

Supplementary I:

Table S.1. The information about the CT acquisition parameters. The slice thickness varied from 0.6mm to 5mm. The dataset also had images acquired from multiple reconstruction kernels. The pixel sizes ranged from 0.44 x 0.44 mm to 0.99 x 0.99mm with an average size of 0.75 x 0.75mm. Each slice had a XY planar resolution of 512 × 512 pixels with a 16-bit gray scale resolution in Hounsfield Units (HU). All images for 90 patients were reconstructed by standard (low pass) convolution kernels (Siemens: 'B321f', 'B35f'; Philips: 'B', 'C', 'D'; Toshiba: 'FC01', 'FC08', 'FC18'; GE: 'SOFT', 'STANDARD') while for 70 cases of 90, CT images were available for sharp (high pass) convolution kernels (Siemens: 'B321s', 'B35fs; GE: 'LUNG').

| Parameters | Acquisition parameters |
|-------------------------|-------------------------------|
| CT vendors | Siemens, Philips, Toshiba, GE |
| Convolution Kernels | STANDARD, SHARP |
| Peak tube voltage (kVp) | Median: 120 (110 – 140) |
| X-ray Tube current (mA) | Mean: 238 (36 – 623) |
| Exposure | Mean: 115 (18 – 327) |
| Slice Thickness | 0.6 – 5 mm |

Table S.2. Patient and treatment characteristics. Median (range) or mean is reported for continuous and counts (percentage) for categorical variables. Statistical difference between major pathological responders vs. non-responders was computed using a Fisher exact test or Wilcoxon-test respectively for categorical and continuous variables. Table S.2.1 shows demographics and clinical characteristics for 90 patients, categorized by responders and non-responders.

Demographics and clinical characteristics for 90 patients, categorized by responders and non-responders

| Characteristics | | All patients (n = 90) | Responders (n = 36) | Non-responders (n = 54) | p-value |
|-------------------------|-------------------------|----------------------------------|--------------------------------|------------------------------------|----------------|
| Sex | Male | 49 (54.4%) | 22 (61.1%) | 27 (50%) | 0.38 |
| | Female | 41 (45.5%) | 14 (38.9%) | 27 (50%) | |
| Age | Median (range) | 64 (38-88) | 61 (43-88) | 65.5 (38-79) | 0.46 |
| Clinical Staging | Stage III A | 85 (94.4%) | 33 (91.7%) | 52 (96.3%) | 0.38 |
| | Stage III B | 5 (5.6%) | 3 (8.3%) | 2 (3.7%) | |
| Histology | Adenocarcinoma | 64 (71.1%) | 18 (50%) | 46 (85%) | 0.0003 |
| | Squamous cell carcinoma | 20 (22.2%) | 12 (33%) | 8 (15%) | |
| | Others | 6 (6.7%) | 6 (17%) | 0 | |
| Procedure Type | Lobectomy | 70 (77.8%) | 24 (66.7%) | 46 (85.2%) | 0.68 |
| | Pneumonectomy | 20 (22.2%) | 12 (33.3%) | 8 (14.8%) | |
| ECOG PS | 0 | 17 (19%) | 8 (22.2%) | 9 (16.6%) | 0.36 |
| | 1 | 66 (73%) | 27 (75%) | 39 (72.2%) | |
| | Unknown | 7 (8%) | 1 (2.7%) | 6 (11.1%) | |

| | | | | | |
|--|--------------------------|------------------|-----------------|-----------------|-------|
| Treatment regimen (chemotherapy) | Carboplatin + Paclitaxel | 59 (65.6%) | 23 (63.9%) | 36 (66.9%) | 0.25 |
| | Carboplatin + Docetaxel | 2 (2.2%) | 0 | 2 (3.7%) | |
| | Cisplatin + Etoposide | 24 (26.7%) | 12 (33.3%) | 12 (22%) | |
| | Carboplatin + Pemetrexed | 2 (2.2%) | 0 | 2 (3.7%) | |
| | Unknown | 3 (3.3%) | 1 (2.8%) | 2 (3.7%) | |
| Radiation dose (Gy) | Mean (range) | 38.3 (30-60) | 39 (30-60) | 38.2 (30-60) | |
| Nodal disease | N2 | 82 (91.1%) | 31 (86%) | 51 (94.4%) | 0.17 |
| | N1 | 2 (2.2%) | 2 (5.7%) | 0 | |
| | N0 | 6 (6.7%) | 3 (8.3%) | 3 (5.6%) | |
| Follow-up [months] | Median (range) | 34.5 (0.13–114) | 41.2 (0.13-114) | 33.2 (0.53-105) | |
| Time to recurrence or distant metastasis [months] | Median (range) | 17.95 (0.2 – 70) | 12.6 (4.2-40.6) | 19.9 (0.2-70) | |
| Recurrence or Distant metastasis [No/Yes] | Recurrence | 38 (42.2%) | 8 (22.2%) | 30 (55.6%) | 0.002 |
| | Non recurrence | 52 (57.8%) | 28 (77.8%) | 24 (44.4%) | |

Table S.3. Demographics and clinical characteristics for 90 patients, categorized by training and test sets.

| Characteristics | | All patients (n = 90) | Training (n = 45) | Testing (n = 45) | p-value |
|---|--------------------------|----------------------------------|------------------------------|-----------------------------|----------------|
| Sex | Male | 49 (54.4%) | 27 | 22 | 0.39 |
| | Female | 41 (45.5%) | 18 | 23 | |
| Age | Median (range) | 64 (38-88) | 65 (38-88) | 64 (43-82) | |
| Clinical Staging | Stage III A | 85 (94.4%) | 41 | 44 | 0.36 |
| | Stage III B | 5 (5.6%) | 4 | 1 | |
| Histology | Adenocarcinoma | 64 (71.1%) | 29 | 35 | 0.44 |
| | Squamous cell carcinoma | 20 (22.2%) | 11 | 8 | |
| | Others | 6 (6.7%) | 4 | 2 | |
| Procedure Type | Lobectomy | 70 (77.8%) | 36 | 34 | 0.8 |
| | Pneumonectomy | 20 (22.2%) | 9 | 11 | |
| ECOG PS | 0 | 17 (19%) | 9 | 8 | 0.11 |
| | 1 | 66 (73%) | 30 | 36 | |
| | Unknown | 7 (8%) | 6 | 1 | |
| Treatment regimen (chemotherapy) | Carboplatin + Paclitaxel | 59 (65.6%) | 30 | 29 | 0.47 |
| | Carboplatin + Docetaxel | 2 (2.2%) | 1 | 1 | |
| | Cisplatin + Etoposide | 24 (26.7%) | 11 | 13 | |
| | Carboplatin + Pemetrexed | 2 (2.2%) | 2 | 0 | |
| | Unknown | 3 (3.3%) | 1 | 2 | |
| Radiation dose (Gy) | Mean (range) | 38.3 (30-60) | 37.1 (30-60) | 40.1 (30-60) | |

| | | | | | |
|--|----------------|------------------|-------------------|----------------|------|
| Nodal disease | N2 | 82 (91.1%) | 42 | 40 | 0.83 |
| | N1 | 2 (2.2%) | 1 | 1 | |
| | N0 | 6 (6.7%) | 2 | 4 | |
| Follow-up [months] | Median (range) | 34.5 (0.13–114) | 35.6 (1.73-105.2) | 33 (0.13-114) | |
| Time to recurrence or distant metastasis [months] | Median (range) | 17.95 (0.2 – 70) | 17.95 (1.5-60) | 17.85 (0.2-70) | |
| Recurrence or Distant metastasis [No/Yes] | Recurrence | 38 (42.2%) | 22 | 16 | 0.28 |
| | Non recurrence | 52 (57.8%) | 23 | 29 | |

Table S.4. Classifier performance based on CT slice thicknesses in terms of AUC mapped out as a function of slice thickness on the test set. With CT scans having a wide range of slice thickness (0.6 mm to 5 mm) the impact of slice thickness on the performance of the classifier was also evaluated. To do this, the classifier was trained using the features extracted from cases with a specific slice thickness in the training set and then evaluated on cases with the same slice thickness on the validation set. As may be observed, the AUC values for the radiomic features drop slightly with increasing slice thickness.

| Slice thickness criteria | Number of studies | AUC |
|--|--------------------------|------------|
| $slc \leq 1.5$ mm | 46 | 0.89 |
| $1.5 < slc \leq 5$ mm | 44 | 0.82 |

Table S.5. Hazard ratios (HR) from univariate Cox Proportional Hazard model on OS for combination of texture and clinicopathologic features.

| Covariate (Feature) | Hazard Ratio (HR) | 95% CI | | p-value |
|----------------------------|--------------------------|---------------|------|----------------|
| <i>Law_laplacian</i> | 2.74 | 1.16 | 6.48 | 0.021 |
| <i>Law</i> | 0.44 | 0.19 | 0.99 | 0.049 |
| <i>Law</i> | 1.51 | 0.91 | 2.48 | 0.11 |
| <i>Gabor</i> | 0.72 | 0.35 | 1.48 | 0.38 |
| <i>Tumor area</i> | 1.44 | 0.93 | 2.24 | 0.098 |
| <i>Law_Laplacian</i> | 0.41 | 0.22 | 0.92 | 0.047 |
| <i>Gabor</i> | 0.83 | 0.43 | 1.60 | 0.58 |
| <i>Law_Laplacian</i> | 1.53 | 0.1.11 | 2.36 | 0.044 |
| <i>Haralick</i> | 0.64 | 0.30 | 1.36 | 0.24 |
| <i>Law_Laplacian</i> | 1.6 | 0.84 | 3.00 | 0.15 |
| <i>Law_Laplacian</i> | 1.15 | 0.64 | 2.07 | 0.62 |
| <i>Gabor</i> | 0.78 | 0.39 | 1.57 | 0.49 |
| <i>Gabor</i> | 0.93 | 0.51 | 1.67 | 0.81 |
| <i>Histology</i> | 0.46 | 0.18 | 1.17 | 0.10 |
| <i>Vascular Invasion</i> | 1.25 | 0.77 | 2.02 | 0.35 |
| <i>Lymphatic Invasion</i> | 1.48 | 0.87 | 2.53 | 0.14 |
| <i>Tumor volume</i> | 0.83 | 0.44 | 1.58 | 0.58 |
| <i>age</i> | 1.00 | 0.95 | 1.06 | 0.84 |
| <i>sex</i> | 1.31 | 0.42 | 4.06 | 0.64 |

Table S.6. Multivariate Cox regression analysis for features that contributed to the risk-score for predicting OS.

| Covariate (Feature) | Hazard Ratio (HR) | 95% CI | p-value |
|----------------------------|--------------------------|---------------|----------------|
| <i>Law</i> | 0.11 | 0.02 0.46 | 0.0024 |
| <i>Gabor</i> | 0.04 | 0.005 0.36 | 0.0038 |
| <i>Tumor area</i> | 2.80 | 1.03 7.61 | 0.042 |
| <i>Law_Laplacian</i> | 11.6 | 1.95 68.8 | 0.007 |

Table S.7. Locoregional and recurrence location for all recurrence patients.

| No. of cases | Locoregional recurrence location | Systemic recurrence location |
|---------------------|---|-------------------------------------|
| 1 | R Supraclavicular Node | Bone |
| 1 | RM lobe | Bone and brain |
| 1 | - | Cervical Spine |
| 1 | Bilateral lung | Brain |
| 10 | - | Brain |
| 1 | - | Abdomen (malignant ascites) |
| 1 | Mediastinal + Hilar Nodes | Brain |
| 1 | Pleura + malignant effusion | - |
| 1 | other lung | Bone |
| 1 | - | Bone |
| 1 | Diffuse + axillary nodes | - |
| 1 | R hilar and paratracheal nodes | - |
| 1 | Diffuse | Brain |
| 1 | Supraclavicular node + Other lung | Brain |
| 1 | new primary vs mets in LUL and LLL | - |
| 1 | Bilateral lung | Brain |
| 1 | - | Adrenal |
| 1 | - | Brain and abdomen |
| 1 | LLL (vs LUL originally) | - |
| 1 | LUL (vs LLL originally) | - |
| 1 | - | Brain + Liver |
| 1 | Supraclavicular node | - |
| 1 | - | Cervical Node |
| 1 | Diffuse | Adrenal |
| 1 | LL lobe (RU originally) | Bone, Brain |
| 1 | Axillary nodes and chest wall | - |
| 1 | - | Liver |
| 1 | R chest wall / Pleura | - |

Table S.8. Multivariate Cox regression analysis for features that contributed to the risk-score for predicting DFS.

| Covariate (Feature) | Hazard Ratio (HR) | 95% CI | | p-value |
|----------------------------|--------------------------|---------------|-------|----------------|
| <i>Law_laplacian</i> | 6.00 | 1.9 | 19.04 | 0.0023 |
| <i>Law</i> | 1.32 | 0.41 | 4.23 | 0.64 |
| <i>Law_Laplacian</i> | 4.78 | 1.41 | 16.1 | 0.011 |
| <i>Haralick</i> | 0.6 | 0.27 | 1.35 | 0.09 |
| <i>Law_Laplacian</i> | 1.4 | 0.33 | 5.9 | 0.65 |
| <i>Law_Laplacian</i> | 0.15 | 0.04 | 0.56 | 0.0053 |
| <i>Gabor</i> | 8.91 | 2.51 | 31.6 | 0.00071 |