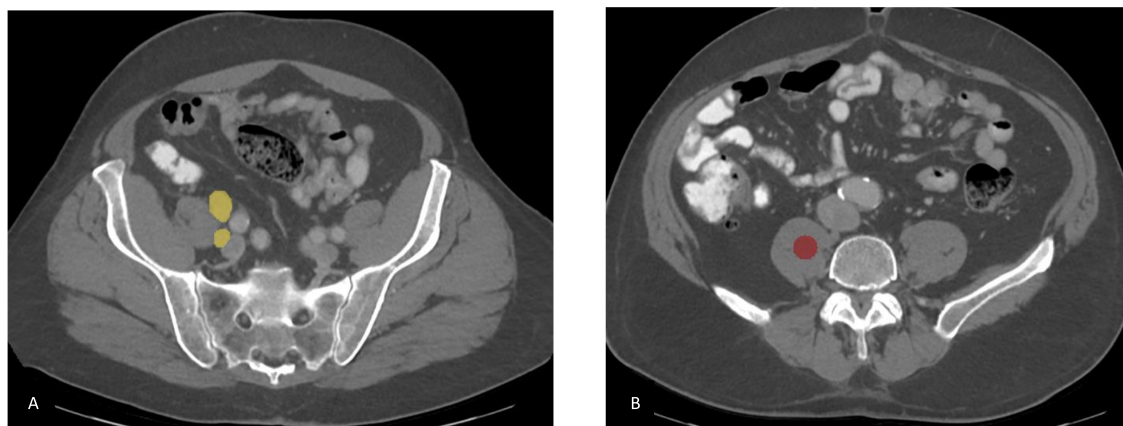
Supplemental Figure S1B. ROC curve representing the performance of the radiomic analysis when feature selection and LOOCV were performed simultaneously in patients assessed by irRECIST

Supplemental Figure S2.

Supplemental Figure S2. CT scans for a 68-year-old man with epithelioid hemangioendothelioma that did not respond to pembrolizumab. Target lesions in the liver (A) and bilateral suprarenal glands (B) were segmented and labeled with yellow. (C) A volume of interest, labeled with red, on the paraspinal muscle was taken for normalization.

Supplemental Figure S3.

Supplemental Figure S3. CT scans for a 60-year-old woman with carcinoma of unknown primary that did not respond to pembrolizumab. Target lesions (shown in yellow) in the pancreas (A), liver (B), and lymph nodes (C) were segmented. (D) A volume of interest on the right paraspinal muscle (shown in red) was taken for normalization.

Supplemental Figure S4.

Supplemental Figure S4. A 68-year old Caucasian male with metastatic squamous carcinoma of the skin, responding to pembrolizumab. (a) Right external iliac lymph nodes selected as target lesions. (b) A volume of interest with red color label on the right paraspinal muscle was taken for normalization