

Supplementary Material

1 Supplementary Data

Supplementary Data S1. Definition of radiomic features

(Adapted from <http://pyradiomics.readthedocs.io/en/latest/features.html> – last accessed June 1, 2020)

a. First Order Features

These statistics describe the central tendency, variability, uniformity, asymmetry, skewness and magnitude of the attenuation values in a given region of interest (ROI), disregarding the spatial relationship of the individual voxels. As such, they describe quantitative and qualitative features of the whole ROI. A total of 19 features were calculated for the original MRI image (Table S1), as follows:

Let:

- X be a set of N_p voxels included in the region of interest (ROI)
- $P(i)$ be the first order histogram with N_g discrete intensity levels, where N_g is the number of non-zero bins, equally spaced from 0 with a width.
- $p(i)$ be the normalized first order histogram and equal to $\frac{P(i)}{N_p}$
- c is a value that shifts the intensities to prevent negative values in X . This ensures that voxels with the lowest gray values contribute the least to Energy, instead of voxels with gray level intensity closest to 0.
- ϵ is an arbitrarily small positive number ($\approx 2.2 \times 10^{-16}$)

Table S1. First-order features

| Radiomic feature | Interpretation |
|--|--|
| $\text{Energy} = \sum_{i=1}^{N_p} (X(i) + c)^2$ | Energy is a measure of the magnitude of voxel values in an image. A larger value implies a greater sum of the squares of these values. |
| $\text{Total Energy} = V_{\text{voxel}} \sum_{i=1}^{N_p} (X(i) + c)^2$ | Total Energy is the value of Energy feature scaled by the volume of the voxel in cubic mm. |
| $\text{Entropy} = - \sum_{i=1}^{N_g} p(i) \log_2 (p(i) + \epsilon)$ | Entropy specifies the uncertainty/randomness in the image values. It measures the average amount of information required to encode the image values |
| Minimum = $\min(X)$ | The minimum gray level intensity within the ROI. |
| The 10th percentile of X | The 10th percentile of X |
| The 90th percentile of X | The 90th percentile of X |
| Maximum = $\max(X)$ | The maximum gray level intensity within the ROI. |
| $\text{Mean} = \frac{1}{N_p} \sum_{i=1}^{N_p} X(i)$ | The average (mean) gray level intensity within the ROI. |
| Median | The median gray level intensity within the ROI. |
| Interquartile range = $P_{75} - P_{25}$ | Here P_{25} and P_{75} are the 25 th and 75 th percentile of the image array, respectively. |
| Range = $\max(X) - \min(X)$ | The range of gray values in the ROI. |

| | |
|--|---|
| $\text{MAD} = \frac{1}{N_p} \sum_{i=1}^{N_p} X(i) - \bar{X} $ | Mean Absolute Deviation (MAD) is the mean distance of all intensity values from the Mean Value of the image array. |
| $\text{rMAD} = \frac{1}{N_{10-90}} \sum_{i=1}^{N_{10-90}} X_{10-90}(i) - \bar{X}_{10-90} $ | Robust Mean Absolute Deviation (rMAD) is the mean distance of all intensity values from the Mean Value calculated on the subset of image array with gray levels in between, or equal to the 10 th and 90 th percentile. |
| $\text{RMS} = \sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) + c)^2}$ | Root Mean Squared (RMS) is the square-root of the mean of all the squared intensity values. It is another measure of the magnitude of the image values. This feature is volume-confounded, a larger value of c increases the effect of volume-confounding. |
| $\text{SD} = \sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^2}$ | Standard Deviation (SD) measures the amount of variation or dispersion from the Mean Value. By definition, standard deviation = $\sqrt{\text{variance}}$ |
| $\begin{aligned} \text{Skewness} &= \frac{\mu_3}{\sigma^3} \\ &= \frac{\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^3}{\left(\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^2 \right)^{3/2}} \end{aligned}$ | Skewness measures the <i>asymmetry</i> of the distribution of values about the Mean value. Depending on where the tail is elongated and the mass of the distribution is concentrated, this value can be positive or negative. (Where μ^3 is the 3 rd central moment). |
| $\begin{aligned} \text{Kurtosis} &= \frac{\mu_4}{\sigma^4} \\ &= \frac{\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^4}{\left(\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^2 \right)^2} \end{aligned}$ | Kurtosis is a measure of the ‘ <i>peakedness</i> ’ of the distribution of values in the image ROI. A higher kurtosis implies that the mass of the distribution is concentrated towards the tail(s) rather than towards the mean. A lower kurtosis implies the reverse: that the mass of the distribution is concentrated towards a spike near the Mean value. (Where μ_4 is the 4th central moment). |
| $\text{Variance} = \frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^2$ | Variance is the the mean of the squared distances of each intensity value from the Mean value. This is a measure of the spread of the distribution about the mean. |
| $\text{Uniformity} = \sum_{i=1}^{N_g} p(i)^2$ | Uniformity is a measure of the sum of the squares of each intensity value. This is a measure of the heterogeneity of the image array, <i>where a greater uniformity implies a greater heterogeneity or a greater range of discrete intensity values.</i> |

b. 2D and 3D Shape Features

Shape features describe the size and shape of a given ROI, without taking into account the attenuation values of its voxels. Since they are independent of the gray level intensities, shape features were consistent across all wavelet transformation and the original MRI image, and therefore were only calculated once. These were defined as follows (Table S2):

Let:

V the volume of the ROI in mm³

A the surface area of the ROI in mm²

Table S2. 2D and 3D shape features

| Radiomic feature | Interpretation |
|------------------|----------------|
|------------------|----------------|

| | |
|--|---|
| $\text{Volume} = \sum_{i=1}^N V_i$ | <p>The volume of the ROI V is approximated by multiplying the number of voxels in the ROI by the volume of a single voxel V_i.</p> |
| <p>Surface Area</p> $= \sum_{i=1}^N \frac{1}{2} a_i b_i \times a_i c_i $ | <p>Surface Area is an approximation of the surface of the ROI in mm^2, calculated using a marching cubes algorithm, where N is the number of triangles forming the surface mesh of the volume (ROI), $a_i b_i$ and $a_i c_i$ are the edges of the i^{th} triangle formed by points a_i, b_i and c_i.</p> |
| <p>Surface to volume ratio = $\frac{A}{V}$</p> | <p>Here, a lower value indicates a more compact (sphere-like) shape. This feature is not dimensionless, and is therefore (partly) dependent on the volume of the ROI.</p> |
| <p>Sphericity = $\frac{\sqrt[3]{36\pi V^2}}{A}$</p> | <p>Sphericity is a measure of the roundness of the shape of the tumor region relative to a sphere. It is a dimensionless measure, independent of scale and orientation. The value range is $0 < \text{sphericity} \leq 1$, where a value of 1 indicates a perfect sphere (a sphere has the smallest possible surface area for a given volume, compared to other solids).</p> |
| <p>Maximum 3D diameter</p> | <p>Maximum 3D diameter is defined as the largest pairwise Euclidean distance between surface voxels in the ROI (Feret Diameter).</p> |
| <p>Maximum 2D diameter (Slice)</p> | <p>Maximum 2D diameter (Slice) is defined as the largest pairwise Euclidean distance between ROI surface voxels in the row-column (generally the axial) plane.</p> |
| <p>Maximum 2D diameter (Column)</p> | <p>Maximum 2D diameter (Column) is defined as the largest pairwise Euclidean distance between ROI surface voxels in the row-slice (usually the coronal) plane.</p> |
| <p>Maximum 2D diameter (Row)</p> | <p>Maximum 2D diameter (Row) is defined as the largest pairwise Euclidean distance between tumor surface voxels in the column-slice (usually the sagittal) plane.</p> |
| <p>Mesh Volume</p> | <p>The Mesh Volume of the ROI is calculated from the triangle mesh of the ROI. For each face in the mesh, defined by points, the (signed) volume of the tetrahedron defined by that face and the origin of the image is calculated.</p> |
| <p>Major axis = $4\sqrt{\lambda_{\text{major}}}$</p> | |
| <p>Minor axis = $4\sqrt{\lambda_{\text{minor}}}$</p> | |
| <p>Least axis = $4\sqrt{\lambda_{\text{least}}}$</p> | |
| <p>Elongation = $\sqrt{\frac{\lambda_{\text{minor}}}{\lambda_{\text{major}}}}$</p> | <p>Here, λ_{major} and λ_{minor} are the lengths of the <u>largest</u> and <u>second largest</u> principal component axes. The values range between 1 (circle-like (non-elongated)) and 0 (single point or 1 dimensional line).</p> |
| <p>Flatness = $\sqrt{\frac{\lambda_{\text{least}}}{\lambda_{\text{major}}}}$</p> | <p>Here, λ_{major} and λ_{minor} are the lengths of the <u>largest</u> and <u>smallest</u> principal component axes. The values range between 1 (non-flat, sphere-like) and 0 (a flat object).</p> |

c. Gray Level Co-occurrence Matrix (GLCM)

In simple words, a GLCM describes the number of times a voxel of a given attenuation value i is located next to a voxel of j . A GLCM of size $N_g \times N_g$ describes the second-order joint probability function of an image region constrained by the mask and is defined as $P(i, j | \delta, \theta)$. The $(i, j)^{\text{th}}$ element of this

matrix represents the number of times the combination of levels i and j occur in two pixels in the image, that are separated by a distance of δ pixels along angle θ . The distance δ from the center voxel is defined as the distance according to the infinity norm. For $\delta=1$, this results in 26 neighbors for each of 13 angles in 3D (26-connectivity) and for $\delta=2$ a 98-connectivity (49 unique angles). In order to get rotationally invariant results, statistics are calculated in all directions and then averaged, to ensure a symmetrical GLCM (Table S3).

Let:

ϵ be an arbitrarily small positive number ($\approx 2.2 \times 10^{-16}$)

$P(i, j)$ be the co-occurrence matrix for an arbitrary δ and θ

$p(i, j)$ be the normalized co-occurrence matrix and equal to $\frac{P(i, j)}{\sum P(i, j)}$

N_g be the number of discrete intensity levels in the image

$p_x(i) = \sum_{j=1}^{N_g} P(i, j)$ be the marginal row probabilities

$p_y(j) = \sum_{i=1}^{N_g} P(i, j)$ be the marginal column probabilities

μ_x be the mean gray level intensity of p_x and defined as $\mu_x = \sum_{i=1}^{N_g} p_x(i) i$

μ_y be the mean gray level intensity of p_y and defined as $\mu_y = \sum_{j=1}^{N_g} p_y(j) j$

σ_x be the standard deviation of p_x

σ_y be the standard deviation of p_y

$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)$, where $i + j = k$, and $k = 2, 3, \dots, 2N_g$

$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)$, where $|i - j| = k$, and $k = 0, 1, \dots, N_g - 1$

$HX = - \sum_{i=1}^{N_g} p_x(i) \log_2 (p_x(i) + \epsilon)$ be the entropy of p_x

$HY = - \sum_{j=1}^{N_g} p_y(j) \log_2 (p_y(j) + \epsilon)$ be the entropy of p_y

$HXY1 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log_2 (p_x(i) p_y(j) + \epsilon)$

$HXY2 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_x(i) p_y(j) \log_2 (p_x(i) p_y(j) + \epsilon)$

For distance weighting, GLCM matrices are weighted by weighting factor W and then summed and normalized. Weighting factor W is calculated for the distance between neighboring voxels by $W = e^{-\|d\|^2}$, where d is the distance for the associated angle.

Table S3. Gray Level Co-occurrence Matrix (GLCM)

| Radiomic feature | Interpretation |
|---|---|
| $\text{Autocorrelation} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) ij$ | Autocorrelation is a measure of the magnitude of the fineness and coarseness of texture. |

$$\text{Joint average} = \mu_x = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) i$$

Returns the mean gray level intensity of the i distribution.

Cluster prominence

$$= \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^4 p(i,j)$$

Cluster Prominence is a measure of the skewness and asymmetry of the GLCM. A higher value implies more asymmetry around the mean while a lower value indicates a peak near the mean value and less variation around the mean.

Cluster tendency

$$= \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^2 p(i,j)$$

Cluster Tendency is a measure of groupings of voxels with similar gray-level values.

Cluster shade

$$= \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^3 p(i,j)$$

Cluster Shade is a measure of the skewness and uniformity of the GLCM. A higher cluster shade implies greater asymmetry about the mean.

$$\text{Contrast} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - j)^2 p(i,j)$$

Contrast is a measure of the local intensity variation, favoring values away from the diagonal ($i=j$). A larger value correlates with a greater disparity in intensity values among neighboring voxels.

$$\text{Correlation} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) ij - \mu_x \mu_y}{\sigma_x(i) \sigma_y(j)}$$

Correlation is a value between 0 (uncorrelated) and 1 (perfectly correlated) showing the linear dependency of gray level values to their respective voxels in the GLCM.

$$\text{Difference average} = \sum_{k=0}^{N_g-1} k p_{x-y}(k)$$

Difference Average measures the relationship between occurrences of pairs with similar intensity values and occurrences of pairs with differing intensity values.

Difference entropy

$$= \sum_{k=0}^{N_g-1} p_{x-y}(k) \log_2 (p_{x-y}(k) + \epsilon)$$

Difference Entropy is a measure of the randomness/variability in neighborhood intensity value differences.

$$\text{Difference variance} = \sum_{k=0}^{N_g-1} (k - DA)^2 p_{x-y}(k)$$

Difference Variance is a measure of heterogeneity that places higher weights on differing intensity level pairs that deviate more from the mean.

$$\text{Joint energy} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (p(i,j))^2$$

Joint energy is a measure of homogeneous patterns in the image. A greater **joint energy** implies that there are more instances of intensity value pairs in the image that neighbor each other at higher frequencies.

$$\text{Joint entropy} = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \log_2 (p(i,j) + \epsilon)$$

Joint entropy is a measure of the randomness/variability in neighborhood intensity values.

$$\text{IMC 1} = \frac{HXY - HXY1}{\max\{HX, HY\}}$$

Informational measure of correlation 1

$$\text{IMC 2} = \sqrt{1 - e^{-2(HXY2 - HXY)}}$$

Informational measure of correlation 2

$$\text{IDM} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i, j)}{1 + |i - j|^2}$$

IDM (inverse difference moment a.k.a Homogeneity 2) is a measure of the local homogeneity of an image. IDM weights are the inverse of the Contrast weights (decreasing exponentially from the diagonal $i=j$ in the GLCM).

$$\text{IDMN} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i, j)}{1 + \left(\frac{|i - j|^2}{N_g^2}\right)}$$

IDMN (inverse difference moment normalized) is a measure of the local homogeneity of an image. IDMN weights are the inverse of the Contrast weights (decreasing exponentially from the diagonal $i=j$ in the GLCM). Unlike Homogeneity2, IDMN normalizes the square of the difference between neighboring intensity values by dividing over the square of the total number of discrete intensity values.

$$\text{ID} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i, j)}{1 + |i - j|}$$

ID (inverse difference a.k.a. Homogeneity 1) is another measure of the local homogeneity of an image. With more uniform gray levels, the denominator will remain low, resulting in a higher overall value.

$$\text{IDN} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i, j)}{1 + \left(\frac{|i - j|}{N_g}\right)}$$

IDN (inverse difference normalized) is another measure of the local homogeneity of an image. Unlike Homogeneity1, IDN normalizes the difference between the neighboring intensity values by dividing over the total number of discrete intensity values.

$$\text{Inverse variance} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i, j)}{|i - j|^2}, i \neq j$$

$$\text{Maximum probability} = \max(p(i, j))$$

Maximum Probability is occurrences of the most predominant pair of neighboring intensity values.

$$\text{Sum average} = \sum_{k=2}^{2N_g} p_{x+y}(k)k$$

Sum Average measures the relationship between occurrences of pairs with lower intensity values and occurrences of pairs with higher intensity values.

$$\text{Sum entropy} = \sum_{k=2}^{2N_g} p_{x+y}(k) \log_2 (p_{x+y}(k) + \epsilon)$$

Sum Entropy is a sum of neighborhood intensity value differences.

$$\text{Sum squares} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu_x)^2 p(i, j)$$

Sum of Squares or Variance is a measure in the distribution of neighboring intensity level pairs about the mean intensity level in the GLCM.

d. Gray Level Run Length Matrix (GLRLM)

A Gray Level Run Length Matrix (GLRLM) describes gray level runs, which are defined as the length in number of pixels, of consecutive pixels that have the same gray level value. In a gray level run length matrix $P(i,j|\theta)$, the $(i,j)^{th}$ element describes the number of runs with gray level i and length j occur in the image (ROI) along angle θ (Table S4).

Let:

N_g be the number of discrete intensity values in the image

N_r be the number of discrete run lengths in the image

N_p be the number of voxels in the image

$N_z(\theta)$ be the number of runs in the image along angle θ , which is equal to $\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j|\theta)$ and $1 \leq N_z(\theta) \leq N_p$

$P(i,j|\theta)$ be the run length matrix for an arbitrary direction θ

$p(i,j|\theta)$ be the normalized run length matrix, defined as $p(i,j|\theta) = \frac{P(i,j|\theta)}{N_z(\theta)}$

ϵ is an arbitrarily small positive number ($\approx 2.2 \times 10^{-16}$).

By default, the value of a feature is calculated on the GLRLM for each angle separately, after which the mean of these values is returned. If distance weighting is enabled, GLRLMs are weighted by the distance between neighbouring voxels and then summed and normalised. Features are then calculated on the resultant matrix. The distance between neighbouring voxels is calculated for each angle using the norm specified in ‘weightingNorm’

Table S4. Gray Level Run Length Matrix (GLRLM)

| Radiomic feature | Interpretation |
|--|---|
| $\text{SRE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j \theta)}{j^2}}{N_z(\theta)}$ | SRE (Short Run Emphasis) is a measure of the distribution of short run lengths, with a greater value indicative of shorter run lengths and more fine textural textures. |
| $\text{LRE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j \theta) j^2}{N_z(\theta)}$ | LRE (Long Run Emphasis) is a measure of the distribution of long run lengths, with a greater value indicative of longer run lengths and more coarse structural textures. |
| $\text{GLN} = \frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_r} P(i,j \theta))^2}{N_z(\theta)}$ | GLN (Gray Level Non-uniformity) measures the similarity of gray-level intensity values in the image, where a lower GLN value correlates with a greater similarity in intensity values. |
| $\text{GLNN} = \frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_r} P(i,j \theta))^2}{N_z(\theta)^2}$ | GLNN (Gray Level Non-uniformity Normalized) measures the similarity of gray-level intensity values in the image, where a lower GLNN value correlates with a greater similarity in intensity values. This is the normalized version of the GLN formula. |
| $\text{RLN} = \frac{\sum_{j=1}^{N_r} (\sum_{i=1}^{N_g} P(i,j \theta))^2}{N_z(\theta)}$ | RLN (Run Length Non-uniformity) measures the similarity of run lengths throughout the image, with a lower value indicating more homogeneity among run lengths in the image. |

$$\text{RLNN} = \frac{\sum_{j=1}^{N_r} (\sum_{i=1}^{N_g} P(i, j|\theta))^2}{N_z(\theta)^2}$$

RLNN (Run Length Non-uniformity) measures the similarity of run lengths throughout the image, with a lower value indicating more homogeneity among run lengths in the image. This is the normalized version of the RLNN formula.

$$\text{RP} = \frac{N_z(\theta)}{N_p}$$

RP (Run Percentage) measures the coarseness of the texture by taking the ratio of number of runs and number of voxels in the ROI. Values are in range $\frac{1}{N_p} \leq RP \leq 1$, with higher values indicating a larger portion of the ROI consists of short runs (indicates a more fine texture).

$$\text{GLV} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta)(i - \mu)^2,$$

$$\text{where } \mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta)i$$

GLV (Gray Level Variance) measures the variance in gray level intensity for the runs.

$$\text{RV} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta)(j - \mu)^2,$$

$$\text{where } \mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta)j$$

RV (Run Variance) is a measure of the variance in runs for the run lengths.

$$\text{RE} = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta) \log_2 (p(i, j|\theta) + \epsilon)$$

RE (Run Entropy) measures the uncertainty/randomness in the distribution of run lengths and gray levels. A higher value indicates more heterogeneity in the texture patterns.

$$\text{LGLRE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i, j|\theta)}{i^2}}{N_z(\theta)}$$

LGLRE (low gray level run emphasis) measures the distribution of low gray-level values, with a higher value indicating a greater concentration of low gray-level values in the image.

$$\text{HGLRE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i, j|\theta)i^2}{N_z(\theta)}$$

HGLRE (high gray level run emphasis) measures the distribution of the higher gray-level values, with a higher value indicating a greater concentration of high gray-level values in the image.

$$\text{SRLGLE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i, j|\theta)}{i^2 j^2}}{N_z(\theta)}$$

SRLGLE (short run low gray level emphasis) measures the joint distribution of shorter run lengths with lower gray-level values.

$$\text{SRHGLE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i, j|\theta)i^2}{j^2}}{N_z(\theta)}$$

SRHGLE (short run high gray level emphasis) measures the joint distribution of shorter run lengths with higher gray-level values.

$$\text{LRLGLRE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i, j|\theta)j^2}{i^2}}{N_z(\theta)}$$

LRLGLRE (long run low gray level emphasis) measures the joint distribution of long run lengths with lower gray-level values.

$$\text{LRHGLRE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i, j|\theta)i^2 j^2}{N_z(\theta)}$$

LRHGLRE (long run high gray level run emphasis) measures the joint distribution of long run lengths with higher gray-level values.

e. Gray Level Size Zone Matrix (GLSZM)

A Gray Level Size Zone (GLSZM) describes gray level zones in a ROI, which are defined as the number of connected voxels that share the same gray level intensity. A voxel is considered connected if the distance is 1 according to the infinity norm (26-connected region in a 3D, 8-connected region in 2D). In a gray level size zone matrix $P(i,j)$ the $(i,j)^{th}$ element equals the number of zones with gray level i and size j appear in image. Contrary to GLCM and GLRLM, the GLSZM is rotation independent, with only one matrix calculated for all directions in the ROI (Table S5).

Let:

N_g be the number of discrete intensity values in the image

N_s be the number of discrete zone sizes in the image

N_p be the number of voxels in the image

N_z be the number of zones in the ROI, which is equal to $\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} P(i,j)$ and $1 \leq N_z \leq N_p$

$P(i,j)$ be the size zone matrix

$p(i,j)$ be the normalized size zone matrix, defined as $p(i,j) = \frac{P(i,j)}{N_z}$

ϵ is an arbitrarily small positive number ($\approx 2.2 \times 10^{-16}$).

Table S5. Gray Level Size Zone Matrix (GLSZM)

| Radiomic feature | Interpretation |
|--|---|
| $SAE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{P(i,j)}{j^2}}{N_z}$ | SAE (small area emphasis) is a measure of the distribution of small size zones, with a greater value indicative of smaller size zones and more fine textures. |
| $LAE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} P(i,j)j^2}{N_z}$ | LAE (large area emphasis) is a measure of the distribution of large area size zones, with a greater value indicative of larger size zones and more coarse textures. |
| $GLN = \frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_s} P(i,j))^2}{N_z}$ | GLN (gray level non-uniformity) measures the variability of gray-level intensity values in the image, with a lower value indicating more homogeneity in intensity values. |
| $GLNN = \frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_s} P(i,j))^2}{N_z^2}$ | GLNN (gray level non-uniformity normalized) measures the variability of gray-level intensity values in the image, with a lower value indicating a greater similarity in intensity values. This is the normalized version of the GLN formula. |
| $SZN = \frac{\sum_{j=1}^{N_s} (\sum_{i=1}^{N_g} P(i,j))^2}{N_z}$ | SZN (size zone non-uniformity) measures the variability of size zone volumes in the image, with a lower value indicating more homogeneity in size zone volumes. |
| $SZNN = \frac{\sum_{j=1}^{N_s} (\sum_{i=1}^{N_g} P(i,j))^2}{N_z^2}$ | SZNN (size zone non-uniformity normalized) measures the variability of size zone volumes throughout the image, with a lower value indicating more homogeneity among zone size volumes in the image. This is the normalized version of the SZN formula. |
| $\text{Zone Percentage} = \frac{N_z}{N_p}$ | ZP (Zone Percentage) measures the coarseness of the texture by taking the ratio of number of zones and number of voxels in the ROI. Values are in range $\frac{1}{N_p} \leq ZP \leq 1$, with higher values indicating a larger portion |

| | |
|---|--|
| | of the ROI consists of small zones (indicates a more fine texture). |
| $\text{GLV} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j)(i - \mu)^2,$ $\text{where } \mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j)i$ | Gray level variance (GLV) measures the variance in gray level intensities for the zones. |
| $\text{ZV} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j)(j - \mu)^2, \text{ where}$ $\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j)j$ | Zone Variance (ZV) measures the variance in zone size volumes for the zones. |
| $\text{ZE} = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j) \log_2 (p(i,j) + \epsilon)$ | Zone Entropy (ZE) measures the uncertainty/randomness in the distribution of zone sizes and gray levels. A higher value indicates more heterogeneity in the texture patterns. |
| $\text{LGLZE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{P(i,j)}{i^2}}{N_z}$ | LGLZE (low gray level zone emphasis) measures the distribution of lower gray-level size zones, with a higher value indicating a greater proportion of lower gray-level values and size zones in the image. |
| $\text{HGLZE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} P(i,j)i^2}{N_z}$ | HGLZE (high gray level zone emphasis) measures the distribution of the higher gray-level values, with a higher value indicating a greater proportion of higher gray-level values and size zones in the image. |
| $\text{SALGLE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{P(i,j)}{i^2 j^2}}{N_z}$ | SALGLE (small area low gray level emphasis) measures the proportion in the image of the joint distribution of smaller size zones with lower gray-level values. |
| $\text{SAHGLE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{P(i,j)i^2}{j^2}}{N_z}$ | SAHGLE (small area high gray level emphasis) measures the proportion in the image of the joint distribution of smaller size zones with higher gray-level values. |
| $\text{LALGLE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{P(i,j)j^2}{i^2}}{N_z}$ | LALGLE (low area low gray level emphasis) measures the proportion in the image of the joint distribution of larger size zones with lower gray-level values. |
| $\text{LAHGLE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} P(i,j)i^2 j^2}{N_z}$ | LAHGLE (low area high gray level emphasis) measures the proportion in the image of the joint distribution of larger size zones with higher gray-level values. |

f. Neighbouring Gray Tone Difference Matrix (NGTDM)

A Neighbouring Gray Tone Difference Matrix quantifies the difference between a gray value and the average gray value of its neighbours within distance δ (Table S6). The sum of absolute differences for gray level i is stored in the matrix. Let \mathbf{X}_{gl} be a set of segmented voxels and $x_{gl}(j_x, j_y, j_z) \in X_{gl}$ be the gray level of a voxel at position (j_x, j_y, j_z) , then the average gray level of the neighbourhood is:

$$\bar{A}_i = \bar{A}(j_x, j_y, j_z)$$

$$= \frac{1}{W} \sum_{k_x=-\delta}^{\delta} \sum_{k_y=-\delta}^{\delta} \sum_{k_z=-\delta}^{\delta} x_{gl}(j_x + k_x, j_y + k_y, j_z + k_z),$$

where $(k_x, k_y, k_z) \neq (0,0,0)$ and $x_{gl}(j_x + k_x, j_y + k_y, j_z + k_z) \in X_{gl}$

Here, W is the number of voxels in the neighbourhood that are also in X_{gl} .

Let:

n_i be the number of voxels in X_{gl} with gray level i

$N_{v,p}$ be the total number of voxels in X_{gl} and equal to $\sum n_i$ (i.e. the number of voxels with a valid region; at least 1 neighbor). $N_{v,p} \leq N_p$, where N_p is the total number of voxels in the ROI.

p_i be the gray level probability and equal to n_i/N_v

$$s_i = \begin{cases} \sum^{n_i} |i - \bar{A}_i| & \text{for } n_i \neq 0 \\ 0 & \text{for } n_i = 0 \end{cases}$$

be the sum of absolute differences for gray level i

N_g be the number of discrete gray levels

$N_{g,p}$ be the number of gray levels where $p_i \neq 0$

Table S6. Neighbouring Gray Tone Difference Matrix (NGTDM)

| Radiomic feature | Interpretation |
|---|---|
| $Coarseness = \frac{1}{\sum_{i=1}^{N_g} p_i s_i}$ | Coarseness is a measure of average difference between the center voxel and its neighbourhood and is an indication of the spatial rate of change. A higher value indicates a lower spatial change rate and a locally more uniform texture. |
| $Contrast = \left(\frac{1}{N_{g,p}(N_{g,p} - 1)} \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_i p_j (i - j)^2 \right) \left(\frac{1}{N_{v,p}} \sum_{i=1}^{N_g} s_i \right),$ | Contrast is a measure of the spatial intensity change, but is also dependent on the overall gray level dynamic range. Contrast is high when both the dynamic range and the spatial change rate are high, i.e. an image with a large range of gray levels, with large changes between voxels and their neighbourhood. |
| where $p_i \neq 0, p_j \neq 0$ | |
| $Busyness = \frac{\sum_{i=1}^{N_g} p_i s_i}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} ip_i - jp_j },$ | A measure of the change from a pixel to its neighbour. A high value for busyness indicates a ‘busy’ image, with rapid changes of intensity between pixels and its neighbourhood. |
| where $p_i \neq 0, p_j \neq 0$ | |

$$\text{Complexity} = \frac{1}{N_{v,p}} \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} |i-j| \frac{p_i s_i + p_j s_j}{p_i + p_j},$$

An image is considered complex when there are many primitive components in the image, i.e. the image is non-uniform and there are many rapid changes in gray level intensity.

where $p_i \neq 0, p_j \neq 0$

$$\text{Strength} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (p_i + p_j)(i-j)^2}{\sum_{i=1}^{N_g} s_i},$$

Strength is a measure of the primitives in an image. Its value is high when the primitives are easily defined and visible, i.e. an image with slow change in intensity but more large coarse differences in gray level intensities.

where $p_i \neq 0, p_j \neq 0$

g. Gray Level Dependence Matrix (GLDM)

A Gray Level Dependence Matrix (GLDM) quantifies gray level dependencies in an image. A gray level dependency is defined as the number of connected voxels within distance δ that are dependent on the center voxel. A neighbouring voxel with gray level j is considered dependent on center voxel with gray level i if $|i-j| \leq \delta$. In a gray level dependence matrix $P(i,j)$ the $(i,j)^{th}$ element describes the number of times a voxel with gray level i with j dependent voxels in its neighbourhood appears in image (Table S7).

N_g be the number of discrete intensity values in the image

N_d be the number of discrete dependency sizes in the image

N_z be the number of dependency zones in the image, which is equal to $\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} P(i,j)$

$P(i,j)$ be the dependence matrix

$p(i,j)$ be the normalized dependence matrix, defined as $p(i,j) = \frac{P(i,j)}{N_z}$

Table S7. Gray Level Dependence Matrix (GLDM)

| Radiomic feature | Interpretation |
|---|--|
| $SDE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{P(i,j)}{i^2}}{N_z}$ | SDE (Small Dependence Emphasis): A measure of the distribution of small dependencies, with a greater value indicative of smaller dependence and less homogeneous textures. |
| $LDE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} P(i,j)j^2}{N_z}$ | LDE (Large Dependence Emphasis): A measure of the distribution of large dependencies, with a greater value indicative of larger dependence and more homogeneous textures. |
| $GLN = \frac{\sum_{i=1}^{N_g} (\sum_{j=1}^{N_d} P(i,j))^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} P(i,j)}$ | GLN (Gray Level Non-Uniformity): Measures the similarity of gray-level intensity values in the image, where a lower GLN value correlates with a greater similarity in intensity values. |
| $DN = \frac{\sum_{j=1}^{N_d} (\sum_{i=1}^{N_g} P(i,j))^2}{N_z}$ | DN (Dependence Non-Uniformity): Measures the similarity of dependence throughout the image, with a lower value indicating more homogeneity among dependencies in the image. |

$$DNN = \frac{\sum_{j=1}^{N_d} (\sum_{i=1}^{N_g} P(i,j))^2}{N_z^2}$$

DNN (Dependence Non-Uniformity Normalized): Measures the similarity of dependence throughout the image, with a lower value indicating more homogeneity among dependencies in the image. This is the normalized version of the DLN formula.

$$GLV = \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} p(i,j)(i - \mu)^2,$$

GLV (Gray Level Variance): Measures the variance in grey level in the image.

$$\text{where } \mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} ip(i,j)$$

$$DV = \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} p(i,j)(j - \mu)^2,$$

DV (Dependence Variance): Measures the variance in dependence size in the image.

$$\text{where } \mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} jp(i,j)$$

$$DE = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} p(i,j) \log_2 (p(i,j) + \epsilon)$$

DE (Dependence Entropy): Measures the entropy in dependence size in the image.

$$LGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{P(i,j)}{i^2}}{N_z}$$

LGLE (Low Gray Level Emphasis): Measures the distribution of low gray-level values, with a higher value indicating a greater concentration of low gray-level values in the image.

$$HGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} P(i,j)i^2}{N_z}$$

HGLE (High Gray Level Emphasis): Measures the distribution of the higher gray-level values, with a higher value indicating a greater concentration of high gray-level values in the image.

$$SDLGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{P(i,j)}{i^2 j^2}}{N_z}$$

SDLGLE (Small Dependence Low Gray Level Emphasis): Measures the joint distribution of small dependence with lower gray-level values.

$$SDHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{P(i,j)i^2}{j^2}}{N_z}$$

SDHGLE (Small Dependence High Gray Level Emphasis): Measures the joint distribution of small dependence with higher gray-level values.

$$LDLGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{P(i,j)j^2}{i^2}}{N_z}$$

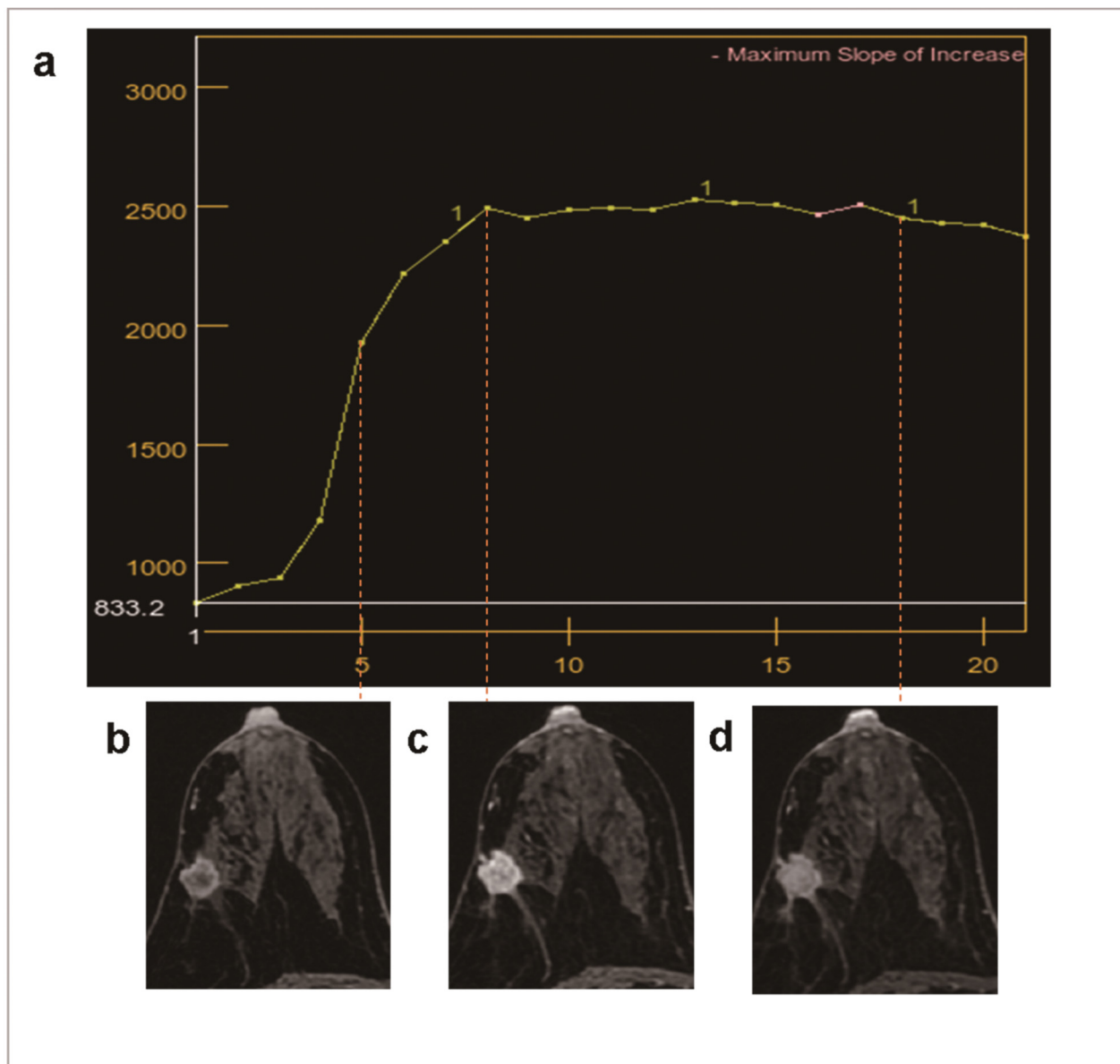
LDLGLE (Large Dependence Low Gray Level Emphasis): Measures the joint distribution of large dependence with lower gray-level values.

$$LDHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} P(i,j)i^2 j^2}{N_z}$$

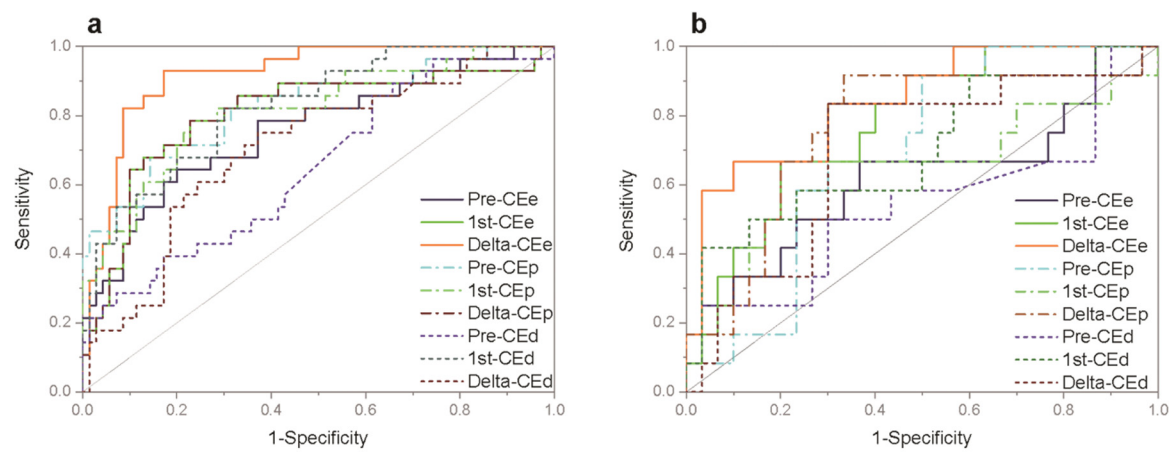
LDHGLE (Large Dependence High Gray Level Emphasis): Measures the joint distribution of large dependence with higher gray-level values.

| | Pre_radiomics features based on CEe (n=9/n=3) | Pre_radiomics features based on CEp (n=10/n=5) | Pre_radiomics features based on CEd (n=2/n=1) |
|------------------------------------|---|---|--|
| Features of LASSO_CV | [1] Original_glcml_ClusterShade [2] Wavelet_HHH_glszm_LALGLE [3] Wavelet_HLH_glrml_GLNUN [4] Wavelet_HLH_glszm_LAHGLE [5] Wavelet.HLL_firstorder_Skewness [6] Wavelet.HLL_glcml_ClusterShade [7]Wavelet.HLL_glcml_JointEnergy [8]Wavelet.LHH_glcml_ClusterShade [9] Wavelet.LHL_gldm_SDLGLE | [1] Wavelet_HLH_glszm_LAHGLE [2] Wavelet_HLH_glszm_LALGLE [3] Wavele_HLH_glszm_ZoneVariance [4] Wavelet_HLH_ngtdm_Busyness [5] Wavelet.HLL_firstorder_Skewness [6] Wavelet.HLL_glcml_ClusterShade [7] Wavelet.HLL_glcml_JointEnergy [8] Wavelet_HLL_glszm_LALGLE [9] Wavelet_HLL_glszm_ZoneVariance [10] wavelet_LHH_glszm_GLV | [1] Wavelet_HLL_gldm_SDLGLE [2] Wavelet_LHL_glcml_ClusterShade |
| Features of Logistic regression_CV | [1] Original_glcml_ClusterShade [2] Wavelet_HLH_glrml_GLNUN" [3] Wavelet.HLL_glcml_ClusterShade | [1] Wavelet_HLH_ngtdm_Busyness [2] Wavelet.HLL_firstorder_Skewness [3] Wavelet.HLL_glcml_JointEnergy [4] Wavelet_HLL_glszm_LALGLE [5] Wavelet_LHH_glszm_GLV | [1] Wavelet_LHL_glcml_ClusterShade |
| | 1st_radiomics features based on CEe (n=10/n=5) | 1st_radiomics features based on CEp (n=9/n=4) | 1st_radiomics features based on CEd (n=13/n=4) |
| Features of LASSO_CV | [1] Original_gldm_LDLGLE [2]Original_shape_SurfaceVolumeRat-io [3] Wavelet_HHH_firstorder_Median [4] Wavelet.HHL_gldm_GLNU [5] Wavelet_HHL_glszm_LAHGLE [6] Wavelet_HLH_glszm_GLNUN [7] Wavelet_HLH_ngtdm_Contrast [8] Wavelet_HLL_firstorder_Kurtosis [9] Wavelet_HLL_glszm_LALGLE [10] Wavelet_LLH_glszm_ZoneEntro-py | [1]Original_firstorder_Mean [2]Original_ngtdm_Complexity [3]Wavelet_HHH_glcml_Imcl [4]Wavelet_HHL_glszm_LAE [5]Wavelet.HHL_glszm_ZoneEntropy [6]Wavelet_HLH_glcml_MaximumProba-bility [7] Wavelet_HLL_firstorder_Kurtosis [8] Wavelet_HLL_glszm_LALGLE [9] Wavelet.LLL_firstorder_Maximum | [1] Original_firstorder_Maximum [2] Original_firstorder_Skewness [3] Original_gldm_LDLGLE [4] Wavelet_HHH_glszm_LALGLE [5] Wavelet_HHL_gldm_LDHGLE [6] Wavelet_HLH_glszm_ZoneEntropy [7] Wavelet_HLL_firstorder_Kurtosis [8] Wavelet_HLL_firstorder_Mean [9] Wavelet_HLL_glcml_Idmn [10] Wavelet_HLL_glszm_LAHGLE [11] Wavelet_LLH_firstorder_Uniformit-y [12] Wavelet.LLL_firstorder_90Percenti-le [13] Wavelet.LLL_glrml_LRE |
| Features of Logistic regression_CV | [1] Original_gldm_LDLGLE [2] Wavelet_HHH_firstorder_Median [3] Wavelet_HLH_glszm_GLNUN [4] Wavelet_HLH_ngtdm_Contrast [5] Wavelet_HLL_firstorder_Kurtosis | [1] Original_firstorder_Mean [2] Wavelet_HHH_glcml_Imcl [3] Wavelet_HHL_glszm_LAE [4] Wavelet_HLL_firstorder_Kurtosis | [1] Wavelet_HHL_gldm_LDHGLE [2] Wavelet_HLH_glszm_ZoneEntropy [3] Wavelet_LLH_firstorder_Uniformity [4] Wavelet.LLL_firstorder_90Percentile |
| | Delta_radiomics features based on CEe (n=13/n=9) | Delta_radiomics features based on CEp (n=12/n=7) | Delta_radiomics features based on CEd (n=3/n=1) |
| Features of LASSO_CV | [1] Original_glcml_Correlation [2] Original_gldm_DependenceEntro-py [3] Original_gldm_LDLGLE [4] Original_glrml_SRLGLE [5] Wavelet_HHH_glcml_Idmn [6] Wavelet_HHH_glszm_LALGLE [7] Wavelet_HLH_glcml_ClusterProminence [8] Wavelet_HLH_gldm_DependenceEntropy [9] Wavelet.HLH_gldm_LDHGLE [10] Wavelet.LHH_gldm_LDHGLE [11] Wavelet.LLH_firstorder_Mean [12] Wavelet.LLH_glcml_Idn [13] Wavelet.LLH_glrml_GLNUN | [1] wavelet_HHH_glcml_Idn [2] Wavelet.HHH_gldm_LDHGLE [3] Wavelet_HLH_ngtdm_Contrast [4] Wavelet_HLH_glszm_ZoneEntropy [5] Wavelet_LHL_glcml_JointAverage [6] Wavelet_LHL_glcml_MCC [7] Wavelet_LLH_firstorder_Range [8] Wavelet_LLH_glcml_DifferenceVari- ance [9] Wavelet_LLH_glszm_GLNUN [10] Wavelet_LLL_glcml_ClusterPromin- ence [11] Wavelet_LLL_glcml_ClusterShade [1 2]" wavelet.LLL_glszm_GLNU | [1] Wavelet_LHH_glszm_ZoneEntropy [2] Wavelet_LLH_firstorder_Maximum [3] Wavelet_LLH_glcml_JointAverage |
| Features of Logistic regression_CV | [1] Original_glcml_Correlation [2] Wavelet_HHH_glcml_Idmn [3] Wavelet_HHH_glszm_LALGLE [4] Wavelet_HLH_glcml_ClusterProminence [5] Wavelet.HLH_gldm_LDHGLE [6] Wavelet.LHH_gldm_LDHGLE [7] Wavelet.LLH_firstorder_Mean [8] Wavelet.LLH_glcml_Idn [9] Wavelet.LLH_glrml_GLNUN | [1] Wavelet.HHH_gldm_LDHGLE [2] Wavelet_HLH_ngtdm_Contrast [3] Wavelet_LHL_glcml_MCC [4] Wavelet_LLH_firstorder_Range [5] Wavelet_LLH_glszm_GLNUN [6] Wavelet_LLL_glcml_ClusterPromine- nce [7] Wavelet.LLL_glszm_GLNU | [1] Wavelet_LLH_firstorder_Maximum |

2 Supplementary Figure



Supplementary Figure S1. The process of differential subsampling with Cartesian ordering dynamic contrast-enhanced (DISCO) DCE-MRI in this study. (a) Images were obtained from a total of 1+20 phases (one pre-contrast phase and 20 post-contrast enhanced phases). (b) The post-contrast early-phase (CEe) scan was started approximately 78 s after the injection of contrast agent. (c) The peak-phase (CEp) scan began 136 s after the injection. (d) The delay phase began approximately 330 s after the injection.



Supplementary Figure S2. ROC curves of the separate radiomic models in training cohort (a) and validation cohort (b). CEe, the early phase of dynamic contrast-enhanced (DCE) MRI; CEp, the peak phase of DCE-MRI; CED, the delay phase of DCE-MRI; Pre, pre-radiomics features; 1st, 1st-radiomics features; Delta, delta-radiomics features.