**Supplementary Material:**

**Supplementary Table S1: List of Radiomic Features**

|  |  |
| --- | --- |
| **Feature Category** | **Features** |
| First Order Statistics | Energy  Total Energy  Entropy  Minimum  10th Percentile  90th Percentile  Maximum  Mean  Median  Interquartile Range  Range  Mean Absolute Deviation  Robust Mean Absolute Deviation  Root Mean Squared  Standard Deviation  Skewness  Kurtosis  Variance  Uniformity |
| 3D Shape | Mesh Volume  Voxel Volume  Surface Area  Surface Area to Volume Ratio  Sphericity  Compactness  Spherical Disproportion  Maximum 3D Diameter  Maximum 2D Diameter (Axial)  Maximum 2D Diameter (Coronal)  Maximum 2D Diameter (Sagittal)  Major Axis Length  Minor Axis Length  Least Axis Length  Elongation  Flatness |
| Gray level Co-occurrence Matrix | Autocorrelation  Joint Average  Cluster Prominence  Cluster Shade  Cluster Tendency  Contrast  Correlation  Difference Average  Difference Entropy  Difference Variance  Difference Average  Joint Energy  Joint Entropy  Informational Correlation  Inverse Difference Moment  Inverse Difference Moment Normalized  Inverse Difference  Inverse Difference Normalized  Inverse Variance  Maximum Probability  Sum Average  Sum Entropy  Sum of Squares |
| Gray Level Size Zone Matrix | Small Area Emphasis  Large Area Emphasis  Gray Level Non-Uniformity  Gray Level Non-Uniformity Normalized  Size-Zone Non-Uniformity  Size-Zone Non-Uniformity Normalized  Zone Percentage  Gray Level Variance  Zone Variance  Zone Entropy  Low Gray Level Zone Emphasis  High Gray Level Zone Emphasis  Small Area Low Gray Level Emphasis  Small Area High Gray Level Emphasis  Large Area Low Gray Level Emphasis  Large Area High Gray Level Emphasis |
| Gray Level Run Length Matrix | Short Run Emphasis  Long Run Emphasis  Gray Level Non-Uniformity  Gray Level Non-Uniformity Normalized  Run Length Non-Uniformity  Run Length Non-Uniformity Normalized  Run Percentage  Gray Level Variance  Run Variance  Run Entropy  Low Gray Level Run Emphasis  High Gray Level Run Emphasis  Short Run Low Gray Level Emphasis  Short Run High Gray Level Emphasis  Long Run Low Gray Level Emphasis  Long Run High Gray Level Emphasis |
| Gray Level Dependence Matrix | Small Dependence Emphasis  Large Dependence Emphasis  Gray Level Non-Uniformity  Dependence Non-Uniformity  Dependence Non-Uniformity Normalized  Gray Level Variance  Dependence Variance  Dependence Entropy  Low Gray Level Emphasis  High Gray Level Emphasis  Small Dependence Low Gray Level Emphasis  Small Dependence High Gray Level Emphasis  Large Dependence Low Gray Level Emphasis  Large Dependence High Gray Level Emphasis |
| Neighboring Gray Tone Difference Matrix | Coarseness  Contrast  Busyness  Complexity  Strength |

Supplementary Table S1: A list of radiomic features extracted from a liver volume. The features include computations related to the statistics, shape, and gray-level relationships of the image.

**Supplementary Table S2: Radiomic Deviations from IBSI Standards**

|  |  |  |
| --- | --- | --- |
| **Computation** | **PyRadiomics Implementation** | **IBSI Guidelines** |
| Binning | Discretizes gray values with fixed bins with edges equally spaced from 0. | Discretizes using fixed bin width equally spaced from minimum of resegmentation range |
| Resampling | Aligns to the corner of the original voxel | Aligns to the center of the image |
| Gray value rounding | Does not implement resampling of similar resolution to original intensity values, with the argument that differences are likely to be minor and may add complexity. | Resamples to similar resolution of original CT image i.e. rounding to integer resolution of Hounsfield Units from the original CT intensity data. |
| Mask resampling | Resamples to nearest neighbor | Allows selection of different interpolators for resampling |

Supplementary Table S2: A list of deviations from the feature extraction guidelines by the Image Biomarker Standardisation Initiative (IBSI).

**Supplementary Equation S3: Random Survival Forest Algorithm**

To build a survival tree that predicts survival from an input vector of radiomic features, the following steps are taken:

1. Select *N* samples from the dataset.
2. For each sample *i = 1, 2, … N,* initialize a binary decision tree with max depth *D.*
3. At each node, iterate through set of features *X = {x­1­, x2, … xN}* and its range of feature values *S = {Smin, Smax}* to select feature *xi* and a threshold split value *si* such that:

Where *L(x,s)* is the log rank test such that

Where at time *ti*, *Ei*is the number of events at time *ti Ei,j* is the number of events at a daughter node *j, Yi* is the number of patients with an events or at risk at time *ti,* and *Yi,j*is the number of patients with an event or at risk at a daughter node *j*

1. Continue to grow children nodes unless the children node has no more than *M* surviving samples, where *M* is a user-defined hyperparameter
2. Calculate the cumulative hazard function for the decision tree with the Nelson-Aalen estimator:

Where *p* is a patient in the set of *M* patients in set *P = {p1,p2, … pM}*, *q* is a node in the set of *N* nodes in set *Q = {q1,q2, … qN}*, *Ep,q* is the number of events at time *tp,q*, and *Yp,q* is the number of patients with an event or at risk at time *tp,q*.

1. Repeat steps 1-5 *K* times to create *K* separately initialized trees, where *K* is a user-defined hyperparameter.
2. Average the cumulative hazard function over all trees to compute the ensembled cumulative hazard.

**Supplementary Table S4: Baseline variable distributions by lesion**

|  |  |
| --- | --- |
| **Characteristics** | **All lesions (n=129)** |
| Number of patients/lesions | 97/129 |
| Sex, n (%) |  |
| Male | 83 (64.3) |
| Female | 46 (36.7) |
| Metastasis at time of diagnosis, n (%) |  |
| M0 | 40 (30.0) |
| M1 | 89 (70.0) |
| Colorectal Histology, n (%) |  |
| Colon | 104 (80.6) |
| Rectal | 20 (15.6) |
| Undetermined | 5 (3.8) |
| Number of liver lesions at diagnosis, n (%) |  |
| 0 | 5 (3.9) |
| 1 | 25 (19.4) |
| 2 | 9 (7.0) |
| 3-5 | 32 (24.8) |
| > 5 | 53 (41.0) |
| Undetermined | 5 (3.9) |
| Other sites at diagnosis, n (%) |  |
| None | 101 (78.3) |
| Lung | 12 (9.3) |
| Non-regional LN | 3 (2.3) |
| Lung and non-regional LN | 4 (3.1) |
| Other | 7 (5.4) |
| Undetermined | 2 (1.6) |
| RT to other sites, n (%) |  |
| No | 75 (58.1) |
| Before liver RT | 28 (21.7) |
| After Liver RT | 21 (16.3) |
| Before and after liver RT | 2 (1.6) |
| Undetermined | 3 (2.3) |
| RT fraction delivered, Median (IQR) | 6 (5-15) |
| RT dose delivered, Median (IQR) | 4500 (3000 - 6000) |
| Dose Painting - Yes, n (%) | 55 (42.6) |
| Intended Dose Median (IQR) | 6000 (4000 - 6750) |
| Mean RT length ± SD (Days) | 11.6 (8.5) |
| PTV volume (cm3), Median (IQR) | 94.4 (39.2 - 174.4) |
| Mean D95 ± SD (% of intended dose) | 97.7 (11.0) |
| Reirradiation - Yes, n (%) | 8 (6.2) |
| Surgery before RT, n (%) | 91 (70.5) |
| Systemic before RT, n (%) | 126 (97.7) |
| Pump before RT, n (%) | 81 (62.8) |
| Lines of Chemo, Median (IQR) | 3 (2 - 4) |
| RFA before RT, n (%) | 45 (34.9) |
| RFA to RT lesions - Yes, n (%) | 13 (10.1) |
| Y90 before RT - Yes, n (%) | 10 (7.8) |
| Embolization before RT, n (%) | 12 (9.3) |
| CEA at diagnosis, Median (IQR) | 15.7 (3.38 - 176.9) |
| CEA at RT, Median (IQR) | 18.7 (4.8 - 127.2) |
| Number of liver lesions at RT, n (%) |  |
| 1 | 57 (44.2) |
| 2 | 43 (33.3) |
| 3 | 12 (9.4) |
| ≥ 4 | 16 (12.4) |
| Undetermined | 1 (0.7) |
| Other sites at RT, n (%) |  |
| None | 52 (40.3) |
| Lung | 27 (21.0) |
| Non-regional LN | 10 (7.8) |
| Lung and non-regional LN | 25 (19.3) |
| Other | 15 (11.6) |
| Mean lesion 1 dimension 2 ± SD | 35.2 (22.3) |
| Mean lesion 1 dimension 1 ± SD | 26.0 (17.9) |
| Freedom from local progression (FFLP), n (%) |  |
| Progression | 55 (42.6) |
| No progression | 67 (52.0) |
| Undetermined | 7 (5.4) |
| Mean FFLP (months) ± SD | 10.5 (0.4) |
| Any hepatic progression (AHP), n (%) |  |
| Progression | 99 (76.8) |
| No progression | 25 (19.4) |
| Undetermined | 5 (3.8) |
| Mean time to AHP (months) ± SD | 7.3 (7.1) |

Abbreviations: LN = lymph node, RT = radiotherapy, PTV = planning target volume, CEA = carcinoembryonic antigen, HAIP = hepatic arterial infusion pump, TARE = transarterial radioembolization.

Supplementary Table S4: A table of baseline clinical variables recorded as part of standard of care, with averages computed from the set of variables per lesion. The clinical variables will be utilized alongside computational radiomic features from computed tomography scans as input data to a machine learning model to predict local progression.

**Supplementary Table S5: Baseline variable distributions by patient**

|  |  |
| --- | --- |
| **Characteristics** | **All Patients (n=97)** |
| Sex, n (%) |  |
| Male | 63 (64.9) |
| Female | 34 (35.1) |
| Metastasis at time of diagnosis, n (%) |  |
| M0 | 32 (33) |
| M1 | 65 (67) |
| Other sites at diagnosis, n (%) |  |
| None | 74 (76.3) |
| Lung | 9 (9.3) |
| Non-regional LN | 3 (3.1) |
| Lung and non-regional LN | 4 (4.1) |
| Other | 5 (5.2) |
| Undetermined | 2 (2.0) |
| RT to other sites, n (%) |  |
| No | 58 (59.8) |
| Before liver RT | 23 (23.7) |
| After Liver RT | 13 (13.4) |
| Before and after liver RT | 2 (2.1) |
| Undetermined | 1 (1) |
| Number of liver lesions at RT, n (%) |  |
| 1 | 56 (57.7) |
| 2 | 25 (25.8) |
| 3 | 6 (6.2) |
| ≥ 4 | 9 (9.3) |
| Undetermined | 1 (1) |
| Other sites at RT, n (%) |  |
| None | 37 (38.1) |
| Lung | 21 (21.7) |
| Non-regional LN | 8 (8.3) |
| Lung and non-regional LN | 17 (17.5) |
| Other | 14 (14.4) |
| Freedom from local progression (FFLP), n (%) |  |
| Progression | 50 (51.6) |
| No progression | 40 (41.2) |
| Undetermined | 7 (7.2) |
| Mean time to local progression (months) ± SD | 10.5 (8.8) |
| Any hepatic progression (AHP), n (%) |  |
| Progression | 76 (78.4) |
| No progression | 16 (16.4) |
| Undetermined | 5 (5.2) |
| Mean time to AHP (months) ± SD | 7.4 (6.9) |

Abbreviations: LN = lymph node, RT = radiotherapy

Supplementary Table S5: A table of baseline clinical variables recorded as part of standard of care, with averages computed from the set of variables per patient. Lesion-specific variables were excluded.

**Supplementary Table S6: Dosages and number of fractions to liver metastases**

|  |  |  |  |
| --- | --- | --- | --- |
| **Total Dose (Gy)** | **Fractions** | **BED10 (Gy)** | **Patient Count** |
| 24 | 1 | 82 | 9 |
| 24 | 3 | 43 | 1 |
| 27 | 3 | 51 | 1 |
| 30 | 3 | 60 | 1 |
| 30 | 5 | 48 | 8 |
| 30 | 10 | 39 | 1 |
| 35 | 3 | 60 | 2 |
| 36 | 6 | 68 | 1 |
| 38 | 15 | 48 | 1 |
| 40 | 5 | 72 | 4 |
| 45 | 3 | 113 | 5 |
| 45 | 5 | 86 | 1 |
| 50 | 5 | 100 | 12 |
| 50 | 10 | 75 | 1 |
| 60 | 3 | 180 | 2 |
| 60 | 5 | 132 | 4 |
| 60 | 6 | 120 | 2 |
| 60 | 10 | 96 | 5 |
| 60 | 15 | 84 | 1 |
| 67.5 | 15 | 98 | 15 |
| 70 | 10 | 119 | 9 |
| 75 | 3 | 263 | 1 |
| 75 | 5 | 188 | 2 |
| 75 | 15 | 113 | 1 |
| 75 | 25 | 98 | 5 |
| 75 | 50 | 86 | 1 |
| 80 | 10 | 144 | 1 |

Supplementary Table S6: A list of doses to liver metastases, fractions, biologically effective dose (BED), and number of patients treated with the combination.

**Supplementary Material S7: Link to survival models**

The survival models from the different feature sets are uploaded to an open repository at: https://github.com/ricky-hu/local\_control\_radiomics\_survival\_model