Table S1 Description of classical metabolic parameters and radiomics features mentioned in this study

Feature	Algorithm	Description
Classical metabolic parameters		
SUV	SUV is defined as: $SUV = \frac{Mean\ radioactivity\ of\ region\ of\ Interest(ROI)(MBq/mL)}{Injection\ activity(MBq)/weight(g)}$	Index of radiation uptake of ROI.
SUV_{max}		Highest SUV of the voxels within the ROI.
SUV_{mean}		Mean SUV of the voxels within the ROI.
SUV_{peak}		Average SUV of the voxels within a fixed ROI of the highest metabolic pixel.
MTV	MTV is defined as: $MTV = 40\%\ SUV_{max} \times Volume(ROI)$	The sum of metabolically active volumes in the lesion.
TLP	TLP is defined as: $TLP = MTV \times SUV_{mean}$	TLP as an equivalent to total lesion glycolysis in PSMA PET, which not only reflects the tumor volume, but also reflects the metabolic uptake of the lesion.
SUV_{max}	SUV_{max} is defined as: $SUV_{max} = \frac{SUVmax\ of\ lesion}{SUVmax\ of\ background}$	The normalized SUV_{max} to remove the effect of background uptake.
SUV_{mean}	SUV_{mean} is defined as: $SUV_{mean} = \frac{SUVmean\ of\ lesion}{SUVmean\ of\ background}$	The normalized SUV_{mean} to remove the effect of background uptake.
Shape-based features		
$Sphericity$	$Sphericity$ is defined as: $sphericity = \frac{\sqrt[3]{36\pi V^2}}{A}$	$Sphericity$ is a measure of the roundness of the shape of the tumor region relative to a sphere.

Surface area

Surface area is defined as:

$$A_i = \frac{1}{2}\left|a_i b_i \times a_i c_i\right| \quad (1)$$

$$A = \sum_{i=1}^{N_f} A_i \quad (2)$$

To calculate the surface area, first the surface area A_i of each triangle in the mesh is calculated (1). The total surface area is then obtained by taking the sum of all calculated sub-areas (2).

Elongation

Elongation is defined as:

$$elongation = \sqrt{\frac{\lambda_{minor}}{\lambda_{major}}}$$

Elongation shows the relationship between the two largest principal components in the ROI shape.

Flatness

Flatness is defined as:

$$flatness = \sqrt{\frac{\lambda_{least}}{\lambda_{major}}}$$

Flatness shows the relationship between the largest and smallest principal components in the ROI shape.

First-order features

Variance

Variance is defined as:

$$variance = \frac{1}{N_p}\sum_{i=1}^{N_p}\left(\mathbf{X}(i) - \bar{X}\right)^2$$

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.

Skewness

Skewness is defined as:

$$skewness = \frac{\mu_3}{\sigma^3} = \frac{\frac{1}{N_p}\sum_{i=1}^{N_p}(\mathbf{X}(i) - \bar{X})^3}{\left(\sqrt{\frac{1}{N_p}\sum_{i=1}^{N_p}(\mathbf{X}(i) - \bar{X})^2}\right)^3}$$

Skewness measures the asymmetry 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.

Kurtosis

Kurtosis is defined as:

$$kurtosis = \frac{\mu_4}{\sigma^4} = \frac{\frac{1}{N_p}\sum_{i=1}^{N_p}(\mathbf{X}(i) - \bar{X})^4}{\left(\frac{1}{N_p}\sum_{i=1}^{N_p}(\mathbf{X}(i) - \bar{X})^2\right)^2}$$

Kurtosis is a measure of the 'peakedness' 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.
Entropy	*Entropy* is defined as: $$entropy = -\sum_{i=1}^{N_g} p(i)\log_2\left(p(i)+\ \right)$$	*Entropy* specifies the uncertainty/randomness in the image values. It measures the average amount of information required to encode the image values.
Uniformity	*Uniformity* is defined as: $$uniformity = \sum_{i=1}^{N_g} p(i)^2$$	*Uniformity* is a measure of the sum of the squares of each intensity value. This is a measure of the homogeneity of the image array.
Total energy	*Total energy* is defined as: $$total\ energy = V_{voxel}\sum_{i=1}^{N_p}\left(\mathbf{X}(i)+c\right)^2$$	Energy is a measure of the magnitude of voxel values in an image. A larger values implies a greater sum of the squares of these values.
High-order features		
Busyness	*Busyness* is defined as: $$Busyness = \frac{\sum_{i=1}^{N_g} p_i s_i}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\left\lvert ip_i - jp_j\right\rvert}, where\ p_i \neq 0, p_j \neq 0$$	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.
Coarseness	*Coarseness* is defined as: $$Coarseness = \frac{1}{\sum_{i=1}^{N_g} p_i s_i}$$	*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.
Complexity	*Complexity* is defined as: $$Complexity = \frac{1}{N_{v,p}}\sum_{i=i}^{N_g}\sum_{j=1}^{N_g}\lvert i-j\rvert\frac{p_i s_i + p_j s_j}{p_i + p_j}, where\ p_i \neq 0, p_j \neq 0$$	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.
Contrast	*Contrast* is defined as: $Contrast = \left(\frac{1}{N_{g,p}(N_{g,p}-1)} \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_i p_j (i-j)^2 \right) \left(\frac{1}{N_{v,p}} \sum_{i=1}^{N_g} s_i \right), where\ p_i \neq 0, p_j \neq 0$	*Contrast* is a measure of the local intensity variation, favoring values away from the diagonal.
Strength	*Strength* is defined as: $Strength = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (p_i + p_j)(i-j)^2}{\sum_{i=1}^{N_g} s_i}, where\ p_i \neq 0, p_j \neq 0$	*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.

The definitions and description of radiomics features mentioned above were derived from https://pyradiomics.readthedocs.io/en/latest/features.html