Supplementary material 1

1.1 Overview over the RF used in this study

SUV-Histogram	Maximum SUV (SUV_{max}), Peak SUV (SUV_{peak}), Minimum SUV (SUV_{max}), Mean SUV (SUV_{mean}), Volume (V), Total Lesion PSMA (TL-PSMA), Area under the curve of the cumulative SUV-volume histogram (aucCSH), Skewedness (S), Coefficient of variance (CoV), Kurtosis (K), Energy(E_H) and Entropy (Ent_H),
Geometry	Solidity (S), Eccentricity (Ecc), Long Diameter (LD), Percent Inactive (PI)
GLCM	Local Homogeneity (LH), Correlation (C_{CM}), Contrast (Con_{CM}), Energy (E_{CM}), Entropy (Ent_{CM}), Variance (Var_{CM}), Autocorrelation (Acor) and Dissimilarity (D)
GLRLM	Short Run Emphasis (SRE), Long Run Emphasis (LRE), Low Gray-Level Run Emphasis (LGRE), High Gray-Level Run Emphasis (HGRE) Short Run Low Gray-Level Emphasis (SRLGE), Short Run High Gray-Level Emphasis (SRHGE), Long Run Low Gray-Level Emphasis (LRLGE), Long Run High Gray-Level Emphasis (LRHGE), Gray-Level Non-uniformity (GLN), Run Percentage (RP) Length Non-uniformity (RLN)
GLSZM	Small Zone Emphasis (SZE), Large Zone Emphasis (LZE), Low Gray-Level Zone Emphasis (LGZE), High Gray-Level Zone Emphasis (HGZE), Small Zone Low Gray-Level Emphasis (SZLGE), Small Zone High Gray-Level Emphasis (SZHGE), Large Zone Low Gray-Level Emphasis(LZLGE), Large Zone High Gray-Level Emphasis (LZHGE), Zone-Size Non-uniformity (ZSN), Zone Percentage (ZP), Gray-Level Variance (GLV) and Zone-Size Variance (ZSV)
NGTDM	Coarseness (Coar), Contrast (Con_{NM}), Complexity (Comp) and Texture Strength (TS)

Abbreviations: SUV: standardized uptake values, GLCM: gray-level co-occurrence matrix, GLRLM: gray-level run length matrix, GLSZM: gray-level size zone matrix, NGTDM: neighborhood gray tone difference matrix

1.2 Results for all RF (analysis 1 and 2)

	Robust on different PET scanners	No intrinsic correlation with volume	Correlation of RF_a ($\rho>0.7$ and $p<0.05$)	Discrimination PCa and non-PCa_b
aucCSH	■	■	■	
Autocorrelation		■	■	■
Busyness	■		■	N/A
Coarseness	■		■	N/A
Complexity		■	■	■
ContrastCM		■	■	■
ContrastNG				N/A
CorrelationCM		■	■	
CoV		■	■	
Dissmilarity		■		■
Eccentricity	■	■	■	
Energy		■	■	■
EnergyCM		■	■	
EntropyCM		■	■	■
EntropyLog2	■	■		
GLN		■	■	
GLN2		■	■	
GLV	■	■		■
GLV2		■		
HGRE		■	■	■
HGZE		■	■	■
HomogeneityCM		■	■	■
Kurtosis	■	■		■
LGRE	■		■	N/A
LGZE	■		■	N/A
LongDiameter	■		■	N/A
LRE			■	N/A
LRHGE		■	■	■
LRLGE	■		■	N/A
LZE			■	N/A
LZHGE	■		■	N/A
LZLGE	■	■	■	
PercentInactive6	■		■	N/A
Qautocorrelation	■	■	■	■
QBusyness				N/A
Qcoarseness	■		■	N/A
QComplexity	■	■	■	■

QContrastCM				
QContrastNG				
QCorrelationCM				
Qdissmilarity				
QEnergyCM				
QEntropyCM				N/A
QGLN				
QGLN2				
QGLV				N/A
QGLV2				N/A
QHGRE				
QHGZE				
QHomogeneityCM				
QLGRE				
QLGZE				N/A
QLRE				
QLRHGE				
QLRLGE				
QLZE				
QLZHGE				N/A
QLZLGE				
QRLN				
QRLV				N/A
QRP				
QSRE				N/A
QSRHGE				
QSRLGE				N/A
QStrength				N/A
QSZE				
QSZHGE				
QSZLGE				N/A
QVarianceCM				
QZP				
QZSN				
QZSV				N/A
RLN				N/A
RLV				N/A
RP				N/A
Skewness				
Solidity				
SRE				N/A
SRHGE				
SRLGE				N/A
Strength				N/A
SUVmax				
SUVmean				
SUVmin				
SUVpeak				
SZE				
SZHGE				

SZLGE				N/A
TL-PSMA				N/A
VarianceCM				N/A
Vol.mm3				N/A
WFAutocorrelation				
WFBusyness				N/A
WFCoarseness				N/A
WFComplexity				
WFContrastCM				
WFContrastNG				N/A
WFCorrelationCM				
WFDissmilarity				
WFEnergyCM				
WFEntropyCM				
WFGLN				
WFGLN2				
WFGLV				
WFGLV2				
WFHGRE				
WFHGZE				
WFHomogeneityCM				
WFLGRE				N/A
WFLGZE				N/A
WFLRE				
WFLRHGE				
WFLRLGE				N/A
WFLZE				
WFLZHGE				N/A
WFLZLGE				
WFRLN				
WFRLV				
WFRP				
WFSRE				
WFSRHGE				
WFSRLGE				N/A
WFStrength				N/A
WFSZE				
WFSZHGE				
WFSZLGE				N/A
WFVarianceCM				N/A
WFZP				
WFZSN				
WFZSV				
ZP				
ZSN				
ZSV				N/A

Black: yes, white: no. Bold: RF with strong correlation with GS. Please notice: (a) Spearman correlations for RF extracted from GTV-Histo and GTV-Exp are shown. (b) RF are marked in black which could discriminate between

PCa and non-PCa tissue considering histology information and expert contours. Abbreviations: PCa: prostate cancer, RF: radiomic features. PET: positron emission tomography. Please see table 1.1 for explanation of RFs.

Supplementary Material 2

Radiomics features selection criteria: Phantom studies

In this study, the impact of different reconstruction parameters, scanner design and segmentation volumes on RF variability was evaluated with phantom measurements. PET images resulted from the EARL (ResEARch for Life – European Association of Nuclear Medicine initiative) accreditation measurement of the NEMA NU2 Phantom (NP) were used. To analyse the robustness of RF for different PET/CT systems 18 contours were delineated: 12 spheres (5.7-8.4 cc) were manually delineated in the background of NP and 6 spheres were delineated by applying a threshold of 40% of SUV_{max} of the NP. Wilcoxon Rank test was employed to identify the RF with comparable values for all the PET/CT systems. To analyse the intrinsic dependency of RF with the volume 102 contours (0.8-234 cc) were delineated in the background of the NP. Spearman's correlation coefficient was calculated and a strong correlation ($\rho>0.8$) was demanded to identify the RF with an intrinsic dependency of the volume. In Figure 2 results are summarized: 52/133 RF were robust to the three different PET systems and 86/133 RF had no strong correlation with the volume.

In order to minimize the impact of bed overlap on RF variability, bed overlap was optimized based on criteria of sensitivity and uniformity for each PET/CT system. The impact of the differences in time acquisition had been previously evaluated by our group using experimental heterogeneous phantoms (coefficient of variance>0.3) [59]. We performed 7 consecutive measurements: 3 measurements with an acquisition time of 1 min, one measurement of 10 min and 3 measurements of 15 min. The variability of RF across the 3 measurements with different time acquisitions (1, 10 and 15 min) was lower than across the 3 measurements of 1 min. We therefore concluded that the impact on RF variability due to the differences in time acquisitions presented in our study (2 to 3 min) could be considered as not statistically significant.

Supplementary Table 1. Patients' characteristics, prospective cohort

	Age (y)	PSA (ng/ml)	TNM	Volume of GTV-Exp / GTV-Histo / GTV-40% (ml)	Gleason score
1	67	6.07	pT3a pN0 cM0	8.3 / 12.4 / 7.5	7a (3+4) and 7b (4+3)
2	61	10.57	pT2c pN0 cM0	5.1 /5 / 5.1	7a (3+4) and 7a (3+4)
3	52	51.13	pT3b pN1 cM0	25.5 / 24.5 / 15.1	9 (5+4)
4	60	49	pT2c pN1 cM0	40.3 / 21.5 / 7.3	7a (3+4)
5	73	25.52	pT2c pN0 cM0	3.9 / 3.2 / 3.2	7a (3+4)
6	59	9.15	pT2c pN0 cM0	4.8 / 3.1 / 1.9	7b (4+3)
7	74	8.82	pT2c pN0 cM0	10.8/ 2.5 / 0	7a (3+4)
8	74	15	pT2c pN0 cM0	1.5 / 2.7 / 3.3	7a (3+4) and 9 (5+4)
9	62	47.17	pT3b pN1 cM0	126.1/ 116.5/ 26.6	8 (4+4)
10	49	5.57	pT2c pN0 cM0	0 / 0 / 0	6 (3+3)
11	68	11.03	pT3a pN0 cM0	10.8/ 2.9 / 1.4	7a (3+4) and 7a (3+4)
12	51	17.4	pT3a pN0 cM0	6.1/ 6.5 / 3.9	7a (3+4) and 7b (4+3)
13	48	23	pT3b pN1 cM0	22.1 / 20.3 / 0	7b (4+3)
14	76	20.7	pT2c pN0 cM0	15 / 13.9 / 2.7	7b (4+3) and 7b (4+3)
15	59	15.8	pT3b pN1 cM0	28.8 / 18.2 / 2.7	9 (4+5)
16	73	40	pT3a pN1 cM0	19.5 / 22 / 4	9 (4+5)
17	53	16.3	pT3a pN0 cM0	8.4 / 9.1 / 1.5	8 (4+4)
18	72	28.9	pT3b pN1 cM0	25.8 / 21 / 0.1	8 (4+4)
19	70	16	pT3a pN0 cM0	3 / 0.2 / 0.7	7b (4+3)
20	67	218	pT3b pN0 cM0	87.7 / 121.2 / 26.6	8 (4+4)
Median	64.5	16.9			

PSMA PET imaging detected all PCa lesions in histopathology reference material with a diameter above 5 mm except one lesion with GS 7a in patient 11. Another lesion with GS 6 (3+3) and a diameter less than 5 mm had no corresponding PSMA uptake in PET in patient 10. The median volume of the prostatic gland among all patients in the prospective cohort was 42.1 (22.4-129.4) ml. GTV-Histo, GTV-Exp, and GTV-40% had a median volume of 7.3 (range: 0.2-121.2) ml, 6.3 (range: 0-87.7) ml, and 2.7 (range: 0-26.6) ml, respectively. PSMA PET detected positive lymph nodes in four patients (patient number: 3, 4, 9 and 18). In three patients (patient number: 13, 15 and 16) lymphadenectomy detected positive lymph nodes without corresponding signal in PET images. Please note that the pT (tumor)-stage, the pN (nodal)-stage and the GS were defined in the prostate specimen after prostatectomy. The PSA value represents the last PSA serum value before PET examination.

Supplementary Table 2. Patients' characteristics, internal validation cohort

	Age (y)	PSA (ng/ml)	TNM	Volume GTV-Exp_val (ml)	Gleason score
1	54	13.4	pT3b pN0 cM0	3.1	7b (4+3)
2	69	5.1	pT2b pN0 cM0	14.4	9 (5+4)
3	69	19.9	pT3a pN0 cM0	6.2	7a (3+4)
4	70	7.5	pT2b pN0 cM0	10.2	7b (4+3)
5	71	13.2	pT2c pN0 cM0	9.8	7a (3+4)
6	65	4.4	pT2c pN0 cM0	1.5	7a (3+4)
7	70	9.3	pT2c pN0 cM0	5.3	7b (4+3)
8	67	41.4	pT3b pN1 cM0	16.5	7b (4+3)
9	64	6.7	pT3a pN1 cM0	5.4	9 (5+4)
10	69	44.8	pT3b pN1 cM1	30.0	9 (4+5)
11	68	9.2	pT3a pN0 cM0	1.5	7b (4+3)
12	66	41	pT3b pN1 cM0	18.5	7a (3+4)
13	57	14.8	pT3b pN1 cM0	4.9	7b (4+3)
14	56	18	pT3b pN1 cM1	38.0	9 (5+4)
15	75	53.26	pT3a pN0 cM0	23.1	8 (4+4)
16	65	12.6	pT3a pN0 cM0	2.6	7b (4+3)
17	52	12.9	pT2a pN0 cM0	0.5	7a (3+4)
18	59	58.4	pT3a pN1 cM0	28.6	8 (4+4)
19	64	139	pT3b pN1 cM0	66.4	9 (4+5)
20	71	6	pT2c pN0 cM0	4.5	7a (3+4)
21	68	15.4	pT3b pN1 cM1	12.9	9 (5+4)
22	67	64	pT3b pN1 cM1	47.6	9 (5+4)
23	58	19.4	pT3b pN1 cM1	36.8	9 (5+4)
24	64	25.2	pT2c pN1 cM1	10.4	7a (3+4)
25	64	6.7	pT3b pN1 cM0	9.1	9 (5+4)
26	64	5	pT3a pN1 cM0	17.5	9 (5+4)
27	79	23.8	pT2c pN0 cM0	14.4	7a (4+3)
28	66	20	pT3a pN0 cM0	4.9	7a (3+4)
29	72	10.1	pT2c pN0 cM0	4.9	7b (4+3)
30	77	15	pT3b pN1 cM0	22.3	9 (4+5)
31	73	18.7	pT3a pN0 cM0	5.4	7a (4+3)
32	77	116	pT3a pN0 cM0	96.6	8 (4+4)
33	59	23.3	pT3b pN1 cM0	4.5	8 (4+4)
34	78	5.9	pT3a pN0 cM0	7.6	7b (4+3)
35	69	23	pT3a pN0 cM0	16.1	9 (5+4)
36	68	7.1	pT3b pN1 cM0	5.9	8 (4+4)
37	67	13.5	pT2c pN0 cM0	3.6	7b (4+3)
38	66	9.8	pT2c pN0 cM0	6.7	9 (5+4)
39	53	8.7	pT3b pN1 cM1	12.2	9 (5+4)
40	62	67	pT3b pN1 cM0	27.6	8 (4+4)
Median	67	15		10.17	

PSMA PET detected PCa in all patients. The median volume of GTV-Exp_val was 10.17 (range: 0.5-96.6) ml. PSMA PET detected positive lymph nodes in 13 patients. In 6 patients lymphadenectomy detected positive lymph nodes without corresponding signal in PET images.

Please note that the pT (tumor)-stage, the pN (nodal)-stage and the GS were defined in the prostate specimen after prostatectomy. The PSA value represents the last PSA serum value before PET examination.

Supplementary Table 3. Relations between TF from GTV-Exp and GS.

	ROC					Spearman correlation		Mann-Whitney U test
TF	AUC	Standard error	p - value	95% CI		ρ	p - value	p - value
QLRHGE	0.912	0.066	0.001	0.785	1.000	0.71	<0.001	<0.01
QHGRE	0.906	0.061	0.001	0.779	1.000	0.7	<0.001	<0.01
QSZHGE	0.912	0.058	0.001	0.801	1.000	0.71	<0.001	<0.01
QSRHGE	0.906	0.067	0.001	0.774	1.000	0.7	<0.001	<0.01

Abbreviations: TF: texture feature, QLRHGE: quantization algorithm + long-run high gray-level emphasis, QHGRE: quantization algorithm + high gray-level emphasis, QSZHGE: quantization algorithm + short zones high gray-level emphasis, QSRHGE: quantization algorithm + short-run high gray-level emphasis, ROC: receiver operating characteristics, AUC: area under the curve, CI: confidence interval.