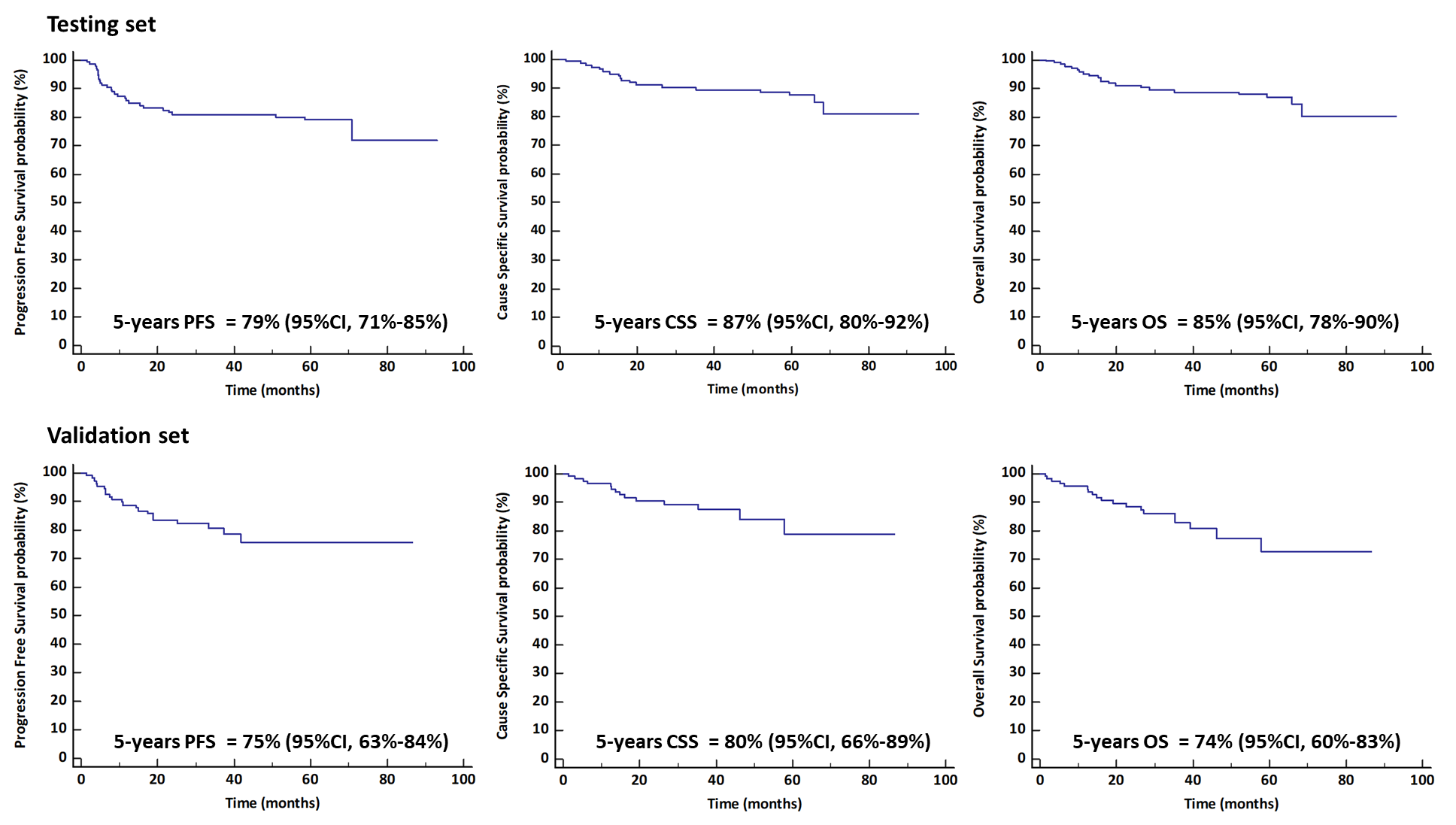
**Supplementary material**

***Treatment outcomes in the entire patient population of each study cohort.***

**Figure S1: Kaplan Meier estimates of progression free survival (PFS), cause-specific survival (CSS) and overall survival (OS), in the testing (upper row) and validation cohort (lower row), respectively.**



***Extraction of radiomics features***

For the extraction of radiomics features from PET/CT imaging, we used the PyRadiomics platform, a software package compliant with the image biomarker standardization initiative (IBSI). PyRadiomics, a flexible open-source platform implemented in the Python programming language and capable of extracting a large panel of engineered features from medical images. These features are subdivided into the following classes:

* First Order Statistics (19 features)
* Shape-based (3D) (16 features)
* Shape-based (2D) (10 features)
* Gray Level Cooccurence Matrix (GLCM, 24 features)
* Gray Level Run Length Matrix (GLRLM, 16 features)
* Gray Level Size Zone Matrix (GLSZM, 16 features)
* Neighbouring Gray Tone Difference Matrix (NGTDM, 5 features)
* Gray Level Dependence Matrix (GLDM, 14 features)

All feature classes, with the exception of shape can be calculated on either the original image and/or a derived image, obtained by applying one of several filters. The shape descriptors are independent of gray value and depict the geometric properties of the lesion. The first-order statistics features describe the presence of areas with statistically different metabolic activity without taking into account their distribution within the lesion, while the matrix-based features analyze the spatial distribution of these areas and their mutual relationship. Table S1 summarizes describe the default parameters we used to extract radiomics features. Table S2 enumerates the extracted features; those included in the LASSO Cox regression are in bold type. Their mathematical description is available in the PyRadiomics online documentation [https://pyradiomics.readthedocs.io/en/latest/features.html].

In this work, the 3D contoured mask, generated after the segmentation, is a Boolean mask including all the lesions with volume > 5ml. The voxels belonging to the lesions are set to True, while all the other voxels are set to False. Further calculation is done using only those voxels from the image that are marked as True in the mask, regardless of where they are. This approach allows to analyze all the voxels as a part of a single volume.

Moreover, it contains all the spatial coordinates of the original reconstructed images allowing the correct localization in the body of each voxel and the measurement of the single volumes and of the distances between lesions.

**Table S1:** **Parameters used to extract radiomics features**

|  |  |
| --- | --- |
| **Calculation setting** | **Default** |
| Approach | 3D |
| Interpolation   * Voxel dimension (mm) * Interpolation method * ROI interpolation method | Yes   * 4x4x4 * B-Spline * Nearest-Neighbor |
| Resegmentation   * Range (SUV) * Outliers | * No * No |
| Discretization | Fixed bin size: 64 bins |

**Table S2: List of extracted radiomics features.**

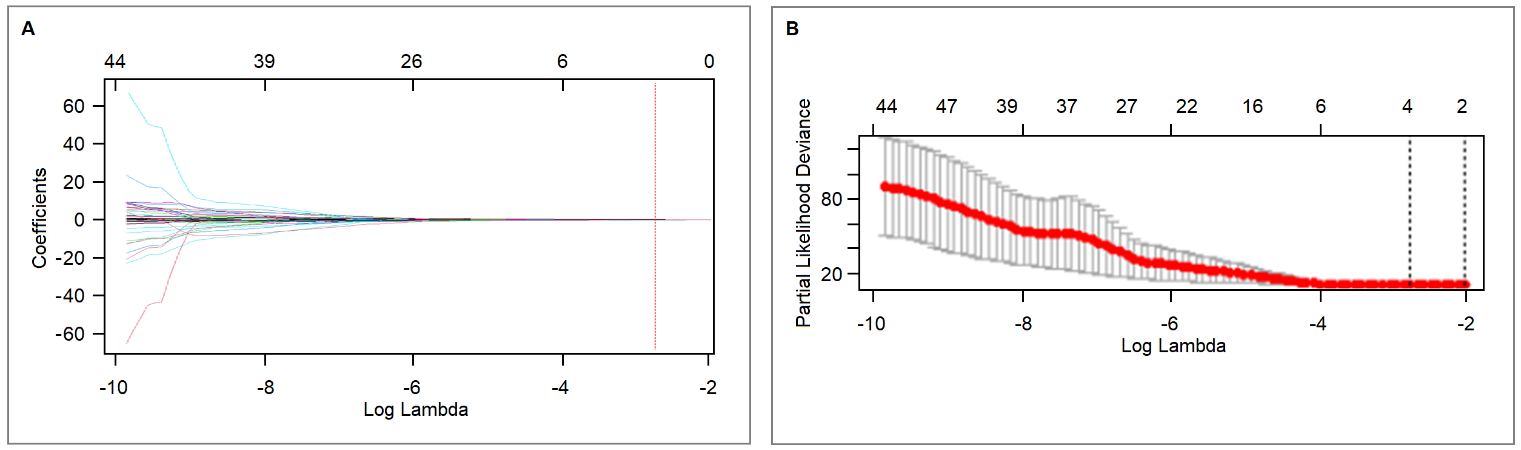
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **First-order features** | | | | |
| *Histogram based features* | | | | |
| **10Percentile\***  90Percentile  Energy  **Entropy\***  InterquartileRange  **Kurtosis\***  **Maximum\***  MeanAbsoluteDeviation  Mean  **Median\***  **Minimum\***  Range  RobustMeanAbsoluteDeviation  RootMeanSquared  **Skewness\***  TotalEnergy  **Uniformity\***  Variance | | | | |
| **Shape-based features** | | | | |
| *3D shape features* | | | | |
| **Elongation\***  **Flatness\***  **LeastAxisLength\***  **MajorAxisLength\***  **Maximum2DDiameterColumn\***  **Maximum2DDiameterRow\***  Maximum2DDiameterSlice  **Maximum3DDiameter\***  MeshVolume  MinorAxisLength  **Sphericity\***  SurfaceArea  SurfaceVolumeRatio  **VoxelVolume\***  **Maximum3DDiameter/BSA\*** | | | | |
| **Second-order (texture) features** | | | | |
| *GLCM* | *GLDM* | *GLRLM* | *GLSZM* | *NGTDM* |
| **Autocorrelation\***  **ClusterProminence\***  **ClusterShade\***  **ClusterTendency\***  Contrast  Correlation  DifferenceAverage  DifferenceEntropy  DifferenceVariance  Id  Idm  Idmn  Idn  **Imc1\***  **Imc2\***  InverseVariance  **JointAverage\***  **JointEnergy\***  **JointEntropy\***  **MCC\***  **MaximumProbability\***  **SumAverage\***  **SumEntropy\***  **SumSquares\*** | DependenceEntropy  DependenceNonUniformity  DependenceNonUniformityNormalized  DependenceVariance  GrayLevelNonUniformity  **GrayLevelVariance\***  **HighGrayLevelEmphasis\***  LargeDependenceEmphasis  **LargeDependenceHighGrayLevelEmphasis\***  LargeDependenceLowGrayLevelEmphasis  **LowGrayLevelEmphasis\***  SmallDependenceEmphasis  **SmallDependenceHighGrayLevelEmphasis\***  SmallDependenceLowGrayLevelEmphasis | GrayLevelNonUniformity  **GrayLevelNonUniformityNormalized\***  **GrayLevelVariance\***  **HighGrayLevelRunEmphasis\***  LongRunEmphasis  **LongRunHighGrayLevelEmphasis\***  **LongRunLowGrayLevelEmphasis\***  LowGrayLevelRunEmphasis  **RunEntropy\***  RunLengthNonUniformity  RunLengthNonUniformityNormalized  RunPercentage  RunVariance  ShortRunEmphasis  **ShortRunHighGrayLevelEmphasis\***  **ShortRunLowGrayLevelEmphasis\*** | GrayLevelNonUniformity  **GrayLevelNonUniformityNormalized\***  **GrayLevelVariance\***  **HighGrayLevelZoneEmphasis\***  LargeAreaEmphasis  LargeAreaHighGrayLevelEmphasis  **LargeAreaLowGrayLevelEmphasis\***  **LowGrayLevelZoneEmphasis\***  SizeZoneNonUniformity  SizeZoneNonUniformityNormalized  SmallAreaEmphasis  **SmallAreaHighGrayLevelEmphasis\***  **SmallAreaLowGrayLevelEmphasis\***  ZoneEntropy  ZonePercentage  ZoneVariance | Busyness  Coarseness  Complexity  Contrast  Strength |

**\***Bold type indicates RFs uncorrelated with tumor volume (MTV) and its uptake (SUVmax), which were included in LASSO Cox regression; they comprise 8 first-order features, 10 3D shape features, and 34 second-order matrix features.

***Lasso Cox regression***

The determination of the optimized regularization parameter λ, which ensured that the model had the least deviance improving both the prediction accuracy and the interpretability, was selected via the minimum criteria i.e., the value of λ that provides the minimum mean cross-validated error. The optimal tuning of this parameter allows obtaining non-zero coefficients only for the more predictive features. The optimal values of the penalty parameter  were determined by ten-fold cross-validations.

**Figure S2:** **Feature selection using the LASSO Cox regression with 10-fold cross-validation.**



(A) In the LASSO Cox model, the minimum standard is adopted to obtain the value of the parameter λ by 10-fold cross-validation. The λ value was confirmed as 0.06. (B) Coefficient sectional view plotted against the log (λ) magnitude. Based on 10-fold cross-validation, the optimal λ corresponding to 4 non-zero coefficients (indicated by a vertical line in the plot) were selected.

***Selected features description [IBSI, pyradiomics documentation]***

[***https://pyradiomics.readthedocs.io/en/latest/features.html#***](https://pyradiomics.readthedocs.io/en/latest/features.html)

[***https://ibsi.readthedocs.io/en/latest/03\_Image\_features.html#***](https://ibsi.readthedocs.io/en/latest/03_Image_features.html)

*GLCM\_SumSquares*

The grey level co-occurrence matrix (GLCM) is a matrix that expresses how combinations of discretized intensities (grey levels) of neighboring voxels are distributed along one of the image directions. By default, the value of a feature is calculated on the GLCM for each angle separately, after which the mean of these values is returned. Sum of Squares or Variance is a measure in the distribution of neighboring intensity level pairs about the mean intensity level in the GLCM. In IBSI is called “Joint Variance”.

*Shape\_Maximum3DDiameter/BSA*

Maximum 3D diameter is defined as the largest pairwise Euclidean distance between tumor surface mesh vertices. In case of multiple lesions, it represents the maximum extension of the disease. The original variable has been normalized by Body Surface Area in order to take into account the patient’s body size as already suggested in literature (Cottereau et al, *Ann Oncol.* 2021; *manuscript reference #28).*

*GLDM\_GrayLevelVariance*

The Gray Level Dependence Matrix (GLDM) aims to capture the coarseness of the overall texture and is rotationally invariant. A gray level dependency is defined as the number of connected voxels within distance certain distance that are dependent on the center voxel. GLDM\_GrayLevelVariance measures the variance in dependence size in the image.

*GLSZM\_GrayLevelNonUniformityNormalized*

The grey level size zone matrix (GLSZM) counts the number of groups (or zones) of linked voxels. Voxels are linked if the neighboring voxel has an identical discretized grey level. The GLSZM is rotation independent, with only one matrix calculated for all directions in the ROI. GLNN measures the variability of grey-level intensity values in the image, which a lower value indicating a greater similarity in intensity values.